МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БУДІВНИЦТВА І АРХІТЕКТУРИ

Кваліфікаційна наукова праця
на правах рукопису

Терейковська Людмила Олексіївна

УДК 004.93

ДИСЕРТАЦІЯ

МЕТОДОЛОГІЯ АВТОМАТИЗОВАНОГО РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО СТАНУ СЛУХАЧІВ СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ

Спеціальність 05.13.06 – Інформаційні технології

Подається на здобуття наукового ступеня доктора технічних наук

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Л.О. Терейковська

Науковий консультант: д. т. н., професор Бушуєв Сергій Дмитрович

Ідентичність усіх примірників дисертації
Засвідчую
Вчений секретар
Спеціалізованої вченої ради Д 26.059.01

_____ О.І. Мельниченко

Київ 2022

АНОТАЦІЯ

Терейковська Л. О. Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання. – Рукопис.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 «Інформаційні технології». – Київський національний університет будівництва і архітектури, МОН України. Київ, 2022.

Дисертація присвячена вирішенню важливої науково-прикладної проблеми, пов'язаної з розпізнаванням емоційного стану слухачів системи дистанційного навчання шляхом розробки відповідної методології, що базується на нейромережевих методах і моделях аналізу біометричних параметрів, та орієнтованої на створення відповідних засобів, які розширюють функціональні можливості системи дистанційного навчання. Визначено та вирішено основні наукові завдання дослідження. *Перше наукове завдання* – побудова методологічного забезпечення технології автоматизованого розпізнавання емоційного стану слухача СДН на основі нейромережевого аналізу біометричних параметрів, реєстрацію яких можливо здійснити за допомогою поширеного апаратно-програмного забезпечення. *Друге наукове завдання* – створення методів формування вхідного поля нейромережевих моделей аналізу біометричних параметрів для розпізнавання емоцій та особи слухача СДН. *Третє наукове завдання* – створення методів нейромережевого розпізнавання емоцій та особи слухача СДН, що враховують можливість використання навчальних прикладів, вимоги щодо відображення в очікуваному вихідному сигналі близькості еталонів класів та необхідність визначення архітектурних параметрів згорткових нейронних мереж. *Четверте наукове завдання* – побудова методології нейромережевого розпізнавання емоційного стану та особи слухача СДН, яка, інтегруючи запропоновані методи нейромережевого розпізнавання, забезпечує розширення функціональних можливостей СДН за рахунок розпізнавання емоційного стану слухача в процесі навчання та підтвердження достовірності його особи. *П'яте наукове завдання* – розробка

архітектури системи нейромережевого розпізнавання емоційного стану та особи слухача СДН. Основні результати дослідження: методологія автоматизованого розпізнавання емоційного стану, яка на відміну від існуючих за рахунок інтегрованого застосування відомих та запропонованих підходів, моделей та методів нейромережевого аналізу біометричних параметрів, забезпечує можливість створення ефективних інструментальних засобів розпізнавання емоційного стану; комплекс методів формування вхідного поля нейромережевих моделей аналізу біометричних параметрів, який передбачає використання процедур подання біометричних параметрів у вигляді, придатному до обробки згортковою нейронною мережею; метод розробки архітектури згорткової нейронної мережі, який дозволяє зменшити обсяг експериментів, пов'язаних з адаптацією значень конструктивних параметрів мережі до умов аналізу біометричних параметрів слухача СДН. Запропоновані нейромережеві моделі та методи розпізнавання емоційного стану слухача СДН використані при виконанні держбюджетних науково-дослідних робіт та впроваджені в навчальний процес Київського національного університету будівництва і архітектури та інших університетів.

Ключові слова: біометричний параметр, нейромережева модель, розпізнавання, система дистанційного навчання, емоційний стан.

ABSTRACT

Tereikovska L. O. Methodology of automated recognition of the emotional state of listeners of the distance learning system. – The manuscript.

Thesis for a Doctor of Science Degree in specialty 05.13.06 "Information Technologies". – Kyiv National University of Construction and Architecture, Kyiv, 2022.

The dissertation is devoted to the solution of an important scientific and applied problem related to the recognition of the emotional state of the listeners of the distance learning system by developing an appropriate methodology based on neural network methods and models for the analysis of biometric parameters, and aimed at creating appropriate tools that expand the functional capabilities of the distance learning system. The main scientific tasks of the research have been defined and solved. The first scientific task is the construction of methodological support for the technology of automated recognition of the emotional state of the listener of the DLS on the basis of neural network analysis of biometric parameters, the registration of which can be carried out with the help of common hardware and software. The second scientific task is the creation of methods for forming the input field of neural network models for the analysis of biometric parameters for recognizing the emotions and personality of a listener of the DLS. The third scientific task is the creation of methods of neural network recognition of emotions and the personality of the listener of the DLS, which take into account the possibility of using educational examples, the requirements for displaying in the expected output signal the proximity of class standards and the need to determine the architectural parameters of convolutional neural networks. The fourth scientific task is the construction of a methodology for neural network recognition of the emotional state and identity of the listener of the DLS, which, integrating the proposed methods of neural network recognition, ensures the expansion of the functional capabilities of the DLS due to the recognition of the emotional state of the listener in the process of learning and confirmation of the authenticity of his identity. The fifth scientific task is the development of the

architecture of the system of neural network recognition of the emotional state and personality of the DLS listener. The main results of the research: the methodology of automated emotional state recognition, which, unlike the existing ones due to the integrated application of known and proposed approaches, models and methods of neural network analysis of biometric parameters, provides the possibility of creating effective instrumental means of emotional state recognition; a set of methods for forming the input field of neural network models for the analysis of biometric parameters, which involves the use of procedures for presenting biometric parameters in a form suitable for processing by a convolutional neural network; a method of developing the architecture of a convolutional neural network, which allows to reduce the amount of experiments related to the adaptation of the values of the constructive parameters of the network to the conditions of the analysis of the biometric parameters of the DLS listener. The proposed neural network models and methods for recognizing the emotional state of the listener of the DLS were used in the implementation of state-budget scientific research works and were introduced into the educational process of the Kyiv National University of Construction and Architecture and other universities.

Keywords: biometric parameter, neural network model, recognition, distance learning system, emotional state.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Монографії

1. Міхайленко В. М., Терейковська Л. О., Терейковський І. А., Ахметов Б. Б. Нейромережеві моделі та методи розпізнавання фонем в голосовому сигналі в системі дистанційного навчання: монографія. Київ: ЦП «Компринт», 2017. 252 с. ISBN 978-966-929-430-2.

Автору належить розробка підходів до формування методологічного базису для нейромережевого аналізу біометричних параметрів слухача системи дистанційного навчання.

Публікації в міжнародних виданнях

1. Tereykovska L., Tereykovskiy I., Aytkhozhaeva E., Tynymbayev S., Imanbayev A. Encoding of neural network model exit signal, that is devoted for distinction of graphical images in biometric authenticate systems. *News of the national academy of sciences of the republic of Kazakhstan series of geology and technical sciences*. 2017. Vol. 6. No 426. P. 217-224.

Scopus, зарубіжне фахове видання.

Автору належить розробка та опис експериментальної верифікації моделі нейромережевого кодування значення очікуваного вихідного сигналу навчальних прикладів.

2. Bapiyev I. M., Aitchanov B. H., Tereikovskyi I. A., Tereikovska L. A., Korchenko A. A. Deep neural networks in cyber attack detection systems. *International Journal of Civil Engineering and Technology (IJCIET)*. November 2017. Vol. 8. Issue 11. P. 1086-1092.

Scopus, зарубіжне фахове видання.

Автору належить розробка критеріїв ефективності застосування глибокої нейронної мережі в очікуваних умовах використання.

3. Zhengbing H., Tereykovskiy I., Tereykovska L., Pogorelov V. Determination of structural parameters of multilayer perceptron designed to estimate parameters of technical systems. *Intelligent Systems and Applications*. 2017. No 10. P. 57-62.

Scopus, зарубіжне фахове видання.

Автору належить розробка підходів до визначення структурних параметрів багатошарового персептрону при його застосуванні для оцінки параметрів технічних систем.

4. Tereikovskyi I., Mussiraliyeva S., Kosyuk Y., Bolatbek M., Tereikovska L. An experimental investigation of infrasound influence hard drives of a computer system. *International Journal of Civil Engineering and Technology*. June 2018. Vol. 9. Issue 6. P. 1558-1566.

Scopus, зарубіжне фахове видання.

Автору належить аналіз моделей реалізації та опис експериментальних досліджень, пов'язаних з оцінкою значущості звукового випромінювання.

5. Akhmetov B., Tereykovsky I., Doszhanova A., Tereykovskaya L. Determination of input parameters of the neural network model, intended for phoneme recognition of a voice signal in the systems of distance learning. *International Journal of Electronics and Telecommunications*. 2018. Vol. 64. No 4. P. 425-432.

Scopus, зарубіжне фахове видання.

Автору належить розробка моделі та опис практичного прикладу представлення нестаціонарного оцифрованого голосового сигналу за допомогою фіксованої кількості мел-кепстральних коефіцієнтів.

6. Tereikovska L., Tereikovskyi I., Beketova A., Karaman G., Makovetska N. Recognition of Speaker's Emotion by Squeezenet Convolutional Neural Network. *Journal of Theoretical and Applied Information Technology*. March 2021. Vol. 99. No 5. P. 1139-1148.

Scopus, зарубіжне фахове видання.

Автору належить розробка згорткової нейромережевої моделі та опис її застосування для розпізнавання емоцій диктора.

7. Hu Z., Tereikovskyi I., Chernyshev D., Tereikovska L., Tereikovskyi O., Wang D. Procedure for Processing Biometric Parameters Based on Wavelet

Transformations. *International Journal of Modern Education and Computer Science(IJMECS)*. 2021. Vol. 13. No 2. P. 11-22.

Scopus, зарубіжне фахове видання.

Автору належить розробка технології застосування вайвлет-перетворень для фільтрації біометричних параметрів, асоційованих з зображеннями та опис практичного прикладу такого застосування.

8. Tereikovskyi I., Chernyshev D., Tereikovska L., Mussiraliyeva S., Akhmed G. The procedure for the determination of structural parameters of a convolutional neural network to fingerprint recognition. *Journal of Theoretical and Applied Information Technology*. April 2019. Vol. 97. No 8. P. 2381-2392.

Scopus, зарубіжне фахове видання.

Автору належить розробка методологічної бази, механізму визначення структурних параметрів згорткової нейронної мережі та опис практичного прикладу застосування цієї моделі для розпізнавання відбитків пальців.

9. Aitkhozhayeva Y., Tynymbayev S., Seilova N., Tereikovska L., Imanbayev A. Method and device for modulus reduction. *Bulletin of the National Academy of Sciences of the Republic of Kazakhstan*. 2019. Vol. 2. No 378. P. 220-225.

Web of Science, зарубіжне фахове видання.

Автору належить розробка методологічного забезпечення для оцінки ефективності застосування технічних систем.

10. Hu Z., Tereykovskiy I., Zorin Y., Tereykovska L., Zhibek A. Optimization of convolutional neural network structure for biometric authentication by face geometry. *Advances in Intelligent Systems and Computing*. 2019. Vol. 754. P. 567-577.

Scopus, зарубіжне фахове видання, опубліковане по матеріалам конференції.

Автору належить розробка та верифікація нейромережевої моделі для біометричної аутентифікації за зображенням обличчя.

11. Hu Z., Tereikovskyi I., Tereikovska L., Tsiutsiura M., Radchenko K.

Applying Wavelet Transforms for Web Server Load Forecasting. *Advances in Intelligent Systems and Computing*. 2020. Springer, Cham. Vol 938. P. 13-22.

Scopus, зарубіжне фахове видання, опубліковане по матеріалам конференції.

Автору належить визначення параметрів ефективності та механізму вибору базисного типу вейвлету.

12. Tereikovskyi I., Tereikovska L., Korystin O., Mussiraliyeva S., Sambetbayeva A. User Keystroke Authentication and Recognition of Emotions Based on Convolutional Neural Network. *Advances in Intelligent Systems and Computing*. 2020. Springer, Cham. Vol. 1126. P. 283-292.

Scopus, зарубіжне фахове видання, опубліковане по матеріалам конференції.

Автору належить розробка та верифікація нейромережевої моделі для розпізнавання особи та емоцій людини за клавіатурним почерком.

13. Tereikovskyi I., Tereykovska L., Mussiraliyeva S., Tsiutsiura M., Achkoski J. Markov Model of Unsteady Profile of Normal Behavior of Network Objects of Computer Systems. *Conflict Management in Global Information Networks*. 2019. CEUR Workshop Proceedings. Vol. 2588. P. 140-152.

Scopus, зарубіжне фахове видання, опубліковане по матеріалам конференції.

Автору належить аналіз моделей та методів оцінки динаміки параметрів мережевих об'єктів.

14. Hu Z., Tereikovskyi I., Korystin O., Mihaylenko V., Tereikovska L. Two-Layer Perceptron for Voice Recognition of Speaker's Identity. *Advances in Intelligent Systems and Computing*. 2021. Springer, Cham. Vol. 1247. P. 508-517.

Scopus, зарубіжне фахове видання, опубліковане по матеріалам конференції.

Автору належить розробка та верифікація нейромережевої моделі на базі двошарового персептрону для ідентифікації диктора.

15. Akhmetov B., Tereikovskyi I., Tereikovska L., Adranova A. Neural

Network User Authentication by Geometry of the Auricle. *Advances in Intelligent Systems and Computing*. 2018. Springer, Cham. Vol. 836. P. 11-19.

Scopus, зарубіжне фахове видання, опубліковане по матеріалам конференції.

Автору належить розробка та верифікація нейромережевої моделі для розпізнавання особи за геометрією вушної раковини.

16. Oksiiuk O., Tereikovska L., Tereikovskiy I. Determination of expected output signals of the neural network model intended for image recognition. *4th International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T).* 2017. P. 596-599.

Scopus, Web of Science

Автору належить розробка та верифікація моделі визначення вихідного сигналу нейронної мережі для розпізнавання зображень.

17. Oksiiuk O., Tereikovska L., Tereikovskiy I. Adaptation of the neural network model to the identification of the cyberattacks type "denial of service". *14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET)*. 2018. P. 502-505.

Scopus, Web of Science

Автору належить розробка методологічного забезпечення визначення значень конструктивних параметрів нейромережевої моделі.

18. Toliupa S., Tereikovskiy I., Dychka I., Tereikovska L., Trush A. The Method of Using Production Rules in Neural Network Recognition of Emotions by Facial Geometry. *3rd International Conference on Advanced Information and Communications Technologies (AICT).* 2019. P. 323-327.

Scopus

Автору належить розробка та верифікація методу використання продукційних правил в нейромережевій моделі для розпізнавання емоцій за зображенням обличчя.

19. Toliupa S., Tereikovska L., Korystin O., Chernyshev D., Tereikovskyi I. Low-Resource Convolution Neural Network for Keyboard Recognition of the

User. *IEEE International Conference on Advanced Trends in Information Theory (ATIT)*. 2019. P. 222-226.

Scopus

Автору належить розробка та верифікація малоресурсної згорткової нейронної мережі для аналізу параметрів клавіатурного почерку.

20. Toliupa S., Nakonechnyi V., Tereikovskyi I., Tereikovska L., Korystin O. One-Periodic Template Marks Model of Normal Behavior of the Safety Parameters of Information Systems Networking Resources. *IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T)*. 2019. P. 764-768.

Scopus

Автору належить аналіз моделей та методів оцінки динаміки параметрів мережевих ресурсів інформаційних систем.

21. Toliupa S., Kulakov Y., Tereikovskyi I., Tereikovskyi O., Tereikovska L., Nakonechnyi V. Keyboard Dynamic Analysis by Alexnet Type Neural Network. *IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*. 2020. P. 416-420.

Scopus

Автору належить розробка та верифікація згорткової нейронної мережі типу Alexnet для аналізу динаміки клавіатурного почерку.

22. Yudin O., Toliupa S., Korchenko O., Tereikovska L., Tereikovskyi I., Tereikovskyi O. Determination of Signs of Information and Psychological Influence in the Tone of Sound Sequences. *IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*. 2020. P. 276-280.

Scopus

Автору належить розробка методологічного забезпечення для визначення ознак інформаційно-психологічного впливу в тональності звукових рядів.

23. Toliupa S., Tereikovska L., Tereikovskyi I., Doszhanova A., Alimseitova Z. Procedure for Adapting a Neural Network to Eye Iris Recognition.

IEEE International Conference on Problems of Infocommunications, Science and Technology (PIC S&T). 2020. P. 167-171.

Scopus

Автору належить розробка та верифікація моделі адаптації параметрів нейронної мережі до розпізнавання райдужної оболонки ока.

24. Tereikovska L., Tereikovskyi I., Mussiraliyeva S., Akhmed G. Recognition of emotions by facial Geometry using a capsule neural network. *International Journal of Civil Engineering and Technology (IJCIET)*. April 2019.Vol. 10. Issue 04. P. 270-279. ISSN: 0976-6308.

Автору належить розробка та верифікація капсульної нейромережевої моделі для розпізнавання емоцій за зображенням обличчя.

25. Tereykovskaya L., Tereykovskiy I., Aytkhozhaeva E., Tynymbayev S., Imanbayev A. Improvement of learning efficiency of the neural networks, intended for recognition of graphic images in systems of biometric authentication. *Computer Modelling & New Technologies*. 2017. 21(2). P. 54-57. ISSN: 1407-5806.

Автору належить розробка та верифікація механізму застосування ймовірнісної нейронної мережі для кодування очікуваного вихідного сигналу.

Публікації у фахових виданнях

1. Терейковский И., Корченко А., Ахметов Б., Бапиев И., Терейковська Л. Определение эффективных видов нейросетевых моделей распознавания кибератак на сетевые ресурсы. *Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні*. 2016. Вип. 2(32). С.56-63.

Автору належить аналіз підходів та розробка методологічного забезпечення для визначення ефективних видів нейромережевих моделей.

2. Терейковський І. А., Терейковська Л. О. Метод кодування вихідного сигналу нейромережевих моделей. *Наукові вісті НТУУ «КПІ»: міжнародний науково-технічний журнал*. 2017. № 5(115). С. 59–64.

Автору належить розробка та верифікація методу кодування вихідного сигналу нейромережевих моделей.

3. Михайленко В.М., Терейковская Л.А., Легеза В.П. Адаптация параметров двухслойного персептрона, предназначенного для биометрической аутентификации пользователей по отпечаткам пальцев. *Управління розвитком складних систем*. 2018. №34. С.130-138.

Автору належить розробка та верифікація методу адаптації параметрів двошарового персептрону для біометричної аутентифікації за відбитками пальців.

4. Терейковський І.А., Заріцький О.В., Терейковська Л.О., Погорелов В.В. Метод розробки архітектури глибокої нейронної мережі, призначеної для розпізнавання комп'ютерних вірусів. *Захист інформації*. 2018. Том 20. № 3. С. 188-199.

Автору належить розробка методологічного забезпечення методу визначення архітектурних параметрів глибокої нейронної мережі.

5. Терейковська Л.О. Аналіз засобів розпізнавання емоційного стану людини за голосом. *Вчені записки Таврійського національного університету імені В.І.Вернадського, серія «Технічні науки»*. 2019. Том 30 (69) Ч. 1. №1. С.120-124.

6. Михайленко В.М., Терейковская Л.А. Обзор средств распознавания эмоционального состояния человека по геометрии лица. *Управління розвитком складних систем*. 2019. № 37. С. 178-184.

Автору належить аналіз засобів автоматизованого розпізнавання емоційного стану за зображенням обличчя, розробка переліку параметрів та оцінка ефективності засобів розпізнавання.

7. Терейковська Л.О. Нейромережева модель розпізнавання емоційного стану операторів автоматизованих робочих місць за клавіатурним почерком. *Вчені записки Таврійського національного університету імені В.І. Вернадського, серія «Технічні науки»*. 2019. Том 30 (69). Ч. 1. №4. С. 129-133.

8. Михайленко В.М., Терейковская Л.А. Концептуальная модель нейросетевого распознавания эмоционального состояния слушателей системы дистанционного обучения. *Управління розвитком складних систем*.

2019. № 38. С. 193-199.

Автору належить розробка концептуальної моделі нейромережевого розпізнавання емоційного стану слухачів системи дистанційного навчання.

9. Терейковська Л.О., Терейковський О.І. Нейромережева модель розпізнавання емоцій по зображенню обличчя. *Вчені записки Таврійського національного університету імені В.І.Вернадського, серія «Технічні науки».* 2019. Том 30 (69). Ч. 1. №2. С. 209-213.

Автору належить розробка та верифікація нейромережевої моделі розпізнавання емоцій по зображенню обличчя

10. Дичка І., Терейковський І., Терейковська Л., Радченко К. Метод визначення ефективного типу базисного вейвлету для застосування в шаблонах нормальної поведінки веб-сервера. *Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні*. 2018. Вип. 2(36). С. 46-55.

Автору належить розробка принципів та критеріїв визначення ефективності типу базисного вейвлету.

11. Терейковська Л. О., Іванченко Є. В., Погорелов В.В. Метод адаптації глибокої нейронної мережі до розпізнавання комп'ютерних вірусів. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2019. Вип. 35. С. 198-205.

Автору належить розробка підходів адаптації глибокої нейронної мережі до умов задачі розпізнавання.

12. Михайленко В. М., Терейковская Л.А. Парадигмы нейросетевого распознавания эмоций. *Управління розвитком складних систем*. 2019. №39. С. 179-186.

Автору належить обґрунтування підходів та парадигм застосування нейронних мереж для розпізнавання емоцій.

13. Терейковська Л. Метод нейромережевого аналізу клавіатурного почерку. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2019. Вип. 37. С. 53-59.

14. Терейковська Л.О. Метод визначення виду згорткової нейронної мережі для аналізу параметрів клавіатурного почерку. *Вчені записки Таврійського національного університету імені В.І.Вернадського, серія «Технічні науки».* 2019. Том 30 (69). Ч. 1. № 6. С. 164-168.

15. Терейковська Л.О. Метод нейромережевого аналізу голосового сигналу. *Кібербезпека: освіта, наука, техніка*. 2020. № 3 (7). С. 31-42.

16. Терейковская Л.А. Нейросетевая модель распознавания эмоций по голосу. *Управління розвитком складних систем*. 2020. №41. С. 95-100.

17. Терейковська Л.О. Модель формування навчальних прикладів нейронної мережі, що призначена для аналізу клавіатурного почерку. *Кібербезпека: освіта, наука, техніка*. 2020. № 1 (9). С. 104-114.

18. Терейковська Л. Метод нейромережевого розпізнавання емоцій по зображенню обличчя. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2020. Вип. 40. С. 146-152.

19. Терейковская Л.А., Терейковский И.А. Применение сверточной нейронной сети для анализа биометрических параметров. *Вчені записки Таврійського національного університету імені В.І.Вернадського, серія «Технічні науки»*. 2020. Том 31 (70). № 5. С. 124-128.

Автору належить розробка та верифікація методу застосування згорткової нейронної мережі для аналізу біометричних параметрів.

20. Чернышев Д.О., Михайленко В.М., Терейковская Л.А. Разработка и исследование инструментальных средств для нейросетевого анализа голоса слушателей системы дистанционного обучения. *Управління розвитком складних систем*. 2020. № 43. С. 123-130.

Автору належить розробка та верифікація інструментальних засобів для нейромережевого аналізу голосу слухачів системи дистанційного навчання.

21. Кулаков Ю., Терейковська Л., Терейковський І. Спосіб застосування згорткової нейронної мережі для розпізнавання особи і емоцій користувача за клавіатурним почерком. *Правове, нормативне та*

метрологічне забезпечення системи захисту інформації в Україні. 2019. Вип. 2 (38). С. 9-17.

Автору належить розробка та верифікація згорткової нейронної мережі для розпізнавання особи і емоцій користувача за клавіатурним почерком.

22. Терейковская Л.А. Архитектура нейросетевого анализатора биометрических параметров слушателей системы дистанционного обучения. *Управління розвитком складних систем*. 2020. № 44. С. 91-99.

23. Терейковська Л. Метод фільтрації біометричних параметрів на основі вейвлет-перетворень. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2021. №42. С. 95-103.

24. Терейковський І.А., Кулаков Ю.О., Терейковська Л.О., Терейковський О.І. Параметри оцінки ефективності засобів розпізнавання особи за райдужною оболонкою та сітківкою ока. *Вчені записки Таврійського національного університету імені В.І.Вернадського, серія «Технічні науки»*. 2022. Том 33 (72). № 1. С. 195-199.

Автору належить розробка підходів та обґрунтування переліку параметрів оцінки ефективності засобів розпізнавання особи за райдужною оболонкою та сітківкою ока.

Матеріали наукових та науково-практичних конференцій

1. Tereikovskyi I., Mussiraliyeva S., Tereikovska L., Chernyshev D., Nyussupov A., Abaiuly Y. The Procedure of Adapting the Design Parameters of the Convolutional Neural Network During the Speaker's Emotions Recognition. *International Conference on Smart Information Systems and Technologies (SIST).* 28-30 April 2022. Nur-Sultan. P. 68-73.

Автору належить розробка підходу щодо адаптації конструктивних параметрів згорткової нейронної мережі.

2. Терейковська Л. О. Проблема створення ефективних засобів розпізнавання голосових сигналів, що адаптовані до умов вітчизняних Інтернет-орієнтованих інформаційних систем. *III Міжнародна науково-*

практична конференція «Актуальні питання забезпечення кібербезпеки та захисту інформації». Закарпатська обл., Міжгірський р-н., с. Верхнє Студене, 22–25 лютого 2017 р.: зб. тез допов. С. 163–165.

3. Terejkowski I., Terejkowska L. Calculation of expected output signal of neural network model for detecting of cyber-attack on network resources. *The 15th International Scientific Conference Information Technologies and Management*. April 27-28, 2017. ISMA University, Riga, Latvia. P. 59-62.

Автору належить розробка підходу щодо визначення вихідного сигналу нейромережевої моделі з урахуванням близькості еталонів класів.

4. Терейковський І.А., Терейковська Л.О., Ахметов Б.Б. Адаптація багатошарового персептрону до умов застосування в системах голосової аутентифікації. *Матеріали XIII міжнародної науково-технічної конференції «ABIA-2017»*. 2017. Київ. С.1.1-1.4.

Автору належить розробка математичного забезпечення адаптації двошарового персептрону до умов задачі голосової аутентифікації.

5. Терейковский И.А., Терейковская Л.А. Многослойный персептрон в задачах оценки параметров технических систем. *Збірка праць XVII Міжнародної наукової конференції імені Т. А. Таран «Інтелектуальний аналіз інформації»*. 17-19 травня 2017 р. м. Київ. С.233-238.

Автору належить розробка математичного забезпечення для розрахунку структурних параметрів багатошарового персептрону.

6. Терейковський І.А., Терейковська Л.О., Терейковський О.І. Визначення ознак деструктивного впливу в невербальному контенті звукових рядів Інтернет-орієнтованих засобів масової інформації. *Кібербезпека в Україні: правові та організаційні питання: матеріали міжн. наук. практ. конф.* 26 листопада 2020 р. м. Одеса. ОДУВС. С.76-79.

Автору належить розробка множини ознак деструктивного впливу.

7. Tereikovska L., Tereikovskyi I. Optimization of Convolutional Neural Network Structure for Biometric Authentication by Face Geometry. *Матеріали V*

Міжнародної науково-практичної конференції "Управління розвитком технологій". 19-20 травня 2018 р. м. Київ. КНУБА. С 42-43.

Автору належить аналіз перспектив адаптації параметрів згорткових нейронних мереж до умов задачі аутентифікації.

8. Міхайленко В. М., Терейковська Л. О. Оцінка ефективності нейромережевих моделей, призначених для розпізнавання емоцій диктора. *Тези доповідей сьомої міжнародної науково-практичної конференції «Управління розвитком технологій»*. 2020. м. Київ. КНУБА. С. 91-92.

Автору належить розробка концепції оцінки ефективності нейромережевих моделей, призначених для розпізнавання емоцій диктора.

9. Терейковський І.А., Терейковська Л.О. Застосування експертних знань для навчання нейронних мереж. *Матеріали III Міжнародної науково-технічної Internet-конференції «Сучасні методи, інформаційне, програмне та технічне забезпечення систем управління організаційно-технічними та технологічними комплексами»*. 23 листопада 2016. м. Київ. НУХТ. С. 147.

Автору належить розробка підходу до навчання нейронних мереж.

10. Міхайленко В.М., Терейковська Л.О., Терейковський О.І. Нейромережеве розпізнавання відбитків пальців. *Матеріали V Міжнародної науково-технічної Internet-конференції «Сучасні методи, інформаційне, програмне та технічне забезпечення систем керування організаційно-технічними та технологічними комплексами»*. 22-23 листопада 2018 р. м. Київ. НУХТ. С. 179-180.

Автору належить розробка нейромережевої моделі для розпізнавання відбитків пальців.

11. Терейковська Л.О. Нейромережева модель розпізнавання емоцій по голосу. *Матеріали VII Міжнародної науково-технічної Internet-конференції «Сучасні методи, інформаційне, програмне та технічне забезпечення систем керування організаційно-технічними та технологічними комплексами»*. 26 листопада 2020. м. Київ. НУХТ. С.124-125.

12. Tereikovska L.A., Radchenko K.O., Kharlamov K.O. The application of

discrete wavelet transform for forecasting the operational load of the Web server of the distance education system. *The Eighth World Congress «Aviation in the XXI-st Century» - «Safety in Aviation and Space Technologies»*. 10.10 – 12.10.2018. Kyiv Aviation University. P. 3.2.13-3.2.17.

Автору належить розробка підходу до застосування вейвлет-перетворень для аналізу нестаціонарних процесів.

13. Терейковський І., Терейковська Л., Терейковський О. Нейромережеві технології біометричної ідентифікації. *Безпека соціально-економічних процесів в кіберпросторі*: матеріали Всеукр. наук.-практ. конф. 27 березня 2019 р. м. Київ. КНТЕУ. 2019. С. 101-103.

Автору належить аналіз нейромережевих технологій ідентифікації.

14. Mikhaylenko V., Tereikovska L. Encoding the output signal of the convolutional neural network intended to recognize handwritten symbol. *Working program and proceedings of international scientific-practical conference of young scientists «Build-Master-Class-2017»*. 16-18 of November, 2017. Kyiv National University of Construction and Architecture. Kyiv. P. 349-350.

Автору належить розробка підходу до визначення вихідного сигналу згорткової нейронної мережі.

15. Терейковська Л. О. Застосування нейронних мереж в програмних системах, що призначені для розпізнавання емоцій диктора. *Тези доповідей першої науково-практичної конференції «Розподілені програмні системи і технології»*. 2020. м. Київ. КНУБА. С. 64.

16. Терейковский И.А., Терейковская Л.А., Картбаев Т.С., Бекетова Г.С. Способ нейросетевого распознавания личности пользователя по радужной оболочке глаза. *Материалы XI Международной научно-технической конференции «Энергетика, инфокоммуникационные технологии и высшее образование».* 16-18 октября 2020. Алматы. С.297-300.

Автору належить розробка моделі нейронної мережі для розпізнавання особи користувача за райдужною оболонкою ока.

ЗМІСТ

АНОТАЦІЇ 2
СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ 6
ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ 23
ВСТУП 24
РОЗДІЛ 1. АНАЛІЗ МОЖЛИВОСТЕЙ ВПРОВАДЖЕННЯ ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО СТАНУ В СИСТЕМУ ДИСТАНЦІЙНОГО НАВЧАННЯ 34
1.1. Передумови розробки інформаційної технології автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання 34
1.2. Аналіз технологій розпізнавання емоцій та особи за зображенням обличчя 44
1.3. Аналіз технологій розпізнавання емоцій та особи за голосом 53
1.4. Аналіз технологій розпізнавання емоцій та особи за клавіатурним почерком 62
1.5. Аналіз технологій розпізнавання особи за додатковими біометричними параметрами 67
1.6. Постановка проблеми дослідження 80
1.7. Висновки до першого розділу 83
РОЗДІЛ 2. ЕЛЕМЕНТИ МЕТОДОЛОГІЧНОГО ЗАБЕЗПЕЧЕННЯ ТЕХНОЛОГІЇ АВТОМАТИЗОВАНОГО РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО СТАНУ СЛУХАЧІВ СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ 85
2.1. Концептуальна модель розпізнавання емоційного стану 85
2.2. Підходи до нейромережевого аналізу біометричних параметрів для розпізнавання емоційного стану 96
2.3. Моделі обробки біометричних параметрів, асоційованих з зображеннями 107
2.4. Нейромережева модель розпізнавання емоцій слухача, адаптована 134

до нівелювання завад, пов’язаних з поворотом обличчя на зображенні
2.5. Висновки до другого розділу 139
РОЗДІЛ 3. МЕТОДИ ФОРМУВАННЯ ВХІДНОГО ПОЛЯ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ РОЗПІЗНАВАННЯ ЕМОЦІЙ СЛУХАЧІВ СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ 142
3.1. Метод застосування вейвлет-перетворень для фільтрації вхідного поля 142
3.2. Метод формування вхідного поля нейромережевої моделі аналізу біометричних параметрів, асоційованих з зображеннями 155
3.3. Метод формування вхідного поля нейромережевої моделі аналізу клавіатурного почерку 170
3.4. Метод формування вхідного поля нейромережевої моделі аналізу голосового сигналу 186
3.5. Висновки до третього розділу 209
РОЗДІЛ 4. МЕТОДИ НЕЙРОМЕРЕЖЕВОГО РОЗПІЗНАВАННЯ ЕМОЦІЙ ТА ОСОБИ СЛУХАЧА СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ 211
4.1. Метод формування навчальних прикладів для нейромережевої моделі аналізу біометричних параметрів 211
4.2. Метод розробки архітектури згорткової нейронної мережі, призначеної для аналізу біометричних параметрів 229
4.3. Метод нейромережевого розпізнавання емоцій 240
4.4. Етапи реалізації методології автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання 250
4.5. Висновки до четвертого розділу 267
РОЗДІЛ 5. СИСТЕМА АВТОМАТИЗОВАНОГО РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО СТАНУ ТА ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ 269
5.1. Архітектура системи автоматизованого розпізнавання емоційного стану 269
5.2. Оцінка ефективності засобів нейромережевого аналізу основних 284

біометричних параметрів
5.3. Оцінка ефективності засобів нейромережевого аналізу додаткових біометричних параметрів 313
5.4. Оцінка ефективності системи автоматизованого розпізнавання емоційного стану та особи слухача 322
5.5. Висновки до п’ятого розділу 332
ВИСНОВКИ 334
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 337
ДОДАТОК А Акти впроваджень результатів дисертаційного дослідження 376
ДОДАТОК Б ПРОГРАМНИЙ КОД ЗАСТОСУНКУ ДЛЯ РОЗПІЗНАВАННЯ ЕМОЦІЙ ТА ОСОБИ ЛЮДИНИ ЗА ГОЛОСОМ 384

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

АПЗ – апаратно-програмне забезпечення

БД – база даних

БП – біометричний параметр

БШП – багатошаровий персептрон

ВШ – вхідний шар

ГНМ – глибока нейронна мережа

ДКП – динаміка клавіатурного почерку

ДШП – двошаровий персептрон

ІС – інформаційна система

ЗАРЕС – засіб автоматизованого розпізнавання емоційного стану

ЗНМ – згорткова нейронна мережа

КП – клавіатурний почерк

НМ – нейронна мережа

НММ – нейромережева модель

НМС – нейромережева система

НМЗ – нейромережевий засіб

САРЕС – система автоматизованого розпізнавання емоційного стану

СДН – система дистанційного навчання

ТМК – термін між послідовним відпусканням однієї клавіші та натисканням наступної клавіші

ТУК – термін утримання клавіші

ВСТУП

Актуальність теми. Відповідно до уявлень сучасної науки, одним із найбільш перспективних напрямків розвитку систем дистанційного навчання (СДН) є застосування в них технологій, що дозволяють ефективно організувати пізнавальну діяльність слухачів для засвоєння ними навчальних матеріалів. При цьому вкрай важливу роль відіграє емоційний стан слухачів, що вказує на доцільність впровадження в СДН відповідних засобів розпізнавання емоційного стану. Водночас в більшості відомих СДН засоби розпізнавання емоційного стану не використовуються, хоча можливість їх впровадження підтверджується відомими фактами використання відповідних технологій та програмних додатків (системи розпізнавання емоцій компаній Microsoft, Taigusys, HireVue, Pymetrics). Вказані системи в основному використовуються для оцінки відповідності респондентів професійним вимогам, а використані в них модулі розпізнавання побудовані на базі нейромережевих моделей (НММ) та методів аналізу біометричних параметрів (БП).

Разом з тим, впровадження відомих засобів автоматизованого розпізнавання емоційного стану у вітчизняні СДН викликає необхідність їх складної адаптації до варіативності умов застосування, пов’язаних з терміном розробки, допустимою величиною похибки розпізнавання, акустичними факторами, особливостями відеопотоку та аудіопотоку, характеристиками сенсорів отримання інформації, ресурсоємністю створення та функціонування. Також недоліками поширених засобів розпізнавання емоційного стану є висока вартість та відсутність детальної науково-технічної документації

В такій постановці є актуальною **науково-прикладна проблема** розробки методології розпізнавання емоційного стану слухачів системи дистанційного навчання, що базується на нейромережевих методах і моделях

аналізу біометричних параметрів, та орієнтована на створення відповідних інструментальних засобів, які розширюють функціональні можливості системи дистанційного навчання. Вказана проблема викликана об’єктивним протиріччям між потребою автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання на основі аналізу його біометричних параметрів та відсутністю в ній засобів такого аналізу.

Теоретичні та практичні результати, що пов'язані з розробкою та впровадженням технологій розпізнавання ЕС отримані такими вітчизняними та іноземними вченими, як Аль-Амморі А. Н., Биков В. Ю., Бодянський Є. В., Бушуєв С. Д., Данчук В.Д., Корж Р. О., Любчик Л. М., Міхайленко В. М., Руденко О. Г., Ekman P., Partila P., Ranganathan H. та інші. Але, незважаючи на запропоновані підходи та рішення, у відомій літературі практично відсутні наукові публікації, в яких об’єктом дослідження були б процеси розпізнавання емоційного стану слухачів СДН. Крім того, потребують доопрацювання і нейромережеві моделі та методи аналізу БП слухача СДН, що використовуються для розпізнавання.

Зв'язок роботи з науковими програмами, планами, темами. Тематика дисертаційної роботи та одержані результати безпосередньо пов’язані з виконанням держбюджетних науково-дослідних робіт:

– Київського національного університету будівництва і архітектури по темі «Нейромережеві методи, моделі та засоби інформаційної технології управління емоційним станом слухачів системи дистанційного навчання» (№0122U001768);

– Київського національного університету будівництва і архітектури по темі «Методи та засоби розпізнавання емоцій диктора» (№0121U110859);

– Київського національного університету будівництва і архітектури по темі «Теоретичні основи інформаційної технології розпізнавання психоемоційного стану слухачів системи дистанційного навчання» (№0121U110668);

– Національного технічного університету України "Київський

політехнічний інститут імені Ігоря Сікорського" по темі «Методи, моделі та комп’ютерні засоби виявлення деструктивного впливу в медіапросторі» (№0121U110662).

Наукова новизна дисертаційного дослідження полягає в тому, що розроблені теоретичні основи, моделі, методи та засоби автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання є новим напрямом досліджень в галузі інформаційних технологій.

Об'єктом дослідження є процеси розпізнавання емоційного стану слухачів системи дистанційного навчання.

Предметом дослідження є моделі, методи та засоби процесів розпізнавання емоційного стану слухачів системи дистанційного навчання.

Основна гіпотеза дослідження полягає у можливості забезпечення ефективного розпізнавання емоційного стану та особи слухача системи дистанційного навчання за рахунок нейромережевого аналізу біометричних параметрів, котрі можуть бути зареєстровані за допомогою загальнопоширеного апаратно-програмного забезпечення.

Метою дослідження є розв’язання науково-прикладної проблеми, пов’язаної з розпізнаванням емоційного стану слухачів системи дистанційного навчання шляхом розробки відповідної методології, що базується на нейромережевих методах і моделях аналізу біометричних параметрів, та орієнтованої на створення відповідних засобів, які розширюють функціональні можливості системи дистанційного навчання.

Завдання дисертаційного дослідження.

Для досягнення поставленої мети необхідно розв'язати наступні **завдання:**

– виконати аналіз передумов та перспектив впровадження технологій розпізнавання емоційного стану в систему дистанційного навчання;

– побудувати концептуальну модель розпізнавання емоційного стану слухача системи дистанційного навчання на основі нейромережевого аналізу біометричних параметрів, що реєструються за допомогою розповсюдженого

апаратного забезпечення;

– удосконалити підходи до застосування нейронних мереж для аналізу біометричних параметрів заради розпізнавання емоційного стану та особи слухача системи дистанційного навчання;

– побудувати моделі обробки біометричних параметрів, асоційованих з зображеннями для нівелювання завад, характерних для умов системи дистанційного навчання;

– розробити метод застосування вейвлет-перетворень для фільтрації вхідного поля нейромережевої моделі, призначеної для розпізнавання емоційного стану та особи слухача системи дистанційного навчання на основі біометричних параметрів, асоційованих з зображеннями;

– створити комплекс методів формування вхідного поля нейромережевих моделей, призначених для аналізу біометричних параметрів задля розпізнавання емоційного стану та особи слухача;

– розробити метод формування навчальних прикладів, призначених для використання в нейромережевих моделях аналізу біометричних параметрів;

– побудувати метод розробки архітектури згорткової нейронної мережі, призначеної для аналізу біометричних параметрів;

– розробити метод нейромережевого розпізнавання емоцій за одним із основних біометричних параметрів;

– розробити методологію автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання;

– побудувати архітектуру системи автоматизованого розпізнавання емоційного стану;

– провести експериментальні дослідження, спрямовані на оцінку ефективності запропонованих рішень.

Методи дослідження. Використовувались методи дослідження, що ґрунтуються на методологічному базисі теорії складних систем, теорії інформації, розпізнавання образів і системному аналізі – для аналізу предметної області досліджень і окреслення перспектив впровадження

технології розпізнавання емоційного стану в систему дистанційного навчання; теорії нейронних мереж, вейвлет-перетворень, кодування, прийняття рішень, моделювання, методи цифрової обробки сигналів та біометрії – для розроблення моделей та методів розпізнавання емоційного стану та особи; методів імітаційного моделювання, об'єктно-орієнтованого проєктування та елементів теорії алгоритмів, програмування та планування експерименту – для розроблення структурних рішень та інструментальних засобів автоматизованого розпізнавання емоційного стану.

Наукова новизна отриманих результатів. Основний науковий результат полягає в тому, що розроблено і науково обґрунтовано методологію автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання, яка включає інтегровану сукупність підходів, моделей, методів нейромережевого аналізу біометричних параметрів та орієнтована на створення відповідних інструментальних засобів, котрі розширюють функціональні можливості системи дистанційного навчання.

В процесі розв'язання означеного проблемного методологічного завдання отримано наступні нові наукові результати:

Вперше розроблено:

– методологію автоматизованого розпізнавання емоційного стану, що за рахунок інтегрованого застосування відомих та запропонованих підходів, моделей та методів нейромережевого аналізу біометричних параметрів, забезпечує можливість створення ефективних інструментальних засобів розпізнавання емоційного стану, які розширюють функціональні можливості системи дистанційного навчання;

– концептуальну модель розпізнавання емоційного стану слухачів системи дистанційного навчання, яка за рахунок конкретизації операцій нейромережевого аналізу біометричних параметрів та за рахунок обґрунтованого переліку параметрів оцінки ефективності розпізнавання забезпечила формалізований опис напрямків досліджень з розробки

відповідних засобів автоматизованого розпізнавання;

– метод формування вхідного поля нейромережевої моделі аналізу клавіатурного почерку слухача системи дистанційного навчання, що за рахунок подання параметрів клавіатурного почерку у вигляді кольорового рисунку, придатного для аналізу за допомогою згорткової нейронної мережі, забезпечує можливість ефективного розпізнавання емоцій та особи слухача системи дистанційного навчання за клавіатурним почерком;

– метод формування вхідного поля нейромережевої моделі аналізу голосового сигналу, що за рахунок застосування процедур перевірки характеристик клієнтського апаратно-програмного забезпечення, валідації величин зареєстрованих параметрів, перевірки відсутності багатоголосся та часу реверберації, приведення параметрів голосового сигналу та за рахунок адаптації параметрів голосового сигналу до аналізу згортковою нейронною мережею забезпечує можливість ефективного розпізнавання емоцій та особи слухача системи дистанційного навчання за голосом.

Удосконалено:

– підходи до визначення конструктивних параметрів, навчання та мінімізації ресурсоємності нейромережевих моделей, що за рахунок теоретичного обґрунтування діапазону значень конструктивних параметрів, кодування очікуваного вихідного сигналу навчальних прикладів із можливістю врахування близькості еталонів визначених класів, навчання нейромережевої моделі з використанням експертних даних, обґрунтованої мінімізації кількості вхідних параметрів, забезпечили можливість розробки ефективних методів формування вхідного поля нейромережевої моделі та ефективних методів нейромережевого розпізнавання емоцій та особи слухача системи дистанційного навчання;

– метод формування вхідного поля нейромережевої моделі аналізу біометричних параметрів, асоційованих з зображеннями, що за рахунок адаптації процедур, пов’язаних із обробкою зображень, нівелюванням впливу типових завад та процедур визначення параметрів ключових і

контрольних точок, до вимог нейромережевого аналізу зображень обличчя, вушної раковини, райдужної оболонки ока, сітківки ока та відбитку пальця, забезпечує можливість ефективного розпізнавання емоцій та особи слухача системи дистанційного навчання за біометричними параметрами, асоційованими з зображеннями;

– метод розробки архітектури згорткової нейронної мережі, призначеної для аналізу біометричних параметрів, який за рахунок використання запропонованих критеріїв ефективності та принципів адаптації конструктивних параметрів нейронної мережі до умов функціонування, дозволяє зменшити обсяг експериментальних досліджень, пов’язаних з розробкою архітектури згорткової нейронної мережі;

– метод нейромережевого розпізнавання емоцій, що за рахунок застосування запропонованих рішень щодо формування вхідного поля та архітектури нейромережевої моделі для аналізу зображення обличчя, голосового сигналу, клавіатурного почерку та рішень щодо відображення у вихідному полі нейромережевої моделі схожості еталонів емоцій, дозволяє за прийнятної ресурсоємності підвищити точність розпізнавання емоцій слухача системи дистанційного навчання за вказаними біометричними параметрами, що реєструються за допомогою розповсюдженого апаратно-програмного забезпечення.

Отримали подальший розвиток:

– моделі обробки біометричних параметрів, асоційованих з зображеннями, що за рахунок врахування особливостей реєстрації зображень в умовах системи дистанційного навчання, вимог до підготовки вхідних даних згорткової нейронної мережі та за рахунок обґрунтованого застосування технологій комп’ютерного зору забезпечили теоретичний базис розробки методу формування вхідного поля нейромережевих моделей аналізу біометричних параметрів, асоційованих з зображеннями;

– метод застосування вейвлет-перетворень для фільтрації вхідного поля нейромережевої моделі аналізу біометричних параметрів, асоційованих

з зображеннями, що за рахунок інтеграції вейвлет-коефіцієнтів, розрахованих з використанням ефективного типу базисного вейвлету для послідовно зареєстрованих зображень, забезпечує можливість фільтрації завад, типових для умов системи дистанційного навчання;

– метод формування навчальних прикладів для нейромережевої моделі аналізу біометричних параметрів, що за рахунок обробки навчальних прикладів ймовірнісною нейронною мережею дозволяє врахувати в очікуваному вихідному сигналі близькість еталонів визначених класів, що забезпечує можливість зменшення ресурсоємності процесу навчання.

Одержані результати загалом утворюють теоретико-інструментальну основу інформаційної технології автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання.

Практичне значення одержаних результатів. Отримані у дисертаційній роботі наукові результати є методологічною базою для розробки і впровадження ефективних інструментальних засобів у вигляді програмних або програмно-апаратних модулів розпізнавання емоційного стану, які мають достатню точність розпізнавання та адаптовані до застосування в умовах системи дистанційного навчання.

Практична цінність полягає у наступному:

– розроблена архітектура системи автоматизованого розпізнавання емоційного стану, що може бути використана для розширення функціональних можливостей сучасних систем дистанційного навчання;

– розроблене алгоритмічне і програмне забезпечення модулів, що реалізують розпізнавання емоцій та особи користувача інформаційної системи за голосовим сигналом, зображенням обличчя та клавіатурним почерком;

– розроблене спеціалізоване програмне забезпечення, що базується на створених нейромережевих методах і моделях, дозволило забезпечити можливість розпізнавання емоційного стану та підвищити оперативність створення алгоритмів функціонування програмних засобів аналізу

біометричних параметрів користувачів інформаційних систем, що підтверджується актами впровадження в діяльність Київського національного університету будівництва і архітектури (акт впровадження від 16.05.2022, 19.05.2022, 20.05.2022) та Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського" (акт впровадження від 18.05.2022).

– розроблені програми, що реалізують запропоновані моделі та методи, впроваджені в навчальний процес на кафедрі інформаційних технологій проєктування та прикладної математики Київського національного університету будівництва і архітектури (акт впровадження від 23.05.2022), кафедрі прикладної математики Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського" (акт впровадження від 25.05.2022) та на кафедрі безпеки інформаційних технологій Національного авіаційного університету (акт впровадження від 26.05.2022).

Особистий внесок здобувача. Всі основні положення та результати дисертаційної роботи, що виносяться на захист, отримані здобувачем самостійно. У дисертаційній роботі автором не використовувалися розробки і ідеї, що належать співавторам спільних наукових праць. Конкретний внесок автора в кожну із вказаних праць визначено у переліку публікацій за темою дисертації. При вирішенні проблеми і задач дисертаційного дослідження із опублікованих у співавторстві робіт здобувачем використано тільки особисто отримані результати.

Апробація результатів дисертації. Основні результати дисертації доповідались та отримали позитивні оцінки на 32 міжнародних наукових конференціях: «Актуальні питання забезпечення кібербезпеки та захисту інформації» (Верхнє Студене, 2017); «Information Technologies and Management» (Riga, 2017); «ABIA-2017» (Київ, 2017); «Кібербезпека в Україні: правові та організаційні питання» (Одеса, 2020); «Управління розвитком технологій» (Київ, 2018, 2020); «Сучасні методи, інформаційне,

програмне та технічне забезпечення систем управління організаційно-технічними та технологічними комплексами» (Київ, 2016, 2018, 2020); «Aviation in the XXI-st Century» (Kyiv, 2018); «Безпека соціально-економічних процесів в кіберпросторі» (Київ, 2019); «Build-Master-Class» (Київ, 2017); «Розподілені програмні системи і технології» (Київ, 2020); XVIII International Conference on Data Science and Intelligent Analysis of Information (Kyiv, 2018); 1st International Conference on Cyber Hygiene and Conflict Management in Global Information Networks (Lviv, 2019); International Conference on Computer Science, Engineering and Education Applications (Kyiv, 2018-2020); міжнародних конференціях IEEE 2017–2022 років.

Публікації. За тематикою дослідження опубліковано 66 наукових робіт (48 основних з яких наведено в авторефераті), у тому числі: 1 монографія, 15 наукових статей в міжнародних виданнях, що входять в бази даних Scopus та Web of Science, 2 статті в іноземних наукових журналах, 24 наукові статті у фахових наукових журналах України, 24 тези доповідей наукових конференцій, 8 з яких проіндексовані в базах даних Scopus та Web of Science.

Структура роботи та її обсяг. Дисертація складається з анотації, змісту, списку скорочень, вступу, п’яти розділів, загальних висновків, списку використаних джерел та додатків. Загальний обсяг дисертації складає 395 сторінок, в тому числі 89 рисунків, 31 таблицю, список літератури з 327 найменувань на 39 сторінках і 2 додатків на 20 сторінках.

РОЗДІЛ 1
АНАЛІЗ МОЖЛИВОСТЕЙ ВПРОВАДЖЕННЯ ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО СТАНУ В СИСТЕМУ ДИСТАНЦІЙНОГО НАВЧАННЯ

1.1. Передумови розробки інформаційної технології автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання

Протягом останніх 10-15 років однією з основних тенденцій розвитку вищої освіти є все більш широке впровадження дистанційних форм навчання [4, 6, 116, 277]. У сучасних умовах зазначена тенденція лише посилюється через необхідність дотримання карантину, спричиненого пандемією коронавірусної інфекції COVID-19. При цьому використання дистанційної форми навчання в якості основної посилює необхідність підвищення ефективності механізмів її реалізації. Вказана проблема пов'язана з тим, що проведення онлайн-трансляцій дистанційних лекцій та практичних занять значно збільшує навантаження на викладача, а при подачі навчальних матеріалів виключно у вигляді ілюстрованого тексту або мультимедійної презентації суттєво погіршується якість сприйняття [2, 63, 79, 140]. На думку ряду авторитетних дослідників [12, 13, 16, 20], погіршення якості сприйняття навчальних матеріалів багато в чому пов'язане з відсутністю моніторингу емоційної складової навчального процесу. Це підтверджується наведеним в [18, 48] законом «Йеркса і Додсона», в якому показана залежність результативності праці людини від ступеню її емоційної напруги. На основі експериментальних даних доведено, що при збільшенні емоційного збудження результативність праці спочатку швидко росте, потім таке зростання дещо стихає, а після досягнення певного критичного рівня починає спочатку плавно, а далі різко зменшуватись. Як свідчать результати [43, 71, 119], в процесі навчання учень може відчувати різні емоції, які по-різному впливають на засвоєння

інформації. Учні можуть дивуватися, відчувати захоплення, наснагу і навіть гнів. Всі ці емоції можуть позитивно впливати на процес навчання, незважаючи на те, що не всі емоції самі по собі позитивні [151, 257]. Зазначається, що головне уникати таких емоцій, як смуток і нудьга, оскільки якщо учнем опанують дані емоції, то він не зможе якісно засвоїти інформацію. А нейтральні події у навчанні взагалі практично не відкладаються в пам'яті учня. Виникнення смутку, нудьги або апатії можливе, наприклад, при нецікавій подачі матеріалу. Ще більш негативною емоцією в процесі навчання являється страх, виникнення якого не тільки нанівець зводить прагнення до знань, але й взагалі блокує здатність до мислення.

Зазначимо, що емоції відображають ставлення людини до явищ навколишнього світу у зв'язку з можливістю задоволення нею своїх потреб. Емоційний стан являється узагальненням емоцій на протязі деякого проміжку часу.

Емоції мають дві сторони - об'єктивну (людина, яка відчуває емоції може посміхатися, хмуритися, плакати, тремтіти, у неї змінюється частота пульсу, ритм дихання і т. ін.), а також суб'єктивну - внутрішні думки та переживання людини з приводу тих чи інших подій [43, 146]. У сучасній психології виділяють від чотирьох до восьми базових емоцій, прояви яких не залежать від раси та культури людини [1, 160, 232]. Базові емоції - це елементарні емоції, які більше ні на що не розщеплюються і самі можуть бути складовими інших складних емоцій. Відповідно до [1, 48, 69], до семи базових емоцій відносяться:

- Радість - позитивний емоційний стан, пов'язаний з можливістю задоволення актуальних потреб людини.

- Гнів - негативний емоційний стан, що викликається раптовим виникненням серйозної перешкоди на шляху задоволення важливих потреб.

- Відраза - негативний емоційний стан, викликане подією, яке різко суперечить моральним або естетичним принципам і установкам людини.

- Смуток - негативна емоція, пов'язана з переживанням негативного факту.

- Страх - негативний емоційний стан, що з'являється при отриманні суб'єктом інформації про реальну або уявну небезпеку.

- Подив - не має чітко вираженого позитивного або негативного знаку емоційної реакції на раптово виниклі обставини.

- Зневага - негативний емоційний стан, який виникає щодо негативного явища, яке вважається негідним того, щоб з ним рахуватися.

Також виділяють нейтральний емоційний стан (нейтральність), який пов'язаний з подіями, що не впливають на настрій та здоров'я людини. Приклади фотографій обличчя людини, що виражає одну з наведених базових емоції, показані на рис. 1.1 (а-ж) [166, 167, 170].

а) Радість б) Гнів в) Відраза г) Печаль

ґ) Страх е) Подив є) Презирство ж) Нейтральність

Рис. 1.1 Демонстрація прояву на обличчі людини базових емоцій

У випадку виділення чотирьох базових емоцій до їх складу відносять: щастя (радість), смуток (сум), переляк (страх), гнів (відраза). При цьому кількість і номенклатура складних емоцій досі чітко не визначена, а відповідна теорія далека від досконалості.

Так, для прикладу на рис. 1.2 показані зображення обличчя, що відображають складні емоції.

а) Гордість б) Прохання в) Втома г) Задоволення

Рис. 1.2 Демонстрація прояву на обличчі складних емоції

Всі емоції поділяються на позитивні і негативні, а також на стенічні і астенічні. Позитивні емоції пов'язані із задоволенням потреб, а негативні викликають почуття незадоволення і вимагають зміни ситуації. Стенічні емоції активізують організм і піднімають настрій (гнів, лють, захват), а астенічні (туга, печаль, смуток, сором) - розслаблюють людину і пригнічують активність організму. При цьому завдання розпізнавання емоцій є слабоформалізованим, а його вирішення ускладнюється невизначеністю характеристик проявів емоцій.

Саме емоції стимулюють центри відповідних відділів мозку, в які надходять сигнали. Цей емоційний каталізатор необхідний, щоб навчальна ситуація мала успіх, оскільки слухач може засвоїти матеріал лише тоді, коли у мозку активуються емоційні центри [1, 43]. Емоційний спектр людини досить широкий і якщо в СДН

буде враховано поточний емоційний стан слухача, то цей факт можна використати для впливу на швидкість подачі навчальних матеріалів та їх обсяг, що забезпечить підвищення ефективності процесу навчання [36, 38, 39].

Базуючись на даних [4, 13, 40], можна визначити такі функції емоційної складової педагогічного супроводу процесу дистанційного навчання:

1. Емоціогенна - покращення настрою, пробудження інтересу до навчальної діяльності та освоєння навчальної дисципліни.

2. Діагностична - розкриття власних резервів внутрішньої активності, можливість самовираження та самопізнання.

3. Релаксаційна - зняття або зменшення тривожності, фізичної та інтелектуальної напруги, відновлення внутрішніх сил та резервів.

4. Терапевтична - подолання труднощів у навчанні.

Як свідчать результати досліджень в області педагогіки, психології та управління проєктами, всебічність охоплення емоційної складової навчальної дисципліни можна забезпечити шляхом включення до змісту навчання емоційних концептів [43, 79]. Власне емоційні концепти можуть бути представлені різними засобами як вербального, так і невербального характеру [2, 105, 296, 322].

На сьогодні найбільш апробований підхід до гармонізації емоційного стану слухача вищого навчального закладу полягає у корегуванні режиму подачі навчальних матеріалів. Наприклад, подача лекційного матеріалу може бути здійснена методом готових знань, що передбачає пасивне сприйняття і запам’ятовування поданої педагогом інформації. Також лекційні матеріали можуть бути подані дослідницьким методом, котрий в свою чергу передбачає самостійне опрацювання запропонованого лекційного матеріалу: аналіз явищ, формулювання проблеми, формування та перевірку гіпотез, самостійне формулювання висновків. Також, відповідно рекомендацій [5, 17, 79, 119], до коригуючих впливів на режим подачі навчальних матеріалів можуть бути віднесені: тривалість проведення заняття, тривалість перерви, вид, контент та обсяг навчальних матеріалів.

Очевидно, що рішення про необхідність реалізації коригуючих впливів слід здійснювати на підставі поточного емоційного стану слухача СДН [159, 213]. Моніторинг емоційного стану доцільно реалізувати на основі аналізу БП, що відображають емоції слухача та можуть бути зареєстровані за допомогою розповсюджених комп'ютерних засобів. Вказане твердження базується на результатах аналізу літературних робіт [8, 19, 75, 168], які вказують на те, що в основі переважної більшості відомих систем автоматизованого та автоматичного розпізнавання емоційного стану покладена технологія розпізнавання емоцій, що базується на результатах аналізу БП.

Узагальнена схема взаємодії компонентів контуру корегування режиму подачі навчальних матеріалів в СДН показана на рис. 1.3.

Передбачено, що слухач в процесі виконання навчальної програми опрацьовує подані йому навчальні матеріали. Режим подачі навчальних матеріалів корегується спеціалізованим модулем – контролером, який може змінювати параметри режиму подачі з урахуванням графіку виконання навчальної програми та поточним емоційним станом слухача СДН. Поточний емоційний стан слухача СДН розпізнається на основі результатів аналізу його БП, що реєструються за допомогою комп'ютерних засобів, що входять до клієнтського АПЗ СДН.

Таким чином, вирішення завдання автоматизації педагогічного супроводу процесу дистанційного навчання з урахуванням емоційної складової може лежати у площині розробки електронних навчальних матеріалів, режим подачі яких буде адаптовано до емоційного стану кожного конкретного слухача СДН [308]. Очевидно, що базуватись такі розробки повинні на ЗАРЕС слухачів СДН. При цьому, як вже було зазначено вище, відомо лише про експериментальне застосування вказаних засобів розпізнавання емоційного стану в системах комп'ютерного супроводу навчального процесу. Так, подібні ЗАРЕС використовується в деяких школах Китайської Народної Республіки для розпізнавання учнів, що відволікаються від виконання навчальних завдань.

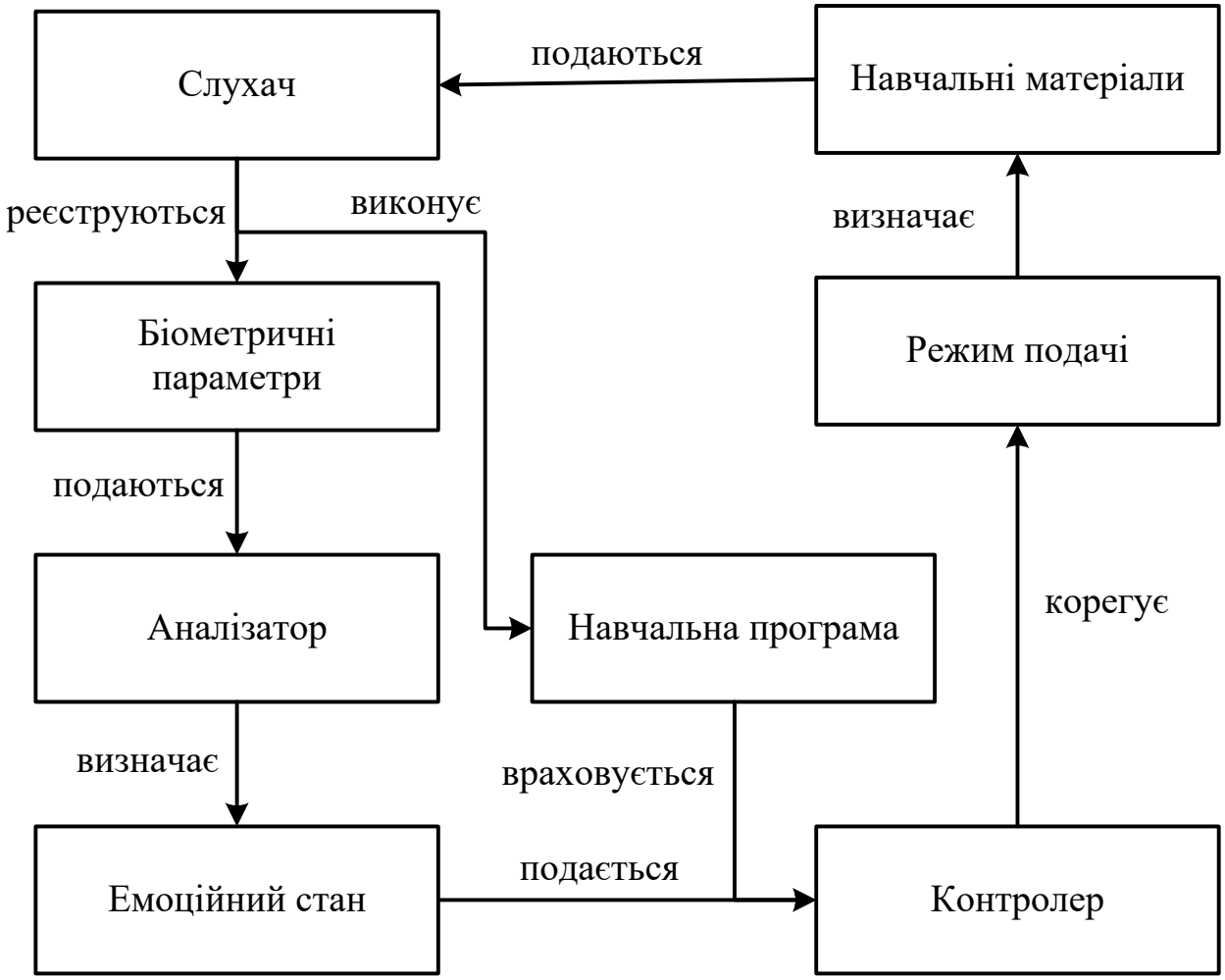

Рис. 1.3 Узагальнена схема взаємодії компонентів контуру корегування режиму подачі навчальних матеріалів в СДН

У поширених СДН типу Moodle (https://moodle.org), ATutor (https://atutor.ca), Blackboard (www.blackboard.com), Lotus LearningSpace (ibm.com/developerworks/lotus/library /ls-quickcourse_LVC), Canvas (https://www.canvas.net), Sakai LMS (https://www.sakailms.org) зазначені засоби моніторингу емоційного стану відсутні. Разом з тим, можливість застосування ЗАРЕС в СДН підтверджується вдалими прикладами в інших прикладних сферах. Так, відомі спроби застосування ЗАРЕС в системах розпізнавання впливу на підсвідомість в мультимедійних повідомленнях засобів масової інформації при

веденні інформаційних війн, для діагностування психічних захворювань, в індустрії розваг та системах управління розумним будинком. Також можна обґрунтовано стверджувати, що в базовому варіанті, як і для більшості універсальних ІС, достатньою є точність розпізнавання емоцій слухача СДН на рівні 0,8, а точність розпізнавання особи слухача в процесі навчання – 0,9. Це висуває на перший план питання розробки та впровадження в СДН інструментальних засобів розпізнавання емоційного стану слухача СДН виходячи з результатів аналізу БП, реєстрацію яких можна реалізувати за допомогою поширеного АПЗ [3, 7, 23, 25]. Насамперед до зазначених БП належать:

- зображення обличчя,
- голос,
- клавіатурний почерк.

Реєстрація таких БП передбачає застосування лише відеокамери, мікрофону та клавіатури, що де-факто завжди входять до складу типового АПЗ слухача СДН, адже навіть при під’єднанні слухача до СДН за допомогою смартфону ним може бути використана віртуальна клавіатура. Зазначимо, що ЗАРЕС на основі вказаних БП мають певні недоліки. Істотним обмеженням на використання засобів аналізу зображень обличчя є необхідність нівелювання характерних перешкод, пов'язаних із ракурсом відеозйомки, нерівномірністю освітлення, роздільною здатністю відеокамери [305, 315, 324]. Недоліки сучасних засобів аналізу КП пов'язані з недосконалістю сучасної теоретичної бази, що значно ускладнює отримання досить точних результатів, які стосуються емоційного стану слухача. Разом з тим, практичний досвід і результати науково-прикладних робіт [125, 177] вказують на необхідність суттєвої модернізації голосових аналізаторів в напрямку зменшення ресурсоємності, збільшення точності розпізнавання, розширення номенклатури емоцій, котрі мають бути розпізнані, скорочення терміну розробки та підвищення рівня адаптації до умов застосування. Нівелювати вказані недоліки можливо як за рахунок комплексного аналізу всіх вказаних БП, так і за рахунок вдосконалення

існуючих теоретичних рішень в області аналізу кожного із БП. При цьому в більшості відомих випадків [182, 185] очікуваним результатом аналізу БП являється інтегральне розпізнавання особи та емоційного стану індивіда. Також слід зазначити, що, відповідно до теоретичних розробок [1, 32, 184, 207], завдання розпізнавання емоційного стану індивіда суттєво спрощується у випадку попереднього розпізнавання особи індивіда. Таким чином, розпізнавання особи слухача являється однією із важливих складових завдання розпізнавання емоційного стану слухача СДН. Для розпізнавання особи слухача СДН доцільно використовувати не тільки такі БП, як зображення обличчя, голос і КП, а й інші БП, що успішно використовуються в засобах біометричної аутентифікації та можуть бути зареєстровані з використанням доступного АПЗ слухача СДН [240]. З урахуванням переліку та характеристик доступного АПЗ, до вказаних БП насамперед можуть бути віднесені:

- відбитки пальців,
- райдужна оболонка ока,
- сітківка ока,
- зображення вушної раковини.

Долучення до складу СДН засобів розпізнавання особи крім очікуваного підвищення точності розпізнавання емоційного стану може викликати ще один позитивний ефект, пов’язаний з підвищенням рівня захисту такої системи та з підвищенням рівня надійності результатів навчання. За рахунок засобів біометрії з’являється можливість реалізації автоматичного розпізнавання особи слухача СДН не тільки в процесі входу в СДН, а і в ході навчального процесу, зокрема під час опрацювання слухачем навчальних матеріалів та під час виконання тестових завдань.

Слід зазначити, що застосування в СДН відомих рішень в області аналізу БП для розпізнавання емоцій та особи слухача потребує їх адаптації до очікуваних умов застосування. Для визначення шляхів вдосконалення ЗАРЕС на першому етапі

досліджень проведено аналіз наукових робіт в даній області [1, 10, 31, 33]. В результаті такого аналізу було визначено групу базових параметрів, що дозволяють в першому наближенні оцінити ефективність ЗАРЕС в умовах СДН [52, 58, 85, 291]. Перелік та пояснення вказаних базових параметрів наведено в табл. 1.1.

Таблиця 1.1

Базові параметри оцінки ефективності засобів розпізнавання емоцій

№	Опис базових параметрів
$A_{i,1}$	Повнота розпізнавання базових емоцій
$A_{i,2}$	Можливість розпізнавання складних емоцій
$A_{i,3}$	Можливість розпізнавання емоційного стану
$A_{i,4}$	Точність розпізнавання
$A_{i,5}$	Обчислювальна складність розпізнавання
$A_{i,6}$	Можливість розпізнавання емоцій без попереднього розпізнавання особи

Наведений в табл. 1.1 перелік базових параметрів оцінки ефективності (група А) є спільним для всіх типів ЗАРЕС, незалежно від виду БП, на базі якого реалізується розпізнавання емоцій. Крім того, спільним для всіх типів ЗАРЕС параметром ефективності може служити параметр, пов’язаний з оцінкою здатності до розпізнавання емоцій в умовах завад. Однак через значну спорідненість завад з природою БП вказаний параметр було конкретизовано, відповідно до виду БП.

Також в результаті проведеного аналізу відомих ЗАРЕС, крім виділених в групу А базових параметрів, визначено необхідність використання додаткових параметрів, що дозволяють оцінити ефективність зазначених засобів. Додаткові параметри ефективності розділено на три групи. Група параметрів В характеризує особливості розпізнавання ЗАРЕС, група параметрів С - особливості технічної реалізації, а група D - сервісні (додаткові) можливості ЗАРЕС.

Для базових параметрів прийнято позначення виду $A_{i,j}$, де i – номер БП, на основі якого реалізується розпізнавання емоцій, j – номер базового параметру ефективності.

Використана наступна нумерація БП: зображення обличчя – 1, ГС – 2, КП – 3. Перелік параметрів, що відносяться до груп B-D залежить від виду БП, на основі якого реалізується розпізнавання. Для вказаних параметрів прийняті позначення виду $B_{i,j}$, $C_{i,j}$, $D_{i,j}$, де i – номер БП, на основі якого реалізується розпізнавання емоцій, j – номер параметру ефективності в межах відповідної групи.

Передбачено, що застосування вказаних параметрів дозволить реалізувати комплексну оцінку ефективності ЗАРЕС. У базовому варіанті виставлення кожної із оцінок пропонується реалізовувати в діапазоні від 0 до 1. Для деякого i-го параметра оцінка 0 означає, що в ЗАРЕС i-та вимога не забезпечується, оцінка 0,5 означає часткове забезпечення, а оцінка 1 - повне забезпечення такої вимоги.

Формування переліку базових параметрів та підходу до визначення додаткових параметрів дозволило перейти до наступного етапу досліджень, пов’язаного з аналізом технологій розпізнавання емоцій та особи для визначення шляхів їх вдосконалення та адаптації до умов СДН.

1.2. Аналіз технологій розпізнавання емоцій та особи за зображенням обличчя

У більшості випадків сучасні ЗАРЕС базуються на розпізнаванні міміки, яка визначається як сукупність рухів частин обличчя людини, що виражає його стан або ставлення до того, що він сприймає. Для цього в ЗАРЕС використовується модуль розпізнавання міміки, функціонування якого засновано на сучасних методах комп'ютерного зору, що в більшості випадків базуються на використанні HMM різних типів.

Найбільш простий спосіб визначення емоції по зображенню особи

заснований на класифікації ключових точок (facial landmarks), координати яких можна отримати, використовуючи різні алгоритми: PDM, CML, AAM, DPM або CNN. Зазвичай розмічають від 5 до 68 точок, прив'язуючи їх до положення брів, очей, губ, носа, щелепи, що дозволяє частково захопити міміку. Нормалізовані координати точок можна безпосередньо подати в класифікатор (наприклад, SVM або Random Forest) і отримати базове рішення. Положення зображення обличчя при цьому має бути вирівняно. Просте використання координат без візуальної компоненти призводить до істотної втрати корисної інформації, тому для поліпшення системи в цих точках обчислюють різні дескриптори: LBP, HOG, SIFT, LATCH і ін. Після конкатенації дескрипторів і редукції розмірності отриманий вектор ознак можна використовувати для класифікації емоцій. Однак такий підхід вже вважається застарілим, так як відомо, що в області аналізу візуальних даних глибокі ЗНМ є найбільш ефективним рішенням.

Відповідно до [8, 52], на сьогодні існує досить велика кількість програмно-апаратних засобів, призначених для розпізнавання емоцій.

Комплекс **Affectiva Affdex** (www.affectiva.com) призначений для відстеження емоцій мобільними додатками. На відміну від більшості подібних програм, Affdex дозволяє здійснювати обробку даних безпосередньо на мобільному пристрої, а не на віддаленому сервері. Для розпізнавання 7 базових емоцій комплекс використовує власну базу даних, що складається з понад 5 млн. зображень обличчя. Робота комплексу полягає у запису зображення особи з подальшим вейвлет-перетворенням і порівнянням із зразками на основі SIFT-дескрипторів.

Подібні рішення використовується і в програмному комплексі **Pupeteer** (www.hikvision.com), призначеному для оцінки поведінки і емоційного стану учнів. Крім зазначеного функціоналу, комплекс дозволяє розпізнати такі рухи голови, як кивок, хитання, повороти з боку в бік. Для розпізнавання змішаних станів, коли не можна точно висловити, яка саме емоція переважає, використовується метод Байєсової фільтрації. В ході експерименту комплекс Pupeteer дозволив з 96-

відсотковою точністю визначити 6 базових емоцій.

Онлайн сервіс **CLMtrackr** (www.auduno.com/clmtrackr) дозволяє в реальному масштабі часу на основі зображення обличчя визначити 7 базових емоцій. Як сенсор використовується веб-камера. Функціонування сервісу базується на відстеженні 70 точок на зображенні обличчя, проведенні між ними ліній і порівнянні цих ліній із зразками різних базових емоцій.

Проєкт **Oxford** фірми Microsoft пропонує каталог API-функцій (azure.microsoft.com/ru-ru/services/cognitive-services), заснованих на алгоритмах комп'ютерного зору для розпізнавання 7 базових емоцій і нейтрального стану. Розпізнавання може проводитися як на одиночній, так і на груповій фотографії, що відповідає файлу розміром близько чотирьох мегапікселів. Розміри фотографії від 36x36 до 4096x4096 пікселів. Кількість зображень, що можуть бути одночасно проаналізовані, до 64. Емоцією, в термінах Oxford, є 8-мірний вектор, нормовані компоненти якого в сумі дорівнюють 1. За даними незалежного тестування точність розпізнавання складає близько 50%. Якість розпізнавання в значній мірі залежить від кута нахилу обличчя, освітленості і наявності перешкод. Найкраще розпізнаються зображення обличчя анфас або напів-анфас.

Сервіс **FaceReader** нідерландської компанії Noldus Information Technology (www.noldus.com/human-behavior-research/products/facereader) використовує метод Active Template для накладення на зображення обличчя деформуємого шаблону і Active Appearance Model для створення моделі ідентифікації особи за контрольними точками. Для класифікації використовуються нейронні мережі, навчені на тренувальному корпусі з 10000 кадрів. Визначаються 7 базових емоцій і нейтральний стан. Крім того, FaceReader здатний за зображенням обличчя людей визначати їх вік, стать і етнічну приналежність. Точність розпізнавання близько 75-80%. Невеликий нахил обличчя практично не впливає на точність розпізнавання. Разом з тим, FaceReader не дозволяє розпізнавати емоції дітей до п'яти років, класифікувати зображення обличчя в окулярах, детектувати повернене обличчя.

Також висуваються досить жорсткі вимоги до освітленості об'єкта, що має бути розпізнаний. Відзначається, що за один раз FaceReader здатний розпізнати тільки зображення обличчя однієї людини. Якщо на зображенні обличь декілька, вони розпізнаються послідовно, при цьому їх положення не повинно змінюватися.

Сервіс **Emovu** (www.eyeris.ai) компанії Eyeris (США) застосовується для визначення зацікавленості людини контентом в потоці відео. Для цього використовуються методи комп'ютерного зору, що базуються на використанні глибоких нейронних мереж. Перспективні розробки цієї компанії пов'язані з впровадженням САРЕС в автономні безпілотні транспортні засоби, де необхідно визначити відчуття, потреби пасажирів, а також динаміку станів людини в часі.

Проєкт університету Торонто **NuraLogix** (www.nuralogix.ai) дозволяє проводити розпізнавання прихованих емоцій на основі аналізу змін кольору підшкірного візерунка кровоносних судин зображення обличчя.

Система **eMotion Software** (www.visual-recognition.nl/eMotion.html), розроблена компанією Visual Recognition (Нідерланди), дозволяє розпізнати 6 базових емоцій (гнів, печаль, страх, здивування, відраза і щастя) і розрахувати їх інтегральне значення. Система створює 3D модель обличчя з виділенням на ній 12 ключових областей, таких як куточки очей та рота. Алгоритм розпізнавання не описаний.

Система **MMER_FEASy** - the FacE Analysis System (www.mmer-systems.eu/en/application-areas). Компанія-розробник MMER-Systems (Німеччина). У розробці використана методологія накладення на зображення обличчя певної деформуючої маски Active Appearance Model methodology, яка дозволяє розрахувати необхідні параметри в реальному часі. Програма розпізнає 6 базових емоцій, а також дозволяє визначити вік, стать і етнічну приналежність. Крім цього система може ідентифікувати особу.

Схожі характеристики властиві ЗАРЕС **FaceSecurity** (www.cognitec.com) компанії Cognitec (Німеччина). Задекларовані можливості системи:

- висока швидкість порівнянь аналізованого зображення з наявними зразками (на апаратному забезпеченні середньої потужності близько 900 000 порівнянь за одну секунду);

- використання широкої номенклатури засобів реєстрації, форматів зображень;

- інтеграція з сучасними системами управління базами даних родин Oracle, IBM DB2, MSSQL Server, що дозволяє ефективно використовувати великі бази даних образів.

До недоліків системи відносять можливість розпізнавання тільки фронтально розташованого обличчя (можливі відхилення до 15 градусів). Ще одним недоліком є висока чутливість до освітленості. Крім веб-додатків і SDK для мобільних телефонів, компанія Cognitec надає API для цифрових рекламних щитів - білбордів (billboard), з метою показу реклами для цільової аудиторії. Крім цього, системи Cognitec застосовуються в автомобілях для аналізу зображення обличчя водіїв і пасажирів, а також для виявлення позиції голови, виявлення розсіяного погляду, виявлення закритих очей.

Крім розглянутих вузькоспеціалізованих систем розпізнавання емоцій, відомий цілий ряд комплексів, функціональність яких також базується на аналізі геометрії зображення обличчя. Ці комплекси призначені для вирішення завдань:

- верифікація користувача на основі зображення обличчя (системи безпеки);

- визначення функціонального стану людини на основі характеристик окремих частин зображення обличчя (системи відеоспостереження);

- пошук певних людей у відеопотоці (системи пошуку);

- анімація і перетворення зображення обличчя з метою створення певних візуальних ефектів (системи морфингу);

- перетворення зображення обличчя в 3D-модель (системи моделювання);

- визначення раси, віку і статі людини (системи гендерної класифікації).

В результаті проведеного аналізу відомих ЗАРЕС на основі аналізу

зображення обличчя сформовано перелік додаткових параметрів, що характеризують особливості процесу розпізнавання, особливості їх технічної реалізації та сервісні (додаткові) можливості.

Короткий опис отриманих додаткових параметрів представлено в табл. 1.2.

Таблиця 1.2

Додаткові параметри оцінки ефективності ЗАРЕС за зображенням обличчя

№	Опис параметру
1	2
Особливості розпізнавання	
$B_{1,1}$	Можливість розпізнавання емоцій за фрагментом зображення обличчя
$B_{1,2}$	Стійкість до зміни освітленості
$B_{1,3}$	Стійкість до перешкод на зображенні обличчя
$B_{1,4}$	Можливість розпізнавання спотвореного обличчя
$B_{1,5}$	Стійкість до повороту зображення
$B_{1,6}$	Стійкість до нахилу зображення
$B_{1,7}$	Можливість аналізу групового зображення
$B_{1,8}$	Можливість аналізу кольорового зображення
$B_{1,9}$	Можливість автоматичного виділення контуру зображення
$B_{1,10}$	Можливість аналізу чорно-білого зображення
$B_{1,11}$	Можливість аналізу зображення у відтінках сірого
$B_{1,12}$	Можливість аналізу двомірного зображення
$B_{1,13}$	Можливість аналізу тривимірного зображення
$B_{1,14}$	Стійкість до зміни розміру зображення
Технічна реалізація	
$C_{1,1}$	Можливість інтеграції з СУБД
$C_{1,2}$	Можливість інтеграції з різними типами відеокамер

Продовж. табл. 1.2

1	2
$C_{1,3}$	Можливість аналізу різних форматів зображень
$C_{1,4}$	Можливість аналізу відеопотоку
Сервісні можливості	
$D_{1,1}$	Розпізнавання особи користувача
$D_{1,2}$	Перетворення двовимірного зображення особи в 3D-модель
$D_{1,3}$	Визначення раси
$D_{1,4}$	Визначення статі

Надалі перелік характеристик, показаний в табл. 1.2, може бути уточнений і доповнений.

Також слід зазначити, що за період 2017-2022 рр. опубліковано декілька нових БД, в яких містяться зображення обличь людей з різними емоціями.

Найбільш відомими є бази OMG-Emotion challenge, EmotiW challenge, AffectNet, AFEW-VA, EmotioNet challenge, EmoReact, Cohn-Kanade. У зазначених базах кількість представлених емоцій варіюється від 6 до 17, а кількість представлених зображень сягає декількох тисяч. Розробка цих БД, з одного боку, сигналізує про великий прогрес в області засобів розпізнавання емоцій, а з іншого боку, зумовлює можливості вдосконалення таких засобів.

Результати проведеного аналізу дозволили визначити значення наведених в табл. 1.1. параметрів ефективності для кожної із проаналізованих ЗАРЕС, що функціонують на основі аналізу зображення обличчя. Отримані оцінки для першої групи характеристик показані в табл. 1.3.

Акцентування уваги на першій групі характеристик продиктовано тим, що тематика дисертаційних досліджень пов'язана з розробкою ефективних моделей і методів розпізнавання емоцій.

Таблиця 1.3

Оцінка першої групи параметрів, що визначають ефективність ЗАРЕС за зображенням обличчя

Назва ЗАРЕС	**Оцінка параметру**					
	$A_{1,1}$	$A_{1,2}$	$A_{1,3}$	$A_{1,4}$	$A_{1,5}$	$A_{1,6}$
Affectiva Affdex	1	0	0	0,7	0,5	0,9
Pupeteer	1	1	1	0,9	0,5	1
CLMtrackr	1	0	0	0,9	0,5	1
Oxford	1	0	0	0,9	0,7	1
FaceReader	1	0	0	0,9	0,7	0,9
Emovu	1	1	1	0,8	0,7	1
NuraLogix	1	1	0	0,7	0,7	1
eMotion Software	1	1	1	0,9	0,8	1
MMER_FEASy	1	0	0	0,8	0,7	1
FaceSecurity	1	1	1	0,9	0,8	1

Як випливає з аналізу даних табл. 1.3, всі проаналізовані ЗАРЕС не пристосовані до розпізнавання емоцій на основі аналізу декількох БП, що унеможливлює їх застосування у випадку відсутності або неякісної відеореєстрації зображення обличчя слухача СДН. При цьому більшість відомих ЗАРЕС не розпізнає складні емоції, а також не дозволяють визначити поточний емоційний стан. Крім того, аналіз означених ЗАРЕС виявив їх низьку пристосованість до зміни рівня освітленості та до можливості розпізнавання емоцій на частково зашумлених зображеннях обличчя. Зазначимо, що не зважаючи на природу завад, їх негативний вплив полягає у тому, що на зображенні з’являються ділянки, котрі власне і спотворюють зображення. Для нівелювання такого впливу в роботах [11, 22, 35, 37] пропонується використовувати різні методи локальної фільтрації та сегментації зображень.

В роботах [22, 68] наведено критичний огляд таких методів фільтрації та сегментації. Відзначено, що в результаті фільтрації характеристики яскравості кожної точки цифрового зображення замінюються іншим значенням яскравості, котре в меншій мірі спотворене перешкодою. При цьому методи фільтрації розділяються на просторові та частотні.

Просторові методи фільтрації зображень передбачають представлення растрового зображення у вигляді двовимірної матриці, що дозволяє застосування до кожної точки зображення операції згортки. Так, в роботі [49] наведено спосіб очищення від шуму за рахунок застосування згладжуючих фільтрів для очистки зображень у відтінках сірого. Проведено порівняння ефективності різних типів згладжуючих фільтрів. За допомогою комп'ютерних експериментів показано, що для слабкозашумлених зображень найбільш ефективним є фільтр Гауса розмірністю 5×5. Відзначається, що просторові методи фільтрації можливо ефективно застосовувати для покращення якості зображень, однак їх застосування малоефективне для нівелювання завад, що викликані спотворенням частини зображення. В цьому випадку більш ефективними вважаються частотні методи фільтрації.

Класичні частотні методи фільтрації базуються на спектральному аналізі зображень з використанням дискретного перетворення Фур'є. Недоліки таких методів викликані обмеженнями методу Фур'є, що призводять до складності локалізації частотних показників спектру у часі [118].

Виправлення вказаного недоліку пов'язують із застосуванням методів фільтрації, що базуються на вейвлет-перетвореннях [109, 127]. Наприклад, в роботі [50] розглянута математична модель параметризації структури райдужної оболонки ока з використанням вейвлет-перетворення на основі похідних функції Гауса. Наведено результати експериментальних досліджень, що доводять ефективність запропонованої моделі.

В роботах [9, 50, 65, 154], що присвячені аналізу та опису технологій

використання вейвлет-перетворень в області кодування зображень, визначено перелік вимог, виконання яких забезпечує ефективність означеного процесу. Вказані вимоги співвідносяться з можливістю:

- кратномасштабного аналізу зображення,
- просторової локалізації,
- ортогональності,
- реалізації алгоритму обчислень з низькою обчислювальною складністю,
- застосування апробованих інструментальних засобів.

Таким чином, для забезпечення можливості ефективного нейромережевого розпізнавання емоцій на основі зображення обличчя необхідно за допомогою апарату вейвлет-перетворень забезпечити очищення зображення від типових нестаціонарних завад.

1.3. Аналіз технологій розпізнавання емоцій та особи за голосом

Проведений аналіз науково-практичних робіт, присвячених розробці засобів аналізу ГС, вказує на те, що традиційно виконувані ними завдання зводяться до розпізнавання окремих слів і особи диктора. Останнім часом коло цих задач істотно розширилось [200, 207]. В першу чергу додалася необхідність розпізнавання емоційного та функціонального стану диктора. Крім цього, вважаються перспективними напрямки визначення психічних відхилень, діагностування серцево-судинних захворювань, оцінки достовірності інформації диктора.

Різні аспекти проблеми аналізу ГС людини розглядались низкою вітчизняних і зарубіжних вчених [47, 59, 187]. Однак у відомих роботах при аналізі ГС акценти ставилися на підвищення ефективності розпізнавання слів і особи диктора [150, 310]. Інша проблематика, в силу своєї новизни, досліджена гірше. Разом з тим, дані [51, 66, 145, 201] вказують на те, що розпізнавання емоцій за голосом пов’язують з аналізом інтонації та тембру ГС. Зазначимо, що інтонація ГС є збірним поняттям,

що включає тональність, силу звучання, темп мовлення, наголоси і паузи. З точки зору розпізнавання емоцій, найбільш важливими складовими є тональність ГС та темп мовлення [143, 233, 285]. Тональність ГС відображає емоційно-вольову установку людини при досягненні конкретної мети. За типом впливу розрізняють три види тональності:

1. Мажорна, для якої характерні позитивні емоції, активна енергія, загальний оптимістичний настрій.

2. Мінорна, для якої характерні негативні емоції, слабка енергія, загальний песимістичний настрій.

3. Нейтральна - характеризується об'єктивністю викладу, гранично малою значимістю психологічної складової мовного спілкування.

Виділяють такі характеристики тональності, як спокій і напруженість. При цьому спокій - відносно рівний стан, а напруженість - нестійкий стан певної тональності, що характеризується накопиченням певної кількості сигналів іншої тональності, подальше збільшення якої здатне привести до якісного стрибка і переходу в іншу тональність.

Темп мовлення відображає швидкість проголошення різних елементів мови (звуків, складів, слів).

Тембр голосу - психоакустична характеристика, обумовлена формантною структурою спектра ГС, ступенем вираженості шумових компонентів, негармонійних обертонів, модуляційними процесами, зростанням і згасанням ГС. Як і тональність ГС, тембр голосу відображає яскравість звуку, його індивідуальність та багато в чому визначається поточним психологічним станом людини [143, 233]. В більшості відомих ЗАРЕС тональність та тембр ГС оцінюють на основі висоти основного тону та обертонів.

Як свідчать дані [129, 136, 205, 249], в основі функціонування поширених систем аналізу ГС покладені математичні моделі, що базуються на байєсовському підході, прихованих марківських процесах, машинах опорних векторів, методах

динамічного програмування, а також теорії НМ. При цьому і дані сучасних літературних джерел [93, 163, 184, 221], і результати дослідження найбільш відомих засобів аналогічного призначення (Google+, Microsoft Office, VoiceNavigator, Siri) дозволяють стверджувати, що найбільш широкі перспективи має нейромережева методологія аналізу ГС.

Разом з тим, в проаналізованій літературі немає єдиної думки про те, який вид НММ доцільно використовувати для вирішення задачі розпізнавання особи та розпізнавання її емоцій.

Також недостатньо повно висвітлено питання попередньої обробки ГС перед його подачею на вхід НМ. При цьому в теоретичних роботах [33, 34, 42] вказується, що обґрунтованість процедури визначення вхідних та вихідних параметрів, правильність вибору архітектури НММ є найважливішими факторами, що зумовлюють ефективність її застосування.

Результати проведеного аналізу літературних робіт [83, 209, 225, 227] дозволяють стверджувати про те, що в основі переважної більшості відомих систем автоматизованого та автоматичного розпізнавання емоційного стану людини покладена технологія розпізнавання емоцій, що базується на результатах обробки спектру ГС. Ці ж результати використовуються для розпізнавання особи людини за ГС [220, 247].

В більшості випадків йдеться про аналіз мел-кепстральних коефіцієнтів ГС, розрахунок яких передбачає виконання таких кроків [59, 72, 173, 323]:

- Вхідний дискретний ГС фільтрується. Для цього використовуються вирази:

$$z(n) = x(n) - 0{,}9x(n-1), \tag{1.1}$$

$$x(n) = z(n)\left(0{,}54 - 0{,}46\cos\left((i-6)\frac{2\pi}{180}\right)\right), \tag{1.2}$$

де $x(n)$ - дискретне значення амплітуди ГС при n-ій реєстрації.

- Розраховується спектр ГС. Для цього в більшості розглянутих робіт використовується метод дискретного перетворення Фур'є:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-i2\pi kn/N} \quad 0 \le k \le N, \tag{1.3}$$

де N - кількість реєстрацій ГС.

Для згладжування спектрограми на кордонах квазістаціонарних фрагментів до значень, отриманих за допомогою виразу (1.3), застосовується віконна функція Хеммінга:

$$X(k) = X(k)H(k), \tag{1.4}$$

де $H(k)$ - функція Хеммінга, що задана за допомогою виразу (1.5).

$$H(k) = 0{,}54 - 0{,}46\cos\left(\frac{2\pi k}{N-1}\right). \tag{1.5}$$

- Використовуючи вирази (1.6-1.10), розраховуються параметри фільтрів:

$$H(k) = 0{,}54 - 0{,}46\cos\left(\frac{2\pi k}{N-1}\right), \tag{1.6}$$

$$B(b) = 1125ln\left(1 + \frac{b}{700}\right), \tag{1.7}$$

$$B^{-1}(b) = 700\left(e^{\frac{b}{1125}} - 1\right), \tag{1.8}$$

$$f(m) = \frac{N}{FB}\left(B(f_1) + m\frac{B(f_h) - B(f_1)}{M+1}\right), \tag{1.9}$$

$$Q_m(k) = \begin{cases} 0, if\ k < f(m-1) \\ \dfrac{k - f(m-1)}{f(m) - f(m-1)}\ if\ f(m-1) \le k < f(m) \\ \dfrac{f(m+1) - k}{f(m+1) - f(m)}\ if\ f(m) \le k \le f(m+1) \\ 0, if\ k > f(m+1) \end{cases}, \tag{1.10}$$

де m - номер фільтра;

M - кількість фільтрів;

F - частота дискретизації.

- Далі, використовуючи вирази (1.11, 1.12), розраховуються значення мел-кепстральних коефіцієнтів:

$$S(m) = ln\left(\sum_{k=0}^{N-1} X(k)^2 Q_m(k)\right) \ 0 \le m \le M, \quad (1.11)$$

$$C(n) = \sum_{n=0}^{M-1} S(m)\cos\left(\pi n(m+0{,}5)\right) \ 0 \le n \le M, \quad (1.12)$$

де $C(n)$ - значення n-го мел-кепстрального коефіцієнта.

Також проведено аналіз найбільш відомих засобів розпізнавання емоцій людини за голосом.

Комп'ютерна система Вищої школи економіки (perm.hse.ru) дозволяє на основі аналізу ГС розпізнавати 8 базових емоцій – нейтральність, спокій, щастя, сум, гнів, страх, відраза та здивування [63]. Заявлена середня точність розпізнавання сягає 71%. При цьому точність розпізнавання нейтральності та спокою є більш високою, а точність розпізнавання страху, суму, здивування та відрази нижча середньої. Система носить експериментальний характер. Основою блоку дикторонезалежного розпізнавання є рекурентна НМ типу LSTM, що застосовується для аналізу ГС, поданого у вигляді набору мел-кепстральних коефіцієнтів. В системі передбачено фільтрацію білого шуму та вирівнювання гучності ГС.

Створена в Ізраїлі **технологія відслідковування емоцій в реальному часі Beyond Verbal** (www.beyondverbal.com) передбачає можливість визначення близько 400 варіантів емоційного стану диктора. Технологія використовується рядом компаній та передбачає використання різноманітних механізмів обробки ГС. Остаточна класифікація емоцій реалізується за допомогою спеціальних шаблонів, для створення яких було проаналізовано 2300000 аудіозаписів ГС із 170 країн світу. Технологія передбачає можливість як дикторозалежного, так і диктронезалежного розпізнавання емоцій.

Мобільний додаток Xpression розроблений компанією EI Technologies із

Великобританії (www.ei-technologies.com) здатен розпізнавати 5 емоцій людини: спокій, щастя, сум, гнів та тривогу.

Додаток має клієнт-серверну архітектуру. Клієнтський модуль використовується для реєстрації ГС та його попередньої обробки. На сервері аналізуються фрагменти ГС тривалістю 200 мілісекунд. Для цього використовується НМ, на вхід якої подаються параметри, що описують гучність ГС, його тональність та темп мовлення. Процес розпізнавання носить дикторозалежний характер, що передбачає створення шаблону конкретного користувача. Надалі додаток стежить за відхиленнями тембру голосу, його гучності, зміною проміжків часу між звуками і словами, та порівнює все це з означеним шаблоном. Класифікація здійснюється на основі розрахованих відхилень. В процесі розробки визначено, що представники різних народів говорять з різною швидкістю та гучністю. Крім того, ГС людей різного віку та різної статі має різні інтонації.

В роботах [9] наведено опис експериментальної **системи розпізнавання емоцій в людській мові** з використанням вейвлетів. Система призначена для розпізнавання в аудіозаписах 4 базових емоцій: гнів, страх, радість та норма. Наведено детальний опис математичного забезпечення системи. Розпізнавання реалізується на основі величин коефіцієнтів кореляції, для розрахунку яких використовується спектральний аналіз ГС на основі вейвлету Морле.

В роботі [51] описана **програма для вимірювання емоцій абонента** мобільного телефону. Результатом роботи програми є графік емоційної кривої абонента, що може бути використаний для розпізнавання його емоцій шляхом порівняння з емоційними кривими базових та складних емоцій.

Програма **EmoWatch** від японської компанії Smartmedical (eng.smartmedical.jp) призначена для ідентифікації і розпізнавання емоцій користувачів мобільних пристроїв компанії Apple. Програма базується на технології Empath, що аналізує основний тон, обсяг і швидкість ГС. Завдяки цій інформації, програма дозволяє аналізувати енергетичні рівні користувача і розпізнавати 4 види емоцій - гнів, спокій,

радість та сум. Особливістю програми є її дикторонезалежність та можливість розпізнавання незалежно від мови користувача.

Крім розглянутих систем розпізнавання емоцій, відомий цілий ряд комплексів, функціональність яких також базується на аналізі ГС. Ці комплекси призначені для вирішення таких завдань: верифікація користувачів, визначення функціонального стану людини, розпізнавання брехні, створення різноманітних голосових ефектів, визначення раси, віку та статі.

В результаті проведеного аналізу відомих засобів розпізнавання емоцій на основі ГС сформовано перелік додаткових параметрів, що характеризують особливості процесу розпізнавання, особливості технічної реалізації та сервісні можливості. Короткий опис отриманих додаткових параметрів представлено в табл. 1.4. Надалі перелік характеристик, показаний в табл. 1.4, може бути уточнений і доповнений.

Таблиця 1.4

Додаткові параметри оцінки ефективності ЗАРЕС за голосом

№	Опис параметру
1	2
Особливості розпізнавання	
$B_{2,1}$	Адаптованість до мови диктора
$B_{2,2}$	Адаптованість до статі диктора
$B_{2,3}$	Адаптованість до віку диктора
$B_{2,4}$	Можливість фільтрації шуму
$B_{2,5}$	Можливість нормалізації гучності ГС
$B_{2,6}$	Можливість розпізнавання в умовах багатоголосся
$B_{2,7}$	Дикторонезалежність процесу розпізнавання
$B_{2,8}$	Можливість аналізу спонтанного мовлення

Продовж. табл. 1.4

1	2
$B_{2,9}$	Необхідність використання навчальних баз даних, що містять приклади ГС з різними емоціями
Технічна реалізація	
$C_{2,1}$	Можливість інтеграції з СУБД
$C_{2,2}$	Можливість інтеграції з різноманітними системами реєстрації аудіопотоку
$C_{2,3}$	Можливість аналізу різноманітних форматів аудіофайлів
Сервісні можливості	
$D_{2,1}$	Розпізнавання особи користувача
$D_{2,2}$	Визначення віку
$D_{2,3}$	Визначення мови диктора
$D_{2,4}$	Визначення статі

Також слід зазначити, що за останні роки опубліковано декілька БД, в яких містяться аудіозаписи голосів людей з різними емоціями. Найбільш відомими є бази Emo-DB, RAVDESS, TESS, UMSSED, RUSLANA [228]. Наприклад, мовний корпус, що входить до бази даних Emo-DB (emodb.bilderbar.info), зібраний в технічному університеті Берліну та містить записи голосів професійних акторів на німецькій мові. База включає 535 записів ГС 10 дикторів (5 чоловіків, 5 жінок), що відтворюють набір дискретних 7 базових емоційних станів. Наявність цих БД, з одного боку, сигналізує про великий прогрес в області засобів розпізнавання емоцій на основі ГС, а з іншого боку, зумовлює можливості вдосконалення таких засобів. Крім того, аналіз доступних БД вказує на досить обмежений обсяг представлених в них ГС, що значно ускладнює розробку дикторонезалежних та текстонезалежних ЗАРЕС за голосом.

Результати проведеного аналізу дозволили за допомогою виділених параметрів

оцінити можливості проаналізованих ЗАРЕС за голосом. Оцінювання проведено аналогічно до оцінювання ЗАРЕС за зображенням обличчя. Акцентування уваги на першій групі характеристик продиктовано тим, що тематика досліджень пов'язана з розробкою ефективних моделей і методів розпізнавання емоцій, виражених в ГС. Отримані оцінки для першої групи характеристик показані в табл. 1.5.

Таблиця 1.5

Оцінка першої групи параметрів, що визначають ефективність ЗАРЕС за голосом

Назва ЗАРЕС за голосом	**Оцінка параметру**					
	$A_{2,1}$	$A_{2,2}$	$A_{2,3}$	$A_{2,4}$	$A_{2,5}$	$A_{2,6}$
Комп'ютерна система Вищої школи економіки	0,9	0	1	0,7	0,8	1
Beyond Verbal	0,9	1	1	0,8	0,7	0,7
Xpression	0,7	0	1	0,8	0,8	0
Система розпізнавання емоцій в людській мові	0,5	0	1	0,8	0,7	1
Програма для вимірювання емоцій абонента	0,5	0	1	0,8	0,7	1
EmoWatch	0,5	0	1	0,8	0,9	1

Як випливає з аналізу даних табл. 1.5, всі проаналізовані ЗАРЕС орієнтовані на розпізнавання емоцій виключно за ГС, що значно звужує сферу їх застосування в СДН. При цьому більшість відомих ЗАРЕС не дозволяють розпізнавати весь спектр базових емоцій та не пристосовані для розпізнавання складних емоцій. Крім того, аналіз літературних джерел [193-195] вказує на необхідність суттєвої модернізації ЗАРЕС в напрямку зменшення ресурсоємності, збільшення точності розпізнавання,

розширення номенклатури як базових, так і складних емоцій, що мають бути розпізнані, скорочення терміну розробки та пристосованості до застосування в умовах СДН. Визначено, що можливості модернізації тісно пов’язані з використанням в ЗАРЕС за голосом сучасних рішень в області нейромережевого аналізу ГС, що базуються на застосуванні ЗНМ та сучасних типів рекурентних НМ.

1.4. Аналіз технологій розпізнавання емоцій та особи за клавіатурним почерком

За останні декілька років засоби аналізу динаміки КП знаходять все більш широке застосування в ІС самого різного призначення. Це пояснюється високою спорідненістю КП з особою користувача, складністю підробки КП, можливістю реєстрації параметрів КП за допомогою тільки стандартного периферійного обладнання комп’ютера, а також широким застосуванням в ІС парольних і технологічних даних у вигляді набору символів [28, 122, 147]. Крім цього, доведено тісний зв'язок КП з фізіологічним і емоційним станом користувача [44, 77, 162]. Вважається, що кожній людині притаманні свої параметри КП, що характеризують швидкість і динаміку набору тексту, термін переходів між кількома (двома і більше) клавішами і помилками при наборі.

В основному засоби аналізу КП використовуються в контурах захисту інформації для біометричної аутентифікації користувачів та контурах прихованого моніторингу емоційного і фізіологічного стану операторів об'єктів критичної інфраструктури [33, 42, 85, 203]. Також відомі спроби впровадження таких засобів в СДН для оцінки рівня сприйняття слухачами навчальних матеріалів з метою підвищення автоматизації навчального процесу [14, 213, 215, 271]. Через те, що значення параметрів КП корелюються як з особою конкретного користувача, так і з його емоційним станом, завдання розпізнавання емоцій може вирішуватися спільно із завданням розпізнавання особи.

Загальноприйнято, що технологія такого розпізнавання включає в себе реєстрацію показників вводу символьних даних з клавіатури, фільтрацію і первинну обробку зареєстрованих показників та порівняння оброблених показників з еталонами, що співвідносяться з емоціями та особою користувача.

При використанні універсальних засобів клавіатурного вводу даних для визначення параметрів КП можуть бути використані зареєстровані коди натиснутих клавіш, символи, що відповідають натиснутим клавішам, час натиснення та час відпускання клавіш. При використанні спеціалізованих систем клавіатурного вводу для визначення параметрів КП додатково можуть бути використані параметри, що характеризують тиск користувача на клавішу, і швидкість/прискорення натискання.

В подальшому, після реєстрації, перед надходженням в систему розпізнавання зазначені параметри підлягають фільтрації і первинній обробці, в процесі якої можуть бути розраховані деякі узагальнені параметри КП.

Типова процедура фільтрації приведена в роботах [73, 130, 313] і полягає в застосуванні так званих частотних, часових і клавіатурних фільтрів. Часові фільтри призначені для фільтрації тривалих клавіатурних подій. Два інших типи фільтрів призначені для фільтрації подій, пов'язаних з клавішами, які рідко використовуються в процесі роботи. Описані фільтри можуть використовуватися як окремо, так і спільно.

Аналіз літературних джерел [251, 252, 261, 313] дозволяє стверджувати, що на сьогоднішній день не існує загальноприйнятої методології формування множини вхідних даних модуля розпізнавання КП. При цьому спільною рисою більшості відомих систем розпізнавання КП є використання в якості вхідних даних відфільтрованих значень терміну утримання окремих клавіш (ТУК) і терміну між послідовним натисненням двох клавіш (ТМК) [87]. Також можна визначити залежність номенклатури вхідних даних від математичного забезпечення процесу розпізнавання, який зводиться до порівняння вхідної послідовності з еталоном КП певної емоції користувача ІС.

Розрізняють два типи моделей розпізнавання:

- по заздалегідь визначеному фрагменту тексту;

- по фрагменту тексту довільного змісту.

В обох випадках для визначення еталонів КП користувачеві ІС необхідно кілька разів ввести один або кілька фрагментів одного і того ж тексту. Більшість відомих еталонів представляють собою статистичні моделі параметрів КП, засновані, наприклад, на нормальному або бімодальному законах розподілу [73, 74].

При розпізнаванні по визначеному фрагменту основу еталонів, як правило, складають показники ТУК та ТМК, які стосуються послідовного порядку натиснення клавіш. При розпізнаванні на основі фрагменту тексту довільного змісту еталони в своїй більшості формуються на основі статистичних показників ТУК та ТМК, окремих стійких послідовностей подій клавіатури, що відображають особливості динаміки набору тексту окремого оператора. Разом з тим, в роботах [122, 130, 312] на основі експериментальних досліджень доведена незадовільна адаптованість статистичних моделей до умов задачі аналізу КП та складність їх формування у випадку аналізу фрагменту тексту довільного змісту, а застосування марківських моделей обмежується складністю їх адаптації до нестаціонарного характеру динаміки КП [299, 301]. Також показано, що складність задачі розпізнавання КП пояснюється необхідністю аналізу достатньо великих обсягів багатовимірних даних.

При цьому в роботах [122, 162, 169, 182] доводиться, що, як раз з точки зору доведеної ефективності аналізу багатовимірних даних, перспективним напрямком вдосконалення систем аналізу КП є застосування НММ. Однак разом з позитивними результатами застосування НММ відзначається їх обмеженість, пов’язана з труднощами формування номенклатури вхідних параметрів та визначенням архітектурних параметрів у випадку необхідності аналізу КП при вводі довільного тексту [144, 203, 245]. Крім того, можна відзначити недоцільність використання застарілих НММ типу БШП, ймовірнісної НМ, карти Кохонена, мережі Хопфілда, мережі РБФ та повнозв’язних ГНМ з прямим розповсюдженням сигналу [298]. В той

же час, базуючись на методах розробки НМЗ захисту інформації, можна припустити доцільність використання ЗНМ для аналізу КП. Так, в роботах [215, 224, 317] запропонований спосіб перетворення даних при натисканні клавіш в зображення для подальшого використання в ЗНМ з метою аутентифікації користувачів. Відзначається можливість досягнути на текстах фіксованої довжини точності розпізнавання КП 96,8%. Схожі результати отримані також в роботах [182, 222], в яких ЗНМ використані для аналізу КП з метою підвищення стійкості парольного захисту користувачів. В роботі [306] запропоновано оригінальний метод збору статистичних даних, адаптований до задачі розпізнавання емоцій при вводі тексту на сенсорному екрані. Разом з тим, в означених роботах недостатньо повно висвітлені питання розробки методу застосування ЗНМ для аналізу КП з метою визначення особи та емоційного стану оператора ІС. Крім того, не знайдено опису процедури кодування параметрів КП до виду, придатного для застосування в ЗНМ.

Більш сучасні ЗАРЕС за КП описані в роботах [175, 181, 211, 219]. Так, в роботі [175] запропоновано **метод репрезентативного нейромережевого розпізнавання емоцій** користувача за рахунок аналізу КП при взаємодії із клавіатурою смартфону. Описано двоетапну процедуру навчання НММ типу БШП. Задекларовано точність розпізнавання 5 емоцій на рівні 81%. В роботі [219] наведено результати досліджень щодо впливу емоційного стану користувача на динаміку параметрів КП. Для визначення величини впливу використано **метод двосторонньої валентності**. Показано можливість розпізнавання збудженого/заторможеного емоційного стану користувача. Використання **лінійної регресивної моделі** для розпізнавання емоцій за КП описано в роботі [211]. **Метод розпізнавання емоційного контексту** на основі параметрів динаміки КП описано в роботі [181]. Наведено результати експериментів, спрямованих на визначення емоційного стану користувача, що виконує письмові завдання після перегляду емоційно-стимулюючих відеокліпів. Також варто відзначити **метод аналізу КП за допомогою LSTM** [317], яка дозволяє обробляти тексти довільної довжини. Однак побудова такої НММ пов'язана із

складністю формування навчальної вибірки та складністю визначення параметрів LSTM. Також в результаті проведеного аналізу відомих ЗАРЕС на основі КП сформовано перелік додаткових параметрів, що характеризують особливості процесу розпізнавання, особливості технічної реалізації та сервісні можливості. Короткий опис отриманих додаткових параметрів представлено в табл. 1.6.

Таблиця 1.6

Додаткові параметри оцінки ефективності засобів розпізнавання емоцій за КП

№	Опис параметру
Особливості розпізнавання	
$B_{3,1}$	Адаптованість до розпізнавання емоцій на текстах з довільним змістом
$B_{3,2}$	Адаптованість до розпізнавання емоцій на текстах довільного обсягу
$B_{3,3}$	Можливість фільтрації шуму, викликаного помилками при наборі тексту
$B_{3,4}$	Необхідність використання навчальних БД, що містять приклади параметрів КП з різними емоціями
Технічна реалізація	
$C_{3,1}$	Можливість інтеграції з СУБД
$C_{3,2}$	Адаптація до виду клавіатури
Сервісні можливості	
$D_{3,1}$	Розпізнавання особи користувача

Результати проведеного аналізу дозволили оцінити ефективність ЗАРЕС за КП. Оцінювання проведено аналогічно до оцінювання ЗАРЕС за зображенням обличчя. Отримані оцінки для першої групи характеристик, показані в табл. 1.7. Результати проведеного аналізу, відображені в табл. 1.7, свідчать про те, що відомі ЗАРЕС на основі КП недостатньо пристосовані для розпізнавання всієї номенклатури базових емоцій, не дозволяють розпізнавати емоції знеособленого

користувача та не пристосовані для інтеграції в ЗАРЕС на основі декількох БП. Як і для ЗАРЕС за голосом, вдосконалення ЗАРЕС на основі КП доцільно реалізувати на основі використання сучасних типів НММ, що ускладнюється через необхідність подачі параметрів КП у вигляді, придатному до нейромережевого аналізу.

Таблиця 1.7

Оцінка першої групи параметрів, що визначають ефективність ЗАРЕС за КП

Назва ЗАРЕС за КП	**Оцінка параметру**					
	$A_{3,1}$	$A_{3,2}$	$A_{3,3}$	$A_{3,4}$	$A_{3,5}$	$A_{3,6}$
Метод репрезентативного нейромережевого розпізнавання емоцій	1	0	1	0,5	0,9	0
Лінійна регресивна модель	1	0	1	0,5	0,9	0
Метод двосторонньої валентності	0	0	1	0,5	0,9	0
Метод розпізнавання емоційного контексту	0	0	1	0,5	0,9	0
Метод аналізу КП за допомогою LSTM	1	1	1	0,7	0,5	0,5

1.5. Аналіз технологій розпізнавання особи за додатковими біометричними параметрами

Як було відзначено в п.1.1, підвищення ефективності ЗАРЕС може бути реалізоване за рахунок врахування особливостей БП конкретного слухача [106]. Для цього доцільно реалізувати процедуру розпізнавання особи на основі додаткових БП, асоційованих з зображеннями. До вказаних БП відносяться

відбитки пальців, райдужна оболонка ока, сітківка ока та вушна раковина. Спільною рисою перерахованих БП є можливість їх реєстрації за допомогою відеокамери, що де-факто входить до стандартного апаратного забезпечення слухача СДН. Перспективність застосування в СДН технологій розпізнавання особи на основі аналізу відбитків пальців підтверджується результатами робіт [53, 60, 76, 180].

В роботі [76] розглядається проблема біометричної аутентифікації користувачів по відбитках пальців. Пропонується НММ аутентифікації на базі ДШП, у якого кількість вхідних нейронів 90, кількість схованих нейронів 10, а кількість вихідних нейронів 4. Відзначено, що зазначені параметри архітектури мережі розраховувалися експериментально. В якості вхідних параметрів використані геометричні характеристики локальних ознак: координата X, координата Y, вектор напрямку Q. Для формування вхідних параметрів запропоновано використовувати 30 локальних ознак. Експериментальним шляхом показано, що нейромережевий класифікатор відхилення вибору має рівень помилок першого роду 5,2%, а рівень помилок другого роду 0%. На підставі цього стверджується, що побудована НММ є ефективною по відношенню до інших біометричних технологій аутентифікації. Схожі результати представлені і в роботі [284], в якій додатково описаний математичний апарат, на якому базується функціонування ДШП.

Стаття [180] присвячена питанню використання НММ в системі ідентифікації відбитків пальців. Вказується, що труднощі ідентифікації викликані тим, що під час процесу сканування відбитку пальця зображення, що генерується сканером, може дещо відрізнятися під час кожного сканування. Для нівелювання цих труднощів пропонується використовувати НММ на базі БШП з одним схованим шаром. Вибір нейромережевої архітектури пояснюється її простотою і апробованістю. Джерелом вхідної інформації для НММ є відскановане зображення відбитка пальця розміром 188x240 пікселів. Надалі це зображення фільтрується і бінаризується. Після цього

обчислюється 12 геометричних моментів, кожен з яких відповідає одному з вхідних нейронів. Кількість вихідних нейронів прийнято рівним шести. Кількість схованих нейронів, що дорівнює 25, визначено, виходячи з теореми Колмогорова. Використано функцію активації типу сигмоїд. Мережа навчалася із застосуванням класичного алгоритму зворотного поширення помилки. Навчальна вибірка складалася зі 100 відбитків пальців. Кількість епох навчання - 1000. Заявлена точність розпізнавання навчальних прикладів склала 100%, що, на думку авторів, свідчить про перспективність НМС розпізнавання відбитків пальців.

В роботі [113] зазначається, що відбитки пальців різних людей можуть мати однакові глобальні ознаки, але абсолютно неможливою є наявність однакових локальних ознак (мінуцій). Тому процес ідентифікації особи пропонується розділити на два етапи. Першим етапом є класифікація відбитків пальців за глобальними ознаками для попереднього віднесення піддослідного відбитка до одного із наперед визначених класів. Другий етап полягає в розпізнаванні відбитка пальця на основі порівняння структури та коефіцієнта схожості мінуцій. Запропоновано алгоритми для класифікації зображень відбитків пальців за типами папілярних візерунків на основі застосування фільтра Габора, вейвлет-перетворення Хаара, Добеші і багатошарової НМ з прямим розповсюдженням сигналу. Проведено чисельні експерименти і представлені результати застосування запропонованих алгоритмів. Показано, що використання алгоритму, заснованого на спільному застосуванні фільтра Габора, п'ятирівневого вейвлет-перетворення Добеші і багатошарової НМ типу ДШП, надає можливість досягти достовірності класифікації близько 75%.

В роботі [24] проведено дослідження методів розпізнавання дактилоскопічних зображень на основі НМ типу ДШП з архітектурою 81-27-3. В якості вхідних параметрів НМ використані модуль і аргумент векторного поля градієнта дактилоскопічних зображень. Сформульовано висновок про необхідність збільшення потужності вхідного вектора НМ до 400.

Робота [112] присвячена аналізу характеристик відбитків пальців, які використовуються в біометричних НМС ідентифікації особи. Відзначається, що на зображенні поверхні пальця можна визначити досить велику кількість дрібних деталей (мінуцій), які є базою для класифікації. Мінуції характеризуються типом, координатами і напрямком.

У статті [72] детально розглянуті локальні і глобальні характеристики відбитка пальця, використовувані в біометричних системах аутентифікації. Показано, що можливості виділення характерних ознак, які в подальшому можна використовувати в цілях ідентифікації, в значній мірі залежать від якості зображення відбитків пальців. Також зазначено, що поширені сканери відбитків пальців забезпечують дозвіл 500 dpi, відсканований образ характеризується 256 рівнями яскравості, а максимальний кут відхилення відбитку від вертикалі зазвичай не перевищує 15 градусів. При цьому в якості характерних ознак відбитку пальця пропонується використовувати кінцеві точки, в яких «чітко» закінчуються папілярні лінії, а також точки розгалуження, в яких папілярні лінії роздвоюються. Відзначено, що в зображеннях поверхні пальця з дозволом близько 1000 dpi можна виявити деталі внутрішньої будови самих папілярних ліній (пори потових залоз) і використовувати вже їх розташування в цілях підвищення точності ідентифікації. Однак сучасний рівень технічного забезпечення поширених систем біометричної аутентифікації не надає можливості отримання зображень такої якості.

В роботі [229] розглянута технологія проєктування і експлуатації НМС дворівневого розпізнавання відбитків пальців. На першому рівні реалізується виділення характерних ознак, а на другому рівні реалізується аналіз виділених ознак, результатом якого є ідентифікація користувача. Показано, що на продуктивність системи розпізнавання впливає точність виділення інформативних ознак відбитків пальців, яка, в свою чергу, в значній мірі залежить від якості зображення. Цим пояснюється наявність в системі модуля підвищення якості початкового (отриманого зі сканера) зображення. Особливість функціонування

цього модуля полягає в тому, що поліпшення якості зображення відбитків пальців не повинно пошкоджувати мінуції. Для цього пропонується використовувати фільтри Габора і так звану скелетизацію, що перетворює сірий масштабний відбиток пальця в чорно-біле зображення, в якому товщина папілярних ліній становить один піксель.

У запропонованій системі НММ використані на обох рівнях розпізнавання. В обох випадках мова йде про ДШП, схований шар якого містить 200 нейронів, з функцією активації сигмоїдального типу, задану виразом виду:

$$y_n = \frac{1}{1 + e^{-\alpha x_{s,n}}} \tag{1.13}$$

$$x_{s,n} = \sum_{i=0}^{I} w_{i,n} X_i \tag{1.14}$$

де α – параметр крутизни;

y_n – вихідний сигнал n-го схованого нейрону;

$x_{s,n}$ – сумарний вхідний сигнал n-го схованого нейрону;

$w_{i,n}$ – ваговий коефіцієнт зв’язку між і-им вхідним та n-им схованим нейроном;

X_i – вхідний сигнал і-го вхідного нейрону.

При цьому кількість вихідних нейронів для ДШП, що використовується на першому рівні розпізнавання, дорівнює 5. Для ДШП, що використовується на другому рівні розпізнавання, кількість вихідних нейронів дорівнює 2. Навчання ДШП реалізоване з використанням апробованого алгоритму зворотнього поширення помилки.

Відзначено емпіричний характер визначення зазначених параметрів. Представлені результати чисельних експериментів показують точність розпізнавання близько 92%. Декларується, що перспективи подальших досліджень пов'язані з використанням ГНМ і розпаралелюванням обчислювальних процесів, пов'язаних з навчанням і розпізнаванням.

Стаття [314] присвячена порівнянню ефективності класифікатора відбитків пальців на основі сучасних НМ та ймовірнісного нечіткого класифікатора. Відзначається, що основною передумовою дослідження є задекларована перспективність використання ГНМ в області аналізу зображень. В якості базової моделі використаний БШП з трьома схованими шарами нейронів. Передбачене попереднє навчання БШП з використанням розрідженого автоендкодера. Отримані результати свідчать про те, що зміна кількості схованих нейронів в межах від 200 до 1250 практично не впливає на точність розпізнавання, яка становить приблизно 93%. При цьому класифікація, проведена за допомогою нечіткого класифікатора, показала більш високі результати, близько 97%. Це дозволило авторам статті сформулювати висновок про недостатню ефективність нейромережевих методів розпізнавання відбитків пальців.

Разом з тим, в статті відсутнє обґрунтування плану проведення експериментів, спрямованих на вивчення того, як зміна величин структурних параметрів НММ впливає на точність розпізнавання. Крім цього, викликає питання доцільності процесу попереднього навчання, який критично необхідний тільки в разі недостатньої кількості навчальних прикладів.

Статті [234, 248] присвячені питанням розпізнавання відбитків пальців за допомогою ЗНМ. При цьому робота [234] в основному орієнтована на доведення доцільності використання ЗНМ. Показано, що ЗНМ дозволяє значно скоротити обчислювальну ресурсоємність процесу розпізнавання. Скорочення обумовлено можливістю відмови від процедур попередньої обробки зображень відбитків пальців. При цьому за допомогою чисельних експериментів доведена можливість досягнення точності розпізнавання відбитка пальця на рівні близько 95%. В експериментах були задіяні ЗНМ, на вхід яких подавався відбиток середнього пальця розміром 224×224 пікселів.

Структура мереж передбачала наявність 19 шарів, з яких 5 шарів були згортковими. У схованих шарах використана функція активації типу ReLU. У

вихідному шарі використана функція активації Softmax, що дозволило інтерпретувати вихідний сигнал у вигляді ймовірності. Навчання проводилося на основі бази даних NIST SD4. В роботі [248] доводиться, що відносно традиційних методів аналізу, використання ЗНМ підвищує точність і швидкість розпізнавання. Крім цього, ЗНМ краще пристосовані для розпізнавання неякісних зображень відбитків пальців. В описаних експериментах використано декілька структурних рішень ЗНМ, на вхід яких подавалися зображення відбитків пальців розміром 227×227 пікселів. Види функцій активації ті ж самі, що і в роботі [234]. Відзначимо, що у всіх проаналізованих ЗНМ розмір ядра згортки становив 3×3. Всі інші параметри підбиралися емпірично.

Також проведено дослідження технологій розпізнавання особи на основі аналізу райдужної оболонки ока та сітківки ока. Визначено, що функціонування засобів розпізнавання райдужної оболонки ока передбачає першочергову фіксацію зображення ока з подальшим аналізом кольорової частини ока навколо зіниці. Цим засоби розпізнавання райдужної оболонки відрізняються від засобів розпізнавання сітківки ока, які орієнтовані на аналіз параметрів, що описують зображення сітки кровоносних судин всередині ока. Важливою перевагою засобів розпізнавання райдужної оболонки є можливість використання виключно пасивних засобів реєстрації параметрів райдужної оболонки без впливу на очі людини.

У засобах розпізнавання райдужної оболонки, які застосовуються в системах біометричної аутентифікації користувачів в якості засобів реєстрації, як правило, використовуються інфрачервоні відеокамери [165], а медичні засоби досить часто орієнтовані на відеокамери видимого діапазону світла [142]. У роботах [158, 197, 327] визначені основні типи характерних ознак райдужної оболонки ока.

Процес аналізу параметрів райдужної оболонки ока та сітківки можна умовно розділити на етапи попередньої обробки і власне розпізнавання [101, 147, 197]. Технологія попередньої обробки райдужної оболонки ока полягає в послідовному виконанні наступних кроків: аналіз якості зображення ока; виділення внутрішнього

і зовнішнього кордонів райдужної оболонки; геометричне перетворення виділеної райдужної оболонки для нормалізації зображення. Особливістю попередньої обробки сітківки ока являється необхідність виявлення кордонів зіниці ока та чіткого окреслення контурів судин ока. При заздалегідь сформованій номенклатурі параметрів, що підлягають аналізу, розпізнавання сітківки та райдужної оболонки ока полягає у визначенні величин вказаних параметрів, формуванні тестового образу і порівнянні цього образу з одним або декількома зразками.

У роботах [165, 256, 262] зазначено, що некоректність розпізнавання райдужної оболонки ока може бути викликана: різними умовами освітлення; використанням різних засобів відеореєстрації; перешкодами від вій і повік; дефектами зіниці; темним кольором райдужної оболонки, через що зіниця і райдужна оболонки зливаються; відблисками на райдужній оболонці; рухами і позиціонуванням голови; перешкодами від лінз і окулярів; змінами райдужної оболонки, які пов'язані з віком, хворобами, психофізіологічним станом людини. Загальновідомі шляхи подолання зазначених труднощів пов'язані як з підвищенням ефективності засобів відеореєстрації, так і з підвищенням ефективності обробки зареєстрованих параметрів. Так, в роботі [147] встановлено, що темний відтінок райдужної оболонки ока, характерний для жителів Південно-Східної Азії, ускладнює виділення кордону зіниці. Показано, що темні очі доцільно реєструвати в ближньому інфрачервоному спектрі, а світлі - у видимому діапазоні. В роботі [165] запропоновано спосіб компенсації високочастотного коливального руху очей за рахунок використання стробоскопічних інфрачервоних камер. Відомі підходи підвищення ефективності обробки зареєстрованих параметрів райдужної оболонки ока базуються на класичних методах цифрової обробки сигналів. В роботі [253] пропонується двоетапна технологія визначення інформативних ознак зображень райдужної оболонки ока. На першому етапі сегментації визначається область райдужної оболонки на зображенні. Для цього реалізується виділення областей, на яких немає відблисків та які не закриті віями і повіками. На другому етапі

параметризації реалізується обчислення інформативних ознак. В роботі [179] запропоновано алгоритм обробки зображення райдужної оболонки ока за допомогою багатомасштабних квадратурних вейвлет-перетворень. Аналіз алгоритмів обчислення порогових значень для бінаризації зображення наведено в [239]. В роботі [231] запропоновано метод параметризації райдужної оболонки на основі виділення ключових точок за допомогою методу перетворення Ерміта. За допомогою комп'ютерних експериментів показано, що помилка системи ідентифікації на основі зображень райдужної оболонки ока при використанні прикладів, представлених в БД CASIA-IrisV3, складає близько 0,23%. В роботі [135] обґрунтований метод виділення зіниці і зовнішнього кордону райдужної оболонки, який базується на детекторах краю і виділення кіл за допомогою перетворення Хафа. Показані обмеження методу, що стосуються зміни освітленості і ракурсу зображення райдужної оболонки.

Робота [238] присвячена нейромережевому аналізу зашумлених зображень райдужної оболонки ока, зареєстрованих відеокамерою видимого діапазону світла без додаткового освітлення. Досліджувалися перешкоди, викликані поворотом очей людини, а також розмитістю зображення. Визначено кроки попередньої обробки зображень, які полягають у розрахунку центру зіниці, нормалізації зображення і приведенні нормалізованого зображення до полярної системи координат. Для розпізнавання використовувався нейромережевий ансамбль, що складається з трьох ЗНМ типу модифікованої LeNet-5. Основні відмінності полягають у зміні кількості шарів згортки і кількості карт згортки. Описано експериментальні дослідження, спрямовані на розпізнавання райдужної оболонки, зображення яких представлені в БД CASIA і UBIRIS. Задекларовано, що досягнута похибка розпізнавання на рівні 3% дозволяє стверджувати про перевагу нейромережевого підходу розпізнавання райдужної оболонки.

Робота [179] присвячена розпізнаванню райдужної оболонки в умовах шуму. Розглянуто особливості викривлення зображення райдужної оболонки, викликані

перекриттям повіками, віями, відблисками світла, окулярами та погіршенням якості реєстрації за рахунок великої відстані до відеокамери, контактними лінзами, поганим освітленням. Запропоновано оригінальний метод виділення інформативних ознак з використанням комбінованого підходу із застосуванням локальної двійкової структури і матриці збігів рівня сірого. Проведені експерименти по розпізнаванню райдужної оболонки ока з використанням наборів даних CASIA і MMUiris. Для розпізнавання використаний класифікатор на базі чотирьохшарової ймовірнісної НМ. Задекларований рівень розпізнавання 96,5%.

Крім того, в статті [110] запропоновано алгоритм розмежування доступу за райдужною оболонкою ока. Для аналізу зображення запропоновано використовувати багатошарову нейромережеву модель з прямим розповсюдженням сигналу. Також в статті [309] описана система біометричної ідентифікації особи на основі розпізнавання райдужної оболонки з використанням штучних НМ. Розроблена система передбачає локалізацію області райдужної оболонки, нормалізацію, а потім і нейромережеве розпізнавання райдужної оболонки. Розпізнавання реалізоване з використанням ДШП, для навчання і тестування якого застосовувалися приклади з БД CASIA. Показано, що мережа чутлива до початкового розподілу вагових коефіцієнтів початкових ваг. Також визначено необхідність своєчасної зупинки навчання. Запропоновано швидкий алгоритм локалізації внутрішніх і зовнішніх кордонів області райдужної оболонки. Задекларована точність розпізнавання навчальних прикладів на рівні 93%. Описана можливість апаратної реалізації системи розпізнавання райдужної оболонки. Розпізнаванню райдужної оболонки ока за допомогою ЗНМ присвячена робота [241]. Показано, що сучасні алгоритми сегментації зображення ока недостатньо адаптовані до типових перешкод. Для вирішення проблем сегментації райдужної оболонки в складних ситуаціях за допомогою датчиків камери видимого та ближнього інфрачервоного світла розроблена ЗНМ IrisDenseNet, яка, на думку авторів, здатна визначити кордон діафрагми ока на зображеннях низької якості.

В роботі [262] розглянуто задачу побудови методу розпізнавання райдужної оболонки ока за допомогою ГНМ з прямим розповсюдженням сигналу, для навчання якої використовується алгоритм зворотного поширення помилки. Метод передбачає виконання етапів, асоційованих з локалізацією райдужної оболонки на зображенні ока, з нормалізацією, кодуванням і нейромережевим порівнянням. Нейромережевий класифікатор пропонується використовувати після нормалізації і виділення інформативних ознак. Відзначимо, що, незважаючи на декларацію про використання ГНМ, в статті при описі НММ наведена структурна схема ДШП з трьома нейронами в одному схованому шарі.

Роботи [147, 259] присвячені розробці нейромережевих систем біометричної аутентифікації користувачів на основі аналізу зображення сітківки ока. Доводяться переваги нейромережевих методів аналізу сітківки ока. В [147] розроблена ГНМ типу трьохшарового персептрону з сигмоїдальною функцією активації схованих нейронів, що застосована для аналізу попередньо оброблених двовимірних зображень сітківки ока. Показана перспективність застосування ЗНМ. В [259] для аналізу сітківки ока використано ДШП з 35 схованими нейронами, на вхід якого подавались вектори ознак, отримані шляхом перетворення сегментованого зображення сітківки в числові значення. Проведені комп'ютерні експерименти, спрямовані на визначення точності розпізнавання 139 осіб на основі аналізу зображень сітківки розміром 768x584 пікселів. Досягнута точність розпізнавання близько 97,5%. Таким чином, доведено, що використання нейромережевих методів дозволило як мінімум на 3% підвищити точність розпізнавання особи за сітківкою ока в порівнянні з іншими методами, що не базуються на теорії НМ.

Також в [147, 259] відзначено важливість реалізації якісного препроцесінгу зображення сітківки ока. Вказаній проблематиці присвячена робота [320], в якій показано, що типові завади для розпізнавання особи людини за сітківкою ока викликані деформацією відповідного зображення. Відзначено, що загальноприйнятому способу нівелювання таких деформацій притаманний

суттєвий недолік, який полягає у складності виділення ознак, необхідних для ідентифікації особи. Для виправлення вказаного недоліку пропонується використовувати метод попередньої обробки, заснований на вдосконаленому круговому просторово-анізотропному перетворенні Габора. Наведені результати експериментальних досліджень свідчать про суттєве зменшення кількості неінформативних ключових точок. Так, у випадку використання порогового значення 0,32, відсоток помилково визначених ключових точок не перевищував 0,1. Дослідження проводились на зображеннях сітківки ока, що представлені в БД VARIA (http://www.varpa.es).

Особливості технології розпізнавання особи за зображенням вушної раковини визначені на основі аналізу робіт [121, 149, 214]. Зроблено висновок про можливість розгляду процедури розпізнавання особи за зображенням вушної раковини у вигляді типової процедури біометричної аутентифікації на основі аналізу біометричних параметрів, асоційованих з зображеннями. Визначено, що в якості вхідних параметрів відповідних засобів розпізнавання використовуються параметри, які характеризують геометрію вушної раковини. Для цього зображення вушної раковини, що підлягає аналізу, подається у вигляді вектору заздалегідь визначених параметрів [128, 153]. Після цього для встановлення особи реалізується порівняння отриманого вектору з векторами-еталонами, що зберігаються в БД.

Розпізнавання особи може бути реалізоване на основі як двовимірної, так і тривимірної хмари точок, що відповідає зображенню поверхні вушної раковини. Також проведений аналіз дає підстави стверджувати про перспективність застосування нейромережевих засобів на основі ЗНМ. При цьому з точки зору використання нейромережевих методів розпізнавання, основними особливостями поставленого завдання є розмір і якість зображення вушної раковини, кількість і параметри характерних ознак вушної раковини, а також кількість еталонних векторів-ознак, які відповідають особам зареєстрованих користувачів [21, 42, 70].

Проведений аналіз засобів розпізнавання особи за відбитками пальців,

райдужною оболонкою ока, сітківкою ока та зображенням вушної раковини дозволив сформулювати висновок про можливість оцінки ефективності вказаних засобів за допомогою додаткових параметрів, що використовуються для оцінки ефективності ЗАРЕС на основі зображення обличчя. Крім того, доцільно додатково застосувати ряд параметрів ефективності, що дозволяють оцінити точність та ресурсоємність розпізнавання особи та враховують специфіку обробки кожного із додаткових БП.

Перелік та пояснення параметрів оцінки ефективності засобів розпізнавання особи наведено в табл.1.8. Відповідно до прийнятої класифікації, вказані параметри ефективності віднесено до групи параметрів В.

Таблиця 1.8

Параметри оцінки ефективності засобів розпізнавання особи

№	Опис параметру
Особливості розпізнавання	
$B_{1,15}$	Точність розпізнавання особи
$B_{1,16}$	Обчислювальна складність розпізнавання особи
$B_{1,17}$	Можливість визначення заданого переліку контрольних точок
$B_{1,18}$	Можливість скелетизації зображення
$B_{1,19}$	Можливість виділення кордонів райдужної оболонки ока
$B_{1,20}$	Можливість виділення кордонів зіниці ока

Також на основі проведеного аналізу технологій розпізнавання особи зроблено висновок, що основним напрямком їх адаптації до застосування в СДН є забезпечення ефективного аналізу БП, що можливо досягти як за рахунок підвищення ефективності обробки первинно зареєстрованих ознак БП, так і за рахунок застосування найбільш досконалих нейромережевих технологій. Оскільки джерелом інформації для аналізу таких БП є монохромні, напівтонові або кольорові

зображення, то напрямки подальших досліджень доцільно співвіднести з розробкою нейромережевих технологій на основі ЗНМ, що в останні декілька років вважаються найбільш ефективним засобом розпізнавання графічної інформації.

1.6. Постановка проблеми дослідження

Результати проведеного аналізу літературних робіт дозволяють стверджувати, що критичні недоліки переважної більшості сучасних ЗАРЕС, які ускладнюють їх застосування в СДН, пов’язані з:

- Низькою точністю розпізнавання номенклатури як базових, так і складних емоцій.

- Високою обчислювальною ресурсоємністю.

- Неможливістю розпізнавання емоцій на основі комплексного аналізу декількох БП.

- Низькою інтеграцією ЗАРЕС з засобами розпізнавання особи.

- Складністю розпізнавання емоційного стану.

Крім того, показано, що підвищення точності розпізнавання пов’язане з необхідністю вдосконалення технології попередньої обробки БП, яка в значній мірі залежить від виду зареєстрованого БП (зображення обличчя, ГС, КП, відбитку пальця, зображення вушної раковини, райдужної оболонки та сітківки ока).

Також обґрунтовано **гіпотезу** дослідження, що полягає у можливості забезпечення ефективного розпізнавання емоційного стану та особи слухача СДН за рахунок нейромережевого аналізу його БП, котрі можуть бути зареєстровані за допомогою загальнопоширеного АПЗ слухача СДН. Відповідно до сформованої гіпотези, на основі результатів аналізу досягнень теорії НМ [33, 64, 185, 189], по аналогії з відомими нейромережевими рішеннями в області захисту інформації [32, 42, 81], визначено шляхи забезпечення ефективного розпізнавання емоційного стану, що пов’язані з розробкою методологічного базису, використанням НММ,

найбільш ефективних в умовах поставленої задачі, та застосуванням сучасних підходів до формування вхідного та вихідного поля НММ, призначених для аналізу БП конкретного виду. Таким чином, науково-прикладна проблема дослідження, яка викликана об'єктивним протиріччям між потребою автоматизованого розпізнавання емоційного стану слухача СДН на основі аналізу його БП та відсутністю в СДН засобів такого аналізу, включає в себе п'ять наукових завдань.

Перше наукове завдання – побудова методологічного забезпечення технології автоматизованого розпізнавання емоційного стану слухача СДН на основі нейромережевого аналізу множини допустимих БП, реєстрацію яких можливо здійснити за допомогою поширеного АПЗ.

Друге наукове завдання – створення методів формування вхідного поля НММ аналізу множини допустимих БП для розпізнавання емоцій та особи слухача СДН, що забезпечують фільтрацію зареєстрованих значень БП та їх первинну обробку для подачі у вигляді, придатному до застосування в НММ.

Третє наукове завдання – створення методів нейромережевого розпізнавання емоцій та особи слухача СДН, що враховують можливість використання навчальних прикладів, сформованих з використанням методів формування вхідного поля НММ та з урахуванням вимоги щодо відображення в очікуваному вихідному сигналі близькості еталонів класів, та враховують необхідність визначення архітектурних параметрів ЗНМ.

Четверте наукове завдання – побудова методології нейромережевого розпізнавання емоційного стану та особи слухача СДН, яка, інтегруючи запропоновані методи нейромережевого розпізнавання, забезпечує розширення функціональних можливостей СДН за рахунок розпізнавання емоційного стану слухача в процесі навчання та підтвердження достовірності його особи.

П'яте наукове завдання – розробка архітектури системи нейромережевого розпізнавання емоційного стану та особи слухача СДН, що базується на побудованій методології нейромережевого розпізнавання та забезпечує можливість

створення відповідних засобів розпізнавання.

Рамкова модель дисертаційного дослідження показана на рис. 1.4.

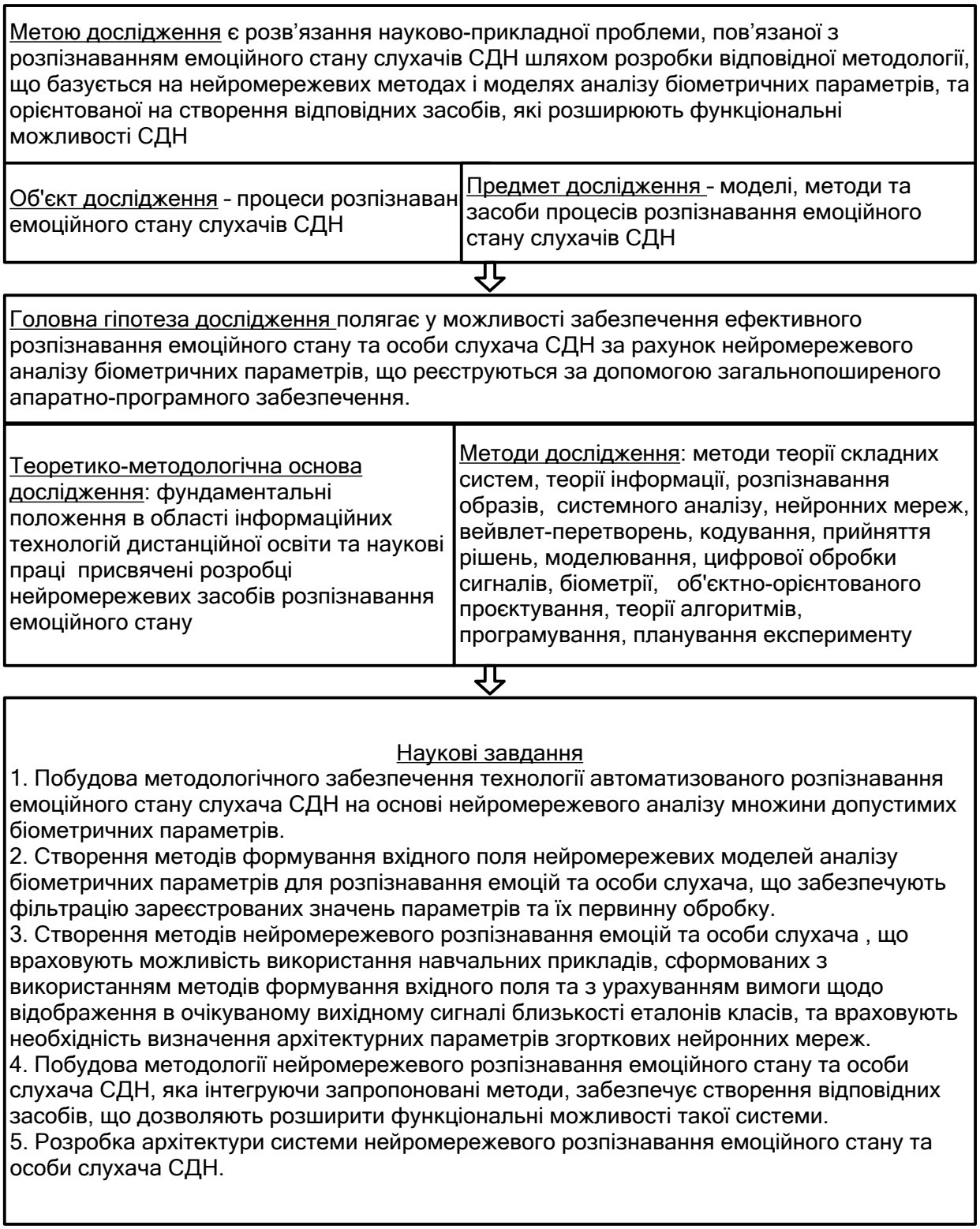

Рис. 1.4 Рамкова модель дисертаційного дослідження

Визначені наукові завдання формують проблемне поле розробки ЗАРЕС слухачів СДН на основі аналізу їх БП з урахуванням умов застосування в СДН та здійснення аналізу БП із застосуванням сучасних рішень в області НМ. Для вирішення вказаних наукових завдань запропоновано використати сучасні досягнення в області обробки та аналізу БП, що базуються на застосуванні експертного оцінювання, цифрової обробки сигналів, оптимізації, вейвлет-перетворень та теорії НМ.

Зазначимо, що хоча предметною областю даного дослідження є освітня сфера, однак отримані результати можуть використовуватися і в інших предметних областях, в яких необхідно вирішити завдання щодо розпізнавання емоційного стану на основі нейромережевого аналізу БП.

1.7. Висновки до першого розділу

Даний розділ присвячено аналізу можливостей впровадження технологій розпізнавання емоційного стану в систему дистанційного навчання. Основні результати розділу наступні:

1. Обґрунтовано доцільність впровадження в систему дистанційного навчання інформаційної технології автоматизованого розпізнавання емоційного стану слухача на основі нейромережевого аналізу зображення обличчя, голосу та клавіатурного почерку.

2. Показано перспективність інтегрування в інформаційну технологію розпізнавання емоційного стану засобів розпізнавання особи слухача системи дистанційного навчання на основі нейромережевого аналізу біометричних параметрів, що підлягають реєстрації за допомогою поширених комп’ютерних засобів. Крім можливості підвищення точності розпізнавання емоційного стану, застосування вказаних засобів дозволить забезпечити можливість підтвердження особи слухача при здійсненні навчального процесу.

3. Сформовано перелік базових та додаткових параметрів, що доцільно використовувати для оцінки ефективності засобів розпізнавання емоцій та засобів розпізнавання особи користувача системи дистанційного навчання.

4. З використанням запропонованих параметрів оцінки ефективності визначено, що застосування в системі дистанційного навчання відомих методів та засобів розпізнавання емоційного стану та особи слухача викликає необхідність їх вдосконалення в напрямку адаптації до очікуваних умов застосування та нівелювання завад, характерних для кожного із біометричних параметрів, які підлягають реєстрації.

5. Визначено недосконаліть методологічного та теоретичного базису розробки нейромережевих засобів, призначених для розпізнавання емоційного стану та особи слухача системи дистанційного навчання на основі комплексного аналізу його біометричних параметрів.

6. Визначена необхідність використання в інформаційній технології розпізнавання емоційного стану слухача засобів, призначених для обробки біометричних параметрів задля нівелювання типових завад та викривлень, що виникають в процесі реєстрації.

7. За результатами проведеного аналізу обґрунтовано постановку наукових завдань дослідження, що дозволило визначити мету, задачі, проблемне поле і методологічну базу дослідження та сформувати відповідну рамкову модель.

РОЗДІЛ 2
ЕЛЕМЕНТИ МЕТОДОЛОГІЧНОГО ЗАБЕЗПЕЧЕННЯ ТЕХНОЛОГІЇ АВТОМАТИЗОВАНОГО РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО СТАНУ СЛУХАЧІВ СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ

2.1. Концептуальна модель розпізнавання емоційного стану

Сучасні дослідження в області психології дозволяють стверджувати про системний характер психічних станів, який включає в себе реагування різних рівнів:

- психічного – переживання, зміна протікання психічних процесів;

- фізіологічного – гуморальна регуляція з боку ендокринної та вегетативної нервової системи;

- поведінкового – психомоторне реагування.

Своєрідність якісних змін на цих рівнях обумовлює специфічну характеристику певного психічного стану. Це дозволяє використовувати для діагностики функціонального стану людини параметри, які описують емоційний стан, що являє собою особливий вид психічних процесів або станів людини, які проявляються в переживанні будь-яких значущих ситуацій (радість, страх, задоволення), явищ і подій протягом життя.

Разом з тим, емоції – це почуття, що виражаються для пред'явлення оточуючим. Оскільки теорія і практика визначення вираження емоцій за рахунок використання теорії НМ досить відпрацьована, це зумовлює можливість їх використання для інструментальної (автоматичної) оцінки функціонального стану. Цей постулат підтверджується великою кількістю відповідних науково-прикладних розробок [1, 3, 5, 9]. При цьому у відомих СДН ЗАРЕС відсутні, хоча можливість їх впровадження підтверджується широким застосуванням програмних додатків, які базуються на нейромережевих моделях. Разом з тим, впровадження відомих нейромережевих ЗАРЕС у вітчизняні СДН викликає необхідність їх складної

адаптації до варіативності умов застосування, пов'язаних з терміном розробки, допустимою величиною похибки розпізнавання, наявним апаратним забезпеченням, ресурсоємністю створення і функціонування. Також недоліками поширених ЗАРЕС є повна або часткова відсутність опису математичного апарату, на якому вони базуються. Цим пояснюється актуальність науково-прикладної проблеми розробки нейромережевих ЗАРЕС, адаптованих до умов СДН.

У загальному випадку концептуальна модель представляє собою модель предметної області, що складається з переліку взаємопов'язаних понять, які використовуються для опису цієї області разом з властивостями і характеристиками, а також для класифікації цих понять за типами, ознаками прикладної області та закономірностями процесів, що в ній протікають [15, 54]. Концептуальна модель є відображенням концепції, під поняттям якої розуміють певний спосіб судження, трактування деяких явищ, основну точку зору, керівну ідею для їх систематичного висвітлення. Оскільки практичний результат розробки ЗАРЕС передбачає створення програмно-апаратного забезпечення для розпізнавання емоційного стану, то для визначення ефективності процесу розпізнавання доцільно використовувати визначення з області комп'ютерної та програмної інженерії. Крім цього, доцільно використовувати визначення з області теорії розпізнавання образів, яка лежить в основі ЗАРЕС, що проєктується. Відповідно до міжнародних стандартів цій галузі, ефективність можливо оцінити шляхом аналізу характеристик, які описують взаємозв'язок рівнів виконання апаратно-програмного комплексу, використання ресурсів (фінансові затрати, апаратне забезпечення, витратні матеріали) і послуг, які виконуються штатним обслуговуючим персоналом та ін. До основних характеристик ефективності АПЗ комп’ютерної системи відносять:

- оперативність – атрибут, що вказує на час відгуку, обробки і виконання функцій;

- ресурсоємність – атрибут, який визначає кількість і тривалість ресурсів, що використовуються в процесі виконання функцій програмної системи;

- узгодженість – атрибут, який вказує на відповідність даного атрибута заданим стандартам, правилам і приписам.

Згідно з наведеними визначеннями, на першому етапі створення концептуальної моделі було проведено гармонізацію термінології, що використовується в області нейромережевого розпізнавання емоцій на основі аналізу БП. Гармонізація проведена з позицій відображення сучасного стану науково-практичних досягнень в галузі розробки ЗАРЕС, адаптованих до використання в СДН.

В результаті визначені такі терміни:

- Емоції – це психічне відображення у формі безпосереднього, упередженого переживання життєвого сенсу явищ і ситуацій, обумовленого відношенням їх об'єктивних властивостей до потреб суб'єкта.

- Базові емоції – гнів, відраза, сум, страх, здивування, презирство, радість.

- Емоційний стан – це складний психічний стан відображення в формі переживання значимості впливу зовнішніх та внутрішніх стимулів. Складність емоційного стану полягає в тому, що один емоційний стан може включати в себе декілька емоцій. Крім того, тривалість емоційного стану більша, ніж тривалість емоції.

- Методологія розпізнавання емоційного стану слухачів СДН – система знань в області автоматизованого розпізнавання емоційного стану, підходів, моделей, методів, засобів та інформації, що використовуються для автоматизованого розпізнавання ЕС слухачів конкретної СДН [15, 16].

- БП – параметри, що визначають фізіологічні і/або поведінкові характеристики людини. До БП, що широко використовуються в сучасних засобах біометрії , відносяться геометричні параметри обличчя, очей, долонь людини, колір її обличчя і очей, акустичні параметри ГС, артеріальний тиск, пульс, ритми

головного мозку, тощо. В даному дослідженні акцент ставиться на БП, реєстрацію яких можливо здійснити виключно за допомогою поширених апаратних засобів: відеокамери, інфрачервоного сканера сітківки ока, мікрофону та сканеру відбитків пальців.

- Обличчя – передня частина голови людини, що зверху обмежена волосяним покровом голови, внизу – кутами та нижнім краєм нижньої щелепи, з боків – краями кордонів нижньої щелепи та основою вушних раковин.

- Відбиток пальця – візерунок папілярних ліній на поверхні епідерміса пальця.

- Мінуції – унікальні для кожного відбитку пальця ознаки, що характеризують зміну структури та орієнтації папілярних ліній. Мінуції описують закінчення, роздвоєння та розриви папілярних ліній.

- Райдужна оболонка ока – це забарвлене кільце в передній частині зіниці, що складається з м'язової та сполучної тканин і пігментних клітин, що змінює розмір зіниці ока.

- Зображення сітківки ока (сітківка ока) – це зображення візерунку кровоносних судин, що знаходяться на дні ока.

- ГС – звуковий сигнал, джерелом якого є голос людини. Діапазон частот ГС відповідає можливостям голосового апарату людини.

- КП – індивідуальна біометрична поведінкова характеристика людини, що визначає особливості набирання ним тексту з клавіатури.

- ДКП (динаміка клавіатурного почерку) – характеристика КП, що описує динаміку введення символів з клавіатури.

- Основний БП – це БП, аналіз якого дозволяє визначити емоційний стан та особу його власника. З позицій доведеної можливості ефективно визначати емоційний стан, до основних БП, які можливо зареєструвати за допомогою доступного апаратного забезпечення, відносяться обличчя людини, ГС та КП [45, 59, 198].

- Допоміжний БП – це БП, аналіз якого дозволяє визначити тільки особу його власника. До допоміжних БП, які можливо зареєструвати за допомогою доступного апаратного забезпечення, відносяться райдужна оболонка ока, відбитки пальців та сітківка ока [196, 197] .

- НМ – мережа штучних нейронів, що моделюють функціонування нейронів головного мозку людини і об’єднані між собою синаптичними зв’язками.

- Штучний нейрон – елементарний процесор, що за допомогою функції активації виконує перетворення інтегрального вхідного сигналу.

- Архітектура НМ – поняття, що включає в себе спосіб розповсюдження сигналів, метод навчання, тип функції активації штучних нейронів, структуру зв’язків в НМ.

- Структура НМ – поняття, що описує взаємозв'язки та взаєморозташування окремих частин НМ.

- НММ – модель, що інтегрує в собі опис архітектури НМ та типів штучних нейронів, що використовуються в даній НМ.

- Конструктивні параметри НММ – параметри, зміна яких впливає на ефективність НМ.

- Навчання НМ – процедура визначення вагових коефіцієнтів синаптичних зв’язків.

Також визначено, що з позиції вирішення проблематики даної наукової роботи, концептуальна модель призначена для формалізації причинно-наслідкових зв'язків, властивих процесу розпізнавання емоцій, визначених необхідністю підвищення ефективності СДН [41, 54]. В концептуальної моделі слід врахувати:

- умови функціонування нейромережевої ЗАРЕС, визначених характером взаємодії її окремих частин з компонентами СДН;

- необхідність реалізації ефективного застосування НММ для розпізнавання емоцій;

- необхідність реалізації ефективного застосування НММ для розпізнавання

особи слухача СДН;

- можливості управління ЗАРЕС і визначення її керованих змінних.

З огляду на загальноприйняту технологію застосування НММ, побудована показана на рис. 2.1 діаграма декомпозиції нейромережевого розпізнавання емоцій.

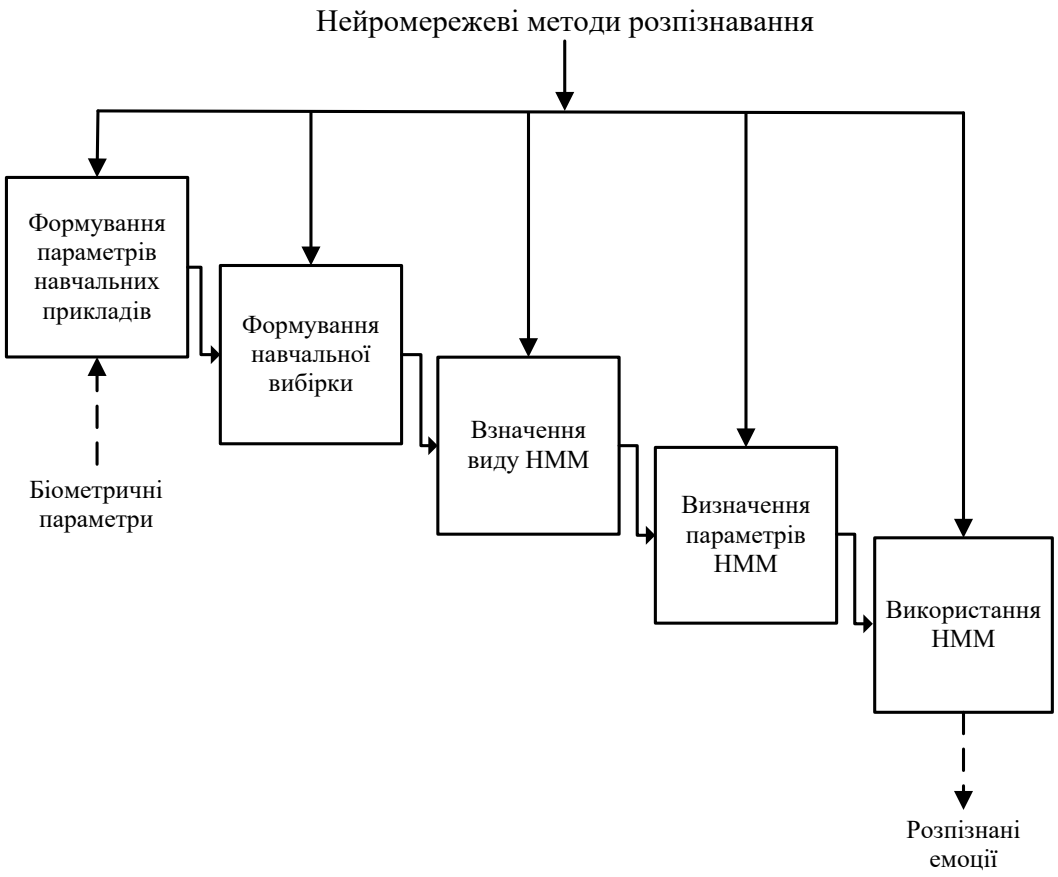

Рис. 2.1. Діаграма декомпозиції нейромережевого розпізнавання емоцій

Призначення складових частин цієї діаграми наступне:

- Формування параметрів навчальних прикладів – визначення для кожної з емоцій, що має бути розпізнана, множини вхідних і вихідних параметрів і способу їх приведення до вигляду, придатного для використання в НММ.

- Формування навчальної вибірки – визначення множини навчальних прикладів, які відповідають еталонам визначених класів емоцій. Кількість, якість і

номенклатура навчальних прикладів повинні бути достатніми для якісного навчання НММ.

- Визначення виду НММ – вибір такої нейромережевої архітектури, яка найбільш повно відповідає умовам завдання розпізнавання емоцій на основі заданих БП для конкретної СДН.

- Визначення параметрів НММ – обґрунтування величин структурних параметрів НММ і розрахунок значень вагових коефіцієнтів.

- Застосування НММ – розпізнавання емоцій на підставі множини наперед визначених зареєстрованих БП.

Наступним етапом створення концептуальної моделі стала розробка показаної на рис. 2.2 схеми компонентів нейромережевого ЗАРЕС.

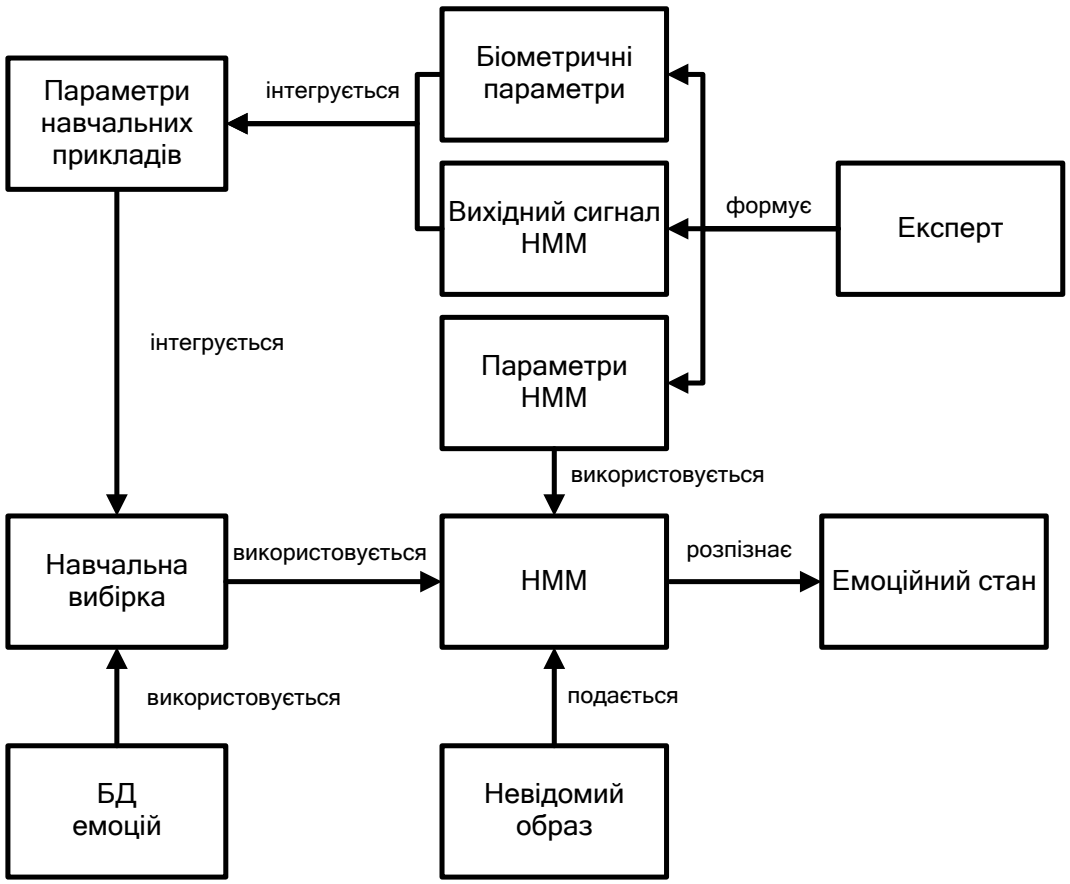

Рис. 2.2. Схема взаємодії компонентів нейромережевого ЗАРЕС в СДН

У схемі враховані особливості реалізації нейромережевої ЗАРЕС, пристосованої до СДН:

- обґрунтування вибору множини БП, на базі яких реалізується розпізнавання емоцій;

- складність визначення виду і параметрів НММ, що адаптовані до умов поставленої задачі розпізнавання емоцій;

- недосконаліть способів формування параметрів навчальних прикладів для НММ, призначених для розпізнавання емоцій.

Тому в схемі передбачена можливість формування параметрів навчальних прикладів і обґрунтування конструктивних параметрів НММ за допомогою експертів.

Аналіз рис. 2.1 і рис. 2.2 дозволив визначити, що ефективність процесу нейромережевого розпізнавання емоційного стану в СДН в основному залежить від ефективності ряду операцій, перелік яких наведено на рис. 2.3.

Слід зазначити, що, на думку авторитетних джерел [7, 11, 202], основним напрямком підвищення ефективності сучасних нейромережевих засобів аналізу БП для розпізнавання емоцій та особи власника є фільтрація типових перешкод, що виникають при їх реєстрації. Прикладами таких перешкод являються:

- при реєстрації обличчя людини – орієнтація голови відносно відеокамери, рівень освітлення, наявність сторонніх предметів, що перекривають зображення (окуляри, зачіска, захисна медична маска);

- при реєстрації відбитків пальців – орієнтація пальця відносно сканеру, забрудненість;

- при реєстрації райдужної оболонки ока – наявність відблисків, перекриття віями, динамічність розміру зіниці ока;

- при реєстрації сітківки ока – динамічність розмірів зіниці ока, зміна положення ока відносно сканеру;

- при реєстрації параметрів КП – «залипання» клавіш, особливості ходу

клавіш універсальної клавіатури, незаплановані перерви при введенні тексту;

- при реєстрації ГС – фоновий шум, музика, наявність одного або декількох голосів сторонніх осіб.

Рис. 2.3. Основні операції, що впливають на ефективність розпізнавання емоційного стану

В результаті встановлено, що в аналітичному вигляді концептуальну модель забезпечення ефективності нейромережевого розпізнавання емоційного стану в умовах СДН можливо записати за допомогою виразу:

$$H_{\Sigma} = f(H_D, H_A, H_R), \tag{2.1}$$

де H_{Σ} – інтегральна ефективність ЗАРЕС;

H_D – ефективність розробки НММ;

H_A – ефективність використання НММ;

H_R – ефективність реєстрації БП.

Деталізувавши компоненти (2.1) отримано такі вирази:

$$H_D = f(h_{D1}, h_{D2}), \tag{2.2}$$

$$H_A = f(h_{A1}, h_{A2}), \tag{2.3}$$

$$H_R = f(h_{R1}, h_{R2}, h_{R3}), \tag{2.4}$$

де h_{D1} – визначення виду НММ;

h_{D2} – визначення параметрів НММ;

h_{A1} – формування параметрів навчальних прикладів;

h_{A2} – формування навчальної вибірки;

h_{R1} – визначення множини БП, що підлягають реєстрації;

h_{R2} – фільтрація зареєстрованих БП;

h_{R3} – нейромережевий аналіз БП для визначення емоційного стану.

Компоненти виразів (2.1-2.4) відповідають процедурам, реалізація яких впливає на ефективність нейромережевих ЗАРЕС в СДН. Крім того, можна стверджувати, що ефективність нейромережевого розпізнавання БП доцільно оцінювати з точки зору ефективності процесу застосування нейромережевої ЗАРЕС і з точки зору навчання НММ [124] . При цьому показники ефективності повинні відображати тривалість, ресурсоємність і точність зазначених процесів.

Таким чином, обґрунтовані показані на рис. 2.4 показники оцінки ефективності нейромережевого розпізнавання емоційного стану. Слід зазначити, що при використанні даної концептуальної моделі для створення ефективних нейромережевих ЗАРЕС необхідно враховувати рівень розвитку нейромережевих засобів розпізнавання емоцій.

В першу чергу це стосується застосування в ЗАРЕС найбільш ефективних типів НММ. Крім того, слід враховувати обсяг необхідних та доступних

матеріальних ресурсів, що виділяються на розробку і впровадження засобів розпізнавання емоцій.

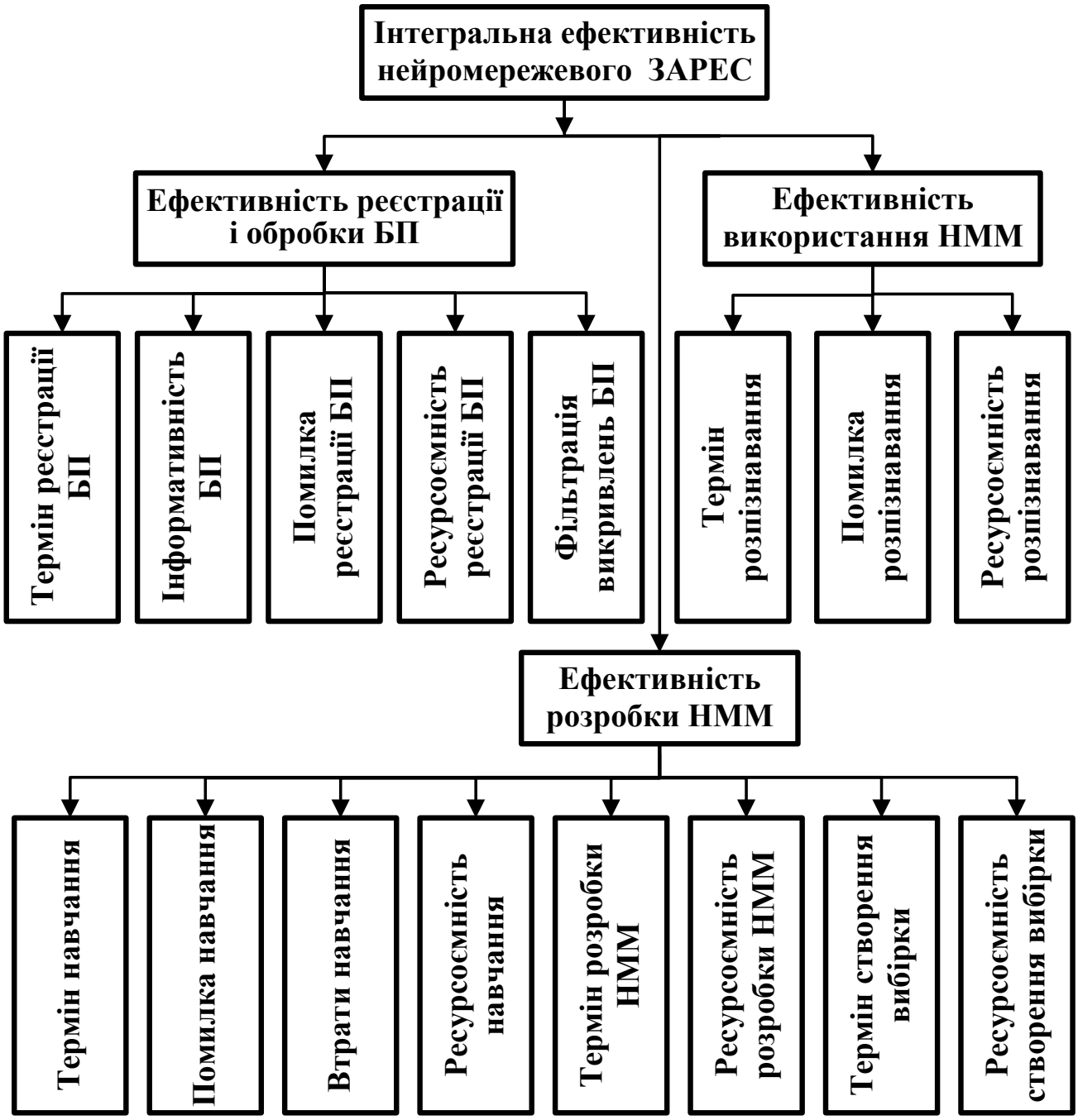

Рис. 2.4. Показники оцінки ефективності нейромережевих ЗАРЕС

Так, наприклад, використавши розроблену концептуальну модель [54],

визначено, що для впровадження в СДН ЗАРЕС за геометрією обличчя необхідно удосконалити існуючу методологічну базу нейромережевого розпізнавання емоцій в напрямку підвищення точності розпізнавання складних емоцій на повернутому зображенні обличчя та на зображенні обличчя, частково закритому типовими перешкодами.

Таким чином, в результаті проведених досліджень вперше розроблено концептуальну модель нейромережевого розпізнавання емоційного стану слухача СДН, яка за рахунок конкретизації операцій, що впливають на нейромережеве розпізнавання емоцій, а також за рахунок обґрунтування переліку параметрів оцінки ефективності, дозволяє деталізувати напрями досліджень з розробки відповідних засобів автоматизованого розпізнавання. Також визначено, що напрямки подальших досліджень пов'язані з удосконаленням методологічної бази розробки та використання нейромережевих засобів розпізнавання емоцій слухачів СДН.

2.2. Підходи до нейромережевого аналізу біометричних параметрів для розпізнавання емоційного стану

Аналіз побудованої концептуальної моделі розпізнавання емоційного стану слухачів СДН, дослідження відомих НМС, призначених для розпізнавання БП людини [4], а також результати теоретичних робіт в цій галузі [6, 8-21, 25] свідчать про те, що, не збільшуючи вимог до ресурсоємності апаратного забезпечення, підвищити ефективність означених НМЗ можливо тільки за рахунок модернізації методологічної бази. При цьому, дотримуючись рекомендацій [4, 14, 24], відправною точкою модернізації повинна стати раціоналізація підходів до нейромережевого розпізнавання ЕС в ІС загального призначення. Зауважимо, що в загальному випадку під терміном підхід до розгляду об’єкта дослідження розуміють основний спосіб сприйняття, оцінки дій, пов'язаний з певним баченням

реальності [1, 55]. Іншими словами, підхід представляє собою сукупність цілісно спрямованих принципів і методів, що визначають схему розуміння і пояснення певних аспектів об'єкта досліджень.

Базуючись на результатах [7, 10], можна стверджувати, що в області нейромережевого розпізнавання емоційного стану підходи до побудови НМС забезпечують:

- Доведення можливості застосування НМЗ для розпізнавання емоцій.

- Визначення принципів побудови НМС.

- Надання можливості використовувати апробовані НММ та нейромережеві методи.

- Адаптація конструктивних параметрів НММ до особливостей поставленої задачі.

Підхід до обґрунтування принципової можливості застосування НМЗ для розпізнавання емоцій. Доведення можливості застосування НМЗ для розпізнавання емоцій слухача СДН базується на даних теоретичних робіт [42, 57], в яких доведена можливість застосування НММ для апроксимації з заданою точністю довільної функції багатьох змінних. Базисом такого результату є теорема Колмогорова-Арнольда, яка свідчить про можливість представлення довільної багатовимірної неперервної функції у вигляді суперпозиції неперервних функцій однієї змінної. В інтерпретації [70] теорема Колмогорова-Арнольда стверджує, що довільна неперервна функція f, визначена на n-мірному одиничному кубі, може бути представлена як $2N+1$ суперпозицій неперервних та монотонних відображень одиничних відрізків:

$$f(x_1, x_2, \dots, x_N) = \sum_{n=1}^{2N+1} G_n \left(\sum_{k=1}^{N} \phi_{n,k}(x_k) \right). \tag{2.5}$$

Зліва в формулі (2.5) – довільна неперервна функція, визначена на багатовимірному кубі, праворуч – функції, визначені на відрізку [0; 1].

Застосування теореми Колмогорова-Арнольда в теорії НМ отримало назву теореми Хехт-Нільсена. В цій теоремі доводиться, що довільну неперервну функцію багатьох змінних можливо з наперед визначеною точністю представити за допомогою НММ з прямим розповсюдженням сигналу. Така НММ повинна містити як мінімум один схований шар нейронів з сигмоїдальними функціями активації. Кількість таких схованих нейронів повинна як мінімум на одиницю перевищувати кількість вхідних нейронів. Вихідний шар нейронів такої НММ може мати нейрони з лінійною або сигмоїдальною функцією активації. Відзначимо, що така НММ класифікується як ДШП, що є спрощеною версією БШП. При цьому результати досліджень Д. Парка та І. Сандберга [42], що також базуються на теоремі Колмогорова-Арнольда, доводять, що при достатній кількості схованих нейронів апроксимувати неперервну функцію також можливо за допомогою НММ типу радіальної базисної функції. Також слід відзначити, що ЗНМ, котрі застосовуються в даній дисертаційній роботі, є однією із модифікацій БШП, пов'язаною з побудовою топології зв'язків по аналогії з зоровим апаратом біологічного прототипу. При цьому в деяких типах ЗНМ вихідні шари нейронів представляють собою фрагмент радіальної базисної функції. Таким чином, можливо вважати, що ЗНМ як інтегральна модифікація ДШП та НМ типу радіальної базисної функції придатна для апроксимації неперервної функції багатьох змінних.

Оскільки процес розпізнавання емоцій та особи слухача зводить до нейромережевого аналізу зареєстрованих величин, що характеризують основні та допоміжні БП, то модель розпізнавання можливо записати у вигляді:

$$\exists\{x_n(t)\} \in \left\{x_n^{E,j}\right\} \overset{NN}{\Longrightarrow} E_j, \tag{2.6}$$

$$\exists\{x_n(t)\} \in \left\{x_n^{I,k}\right\} \overset{NN}{\Longrightarrow} I_k, \tag{2.7}$$

де $\{x_n(t)\}$ – множина зареєстрованих в момент часу t значень, що описують n-ий БП;

$\left\{x_n^{E,j}\right\}$ – множина значень n-го БП, характерних для j-ої емоції;

$x_n^{I,k}$ – множина значень n-го БП, характерних для особи k-го слухача СДН;

$\overset{NN}{\Longrightarrow}$ – оператор нейромережевого розпізнавання;

E_j – j-та емоція слухача СДН;

I_k – особа k-го слухача СДН.

При цьому елементи множини $\{x_n(t)\}$ відображають значення одного із основних або допоміжних БП, що зафіксовані з певною дискретизацією в t-ий момент часу. При достатній частоті дискретизації можливо стверджувати, що $\{x_n(t)\}$ співвідноситься з неперервною функцією, що визначена на скінченому інтервалі. Таким чином, співставлення виразів (2.5) та (2.6, 2.7) дозволяє стверджувати про принципову можливість розпізнавання емоцій та особи слухача СДН на основі нейромережевого аналізу основних та допоміжних БП, зафіксованих в певний момент часу.

Результати [3, 10, 11] вказують на те, що при створенні НМС розпізнавання емоцій в СДН, як і у випадку побудови універсальних НММ, слід використати два базових підходи. Перший базовий підхід пов’язаний з побудовою архітектури НММ, що визначає її структуру та спосіб поширення сигналу між нейронами. Найбільш апробованими класичними архітектурами є БШП, неокогнітрон, мережа радіальної базисної функції, мережа адаптивної резонансної теорії, імовірнісна нейронна мережа, мережі Кохонена, Хопфілда, Хеммінга і Коско. Останні чотири архітектури об'єднуються принципом створення рекурентних НММ, що передбачає зворотний зв'язок між нейронами. Решта архітектур побудовані відповідно до принципу створення НМ з прямим поширенням сигналу, що передбачає наявність тільки прямих зв'язків, за якими сигнал поширюється виключно в напрямку від вхідних нейронів до вихідних. Другий базовий підхід використовується при навчанні НММ. З позицій наявності в навчальних прикладах очікуваного вихідного сигналу, в межах цього підходу використовують три основних методи навчання: з учителем, без учителя та змішаний. У свою чергу, метод навчання з учителем можна

розділити з позицій механізму корекції вагових коефіцієнтів. Наприклад, найбільш відомий метод зворотного поширення помилки реалізує принцип навчання з урахуванням локального градієнта функції помилки. До основних методів навчання без вчителя відносяться машина Больцмана та конкурентне навчання. Також виділяють ітераційне навчання і навчання шляхом безпосереднього запам'ятовування даних.

В останній час при розробці НММ досить широко використовуються підхід до навчання глибоких НМ, підхід до навчання з підкріпленням, підхід до побудови ЗНМ, а також підхід до створення мереж довгої короткострокової пам'яті. Підхід до навчання глибоких НМ використовується для нівелювання процесу згасання градієнта в НММ на базі багатошарового персептрону з більш, ніж одним шаром схованих шарів, що зумовлює можливість здійснення ієрархічного процесу розпізнавання. Цей підхід передбачає:

- використання в нейронах схованих шарів функції активації типу ReLU;

- двоетапний режим навчання (спочатку самонавчання з використанням машини Больцмана, донавчання за допомогою алгоритму зворотного поширення помилки);

- набір допоміжних прийомів для поліпшення якості навчання.

Підхід до навчання з підкріпленням передбачає корекцію вагових коефіцієнтів в процесі взаємодії НММ з зовнішнім середовищем.

Підхід до побудови ЗНМ передбачає адаптацію структури БШП до задачі розпізнавання зображень та полягає у реалізації:

- локального сприйняття;

- розділення вагових коефіцієнтів;

- субдискретизації.

Локальне сприйняття означає, що на вхід нейрона подаються не всі виходи попереднього шару, а лише деяка їх частина, що дозволяє зберігати топологію зображення від шару до шару.

В ракурсі розділення вагових коефіцієнтів підхід припускає, що для великої кількості зв'язків використовується дуже невеликий набір вагових коефіцієнтів, які називаються ядрами згортки. Це значно зменшує ресурсоємність мережі. Суть субдискретизації полягає у зменшенні просторової розмірності зображення, що забезпечує інваріантність мережі до масштабу зображення.

Підхід до створення мереж довгої короткострокової пам'яті зводиться до того, що в мережу інтегрується рекурентний модуль, здатний запам'ятовувати значення як на короткі, так і на довгі проміжки часу. Зазначена можливість реалізується за рахунок використання в рекурентному модулі набору фільтрів, що регулюють процес запам’ятовування інформації. Таким чином, збережене значення не розмивається в часі, а при навчанні мережі за допомогою алгоритму зворотного поширення помилки негативний ефект згасання градієнту значно послаблюється.

Також в даному досліджені використано підхід до побудови капсульної нейронної мережі, а також підхід до пропорційного масштабування ЗНМ. Зазначені підходи спрямовані на підвищення стійкості ЗНМ до перешкод, які характерні при реєстрації зображень в ІС загального призначення. Відзначимо, що з позицій представленого дослідження, характерною рисою таких ІС є відсутність специфічних вимог, визначених у нормативній документації.

Крім наведених, розроблено ряд нових підходів, що за рахунок врахування особливостей поставленої задачі створення ЗАРЕС, забезпечили ефективність розробленої НМС розпізнавання емоційного стану слухачів СДН.

Використовуючи запропоновані показники ефективності, визначена необхідність того, що підходи до нейромережевого розпізнавання емоцій повинні забезпечувати ефективність механізмів вибору виду і параметрів НММ, вибору множини БП, що мають бути зареєстровані, а також розробки ефективної технології першочергової обробки зареєстрованих даних. Інтеграція цього твердження з результатами робіт [4-6] дозволила розширити перелік базових підходів до побудови НММ, використання яких забезпечує ефективність нейромережевого

розпізнавання емоційного стану в ІС загального призначення.

Підхід до визначення конструктивних параметрів. Вид та параметри нейромережевої архітектури повинні адекватно відповідати значущим умовам завдання розпізнавання емоцій.

Використовуючи вирази (2.8-2.11), в аналітичному вигляді сформульований підхід можливо записати так:

$$h_{D1}(v_i) \to max, v_i \in \{V\}_I, \tag{2.8}$$

$$h_{D2}(v_{max}, m_k) \to max, m_k \in \{M\}_K, v_{max} \in \{V\}_I, \tag{2.9}$$

де v_i – і-ий вид НММ;

$\{V\}_I$ – множина прийнятних видів нейромережевих архітектур;

I – кількість прийнятних видів НММ;

v_{max} – найбільш ефективний вид НММ;

m_k – *к*-та комбінація параметрів найбільш ефективного виду НММ;

$\{M\}_K$ – множина допустимих комбінацій параметрів.

Формування множини допустимих видів НММ доцільно здійснювати з позицій забезпечення ефективного навчання. Для цього необхідно, використовуючи наявні ресурси, в допустимі строки створити представницьку навчальну вибірку і реалізувати процес навчання. Тому допустимість виду НММ можна визначити за допомогою виразу:

$$if\ \left(q_j \in \{Q\}_J\right) \wedge \left(t_s(q_j) + t_e(q_j)\right) \to q_j \in \{V\}_I, \tag{2.10}$$

де $t_s(q_j)$ – термін формування навчальної вибірки;

$t_e(q_j)$ – термін навчання НММ;

$\{Q\}_J$ – множина апробованих видів НММ.

При реалізації тієї частини підходу, яка відповідає виразу (2.5), можливо використовувати результати [3], в яких сформовано перелік критеріїв ефективності, а також метод, що дозволяє визначити ефективність різних нейромережевих архітектур в різних умовах використання.

Відповідно до цього методу, для розрахунку ефективності і-го виду НСМ використовується вираз виду:

$$h_{D1}(v_i) = f(\{\alpha\}_N, \{R(v_i)\}_N), \quad (2.11)$$

де f – функція ефективності;

$\{\alpha\}_N$ – множина вагових коефіцієнтів критеріїв ефективності;

$\{R(v_i)\}_N$ – множина критеріїв ефективності;

$\alpha_n \in [0..1]$ – ваговий коефіцієнт n-го критерію ефективності;

$R_n(v_i) \in \{-1; 0; 1\}$ – значення n-го критерію для і-го виду НММ.

З огляду на можливість негативних значень критеріїв ефективності і небажаність від'ємного значення h_{D1}, для оцінки ефективності можливо використовувати функцію виду $y = 2^x$.

Це дозволяє записати вираз (2.11) у вигляді:

$$h_{D1}(v_i) = \sum_{n=1}^{N} \left(\alpha_n \times 2^{R_n(v_i)}\right). \quad (2.12)$$

Реалізація другої частини підходу до визначення конструктивних параметрів багато в чому залежить від виду НММ і при рівності інших умов в основному зводиться до визначення таких параметрів моделі, при яких досягається максимальна точність розпізнавання емоцій.

У задачі розпізнавання особи та емоцій слухача СДН на основі зображення БП, що співвідносяться з монохромним, напівтоновим або кольоровим зображенням та характеризуються за допомогою геометричних показників (обличчя людини, відбиток пальця, райдужна оболонка ока), ефективність ЗНМ вважається доведеною. У цьому випадку, за аналогією з [12, 23], адаптувати структурні параметри НММ пропонується виходячи з того, що процес розпізнавання емоцій повинен бути максимально наближений до свого біологічного прототипу. Під біологічним прототипом мається на увазі процес розпізнавання БП експертом. Наприклад, на першому рівні розпізнавання складних

геометричних об’єктів, що співвідноситься з першим згортковим шаром ЗНМ, експерт може виділити шість елементарних фігур, показаних на рис. 2.5.

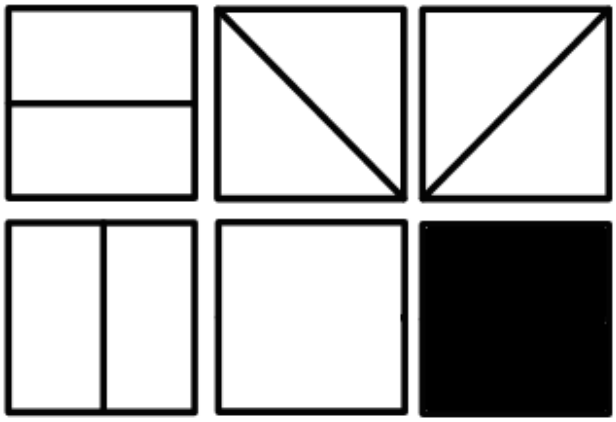

Рис. 2.5. Відображення елементарних фігур при розпізнаванні БП

Провівши експертне оцінювання БП, що співвідносяться з монохромним, напівтоновим або кольоровим зображенням, інтеграція сформульованої передумови з концепцією функціонування ЗНМ дозволила сформулювати наступну групу принципів адаптації ЗНМ до умов застосування:

Принцип 1. Кількість згорткових шарів нейронів повинна відповідати кількості рівнів розпізнавання БП експертом.

Принцип 2. Кількість карт ознак у n-му згортковому шарі має дорівнювати кількості ознак на n-му рівні розпізнавання.

Принцип 3. Карта ознак n-го шару, відповідна j-ій ознаці розпізнавання, зв'язується тільки з тими картами ознак попереднього шару, які використовуються для побудови зазначеної фігури.

Принцип 4. Розмір ядра згортки для n-го згорткового шару повинен дорівнювати розміру ознак, що розпізнаються на n-му ієрархічному рівні.

Принцип 5. Використання згорткових шарів не повинно спотворювати геометричні параметри ознак, що використовуються для розпізнавання БП.

Підхід до навчання. Для зменшення часу навчання НММ доцільно:

- в процедурі кодування очікуваного вихідного сигналу навчальних прикладів

передбачити можливість врахування близькості визначених еталонів образів;

- в архітектурі НММ передбачити можливість використання експертних даних.

Математичною інтерпретацією першої частини цього підходу є вираз:

$$y\left(u \in \{U_{E_i}\}\right): y\left(u \in \{U_{E_n}\}\right) \sim \chi(E_i, E_n), \quad (2.13)$$

де $y\left(u \in \{U_{E_i}\}\right)$ – очікуваний вихідний сигнал для навчальних прикладів, що описують і-ту емоцію, яка має бути розпізнана;

$\chi(E_i, E_n)$ – відстань між еталонами і-ої та n-ої емоції, що має бути розпізнана.

За аналогією з [22, 82], для розрахунку $\chi(E_i, E_n)$ пропонується використовувати НММ типу PNN, навчальні приклади якої містять очікуваний вихідний сигнал у вигляді назви класу, до якого вони належать. У той же час при розпізнаванні можливо в вихідному сигналі модифікованої мережі відобразити близькість вхідного образу до кожного класу, який був закладений в неї при навчанні. Друга частина підходу до навчання співвідноситься з відомими рішеннями в області навчання НММ за допомогою продукційних правил виду:

$$if\ (x_1 = L_1) \wedge (x_2 = L_2) \wedge \ldots (x_K = L_K) \rightarrow P(m), \quad (2.14)$$

де $x_1, x_2, \ldots, x_K$ – значення показників, що характеризують БП, який відповідає m-ій емоції, котра має бути розпізнана;

$L_1, L_2, \ldots, L_K$ – фіксовані величини;

$P(m)$ – ймовірність відображення m-ій емоції, яка має бути розпізнана.

Підхід до мінімізації ресурсоємності НМЗ. Для зменшення ресурсоємності і часу навчання НММ доцільно мінімізувати кількість вхідних параметрів НММ.

При розпізнаванні особи та емоцій слухача СДН на основі БП, що співвідносяться з монохромними, напівтоновими та кольоровими зображеннями, мінімізацію доцільно реалізувати, виходячи з можливості аналізу нормалізованих координат ключових точок, визначених у роботах [4, 5].

У випадку аналізу двомірного зображення обличчя слухача множину вхідних параметрів НММ можливо визначити так:

$$x_1 = a_1^{kp1}, x_2 = b_1^{kp1}, x_3 = c_{kp_1}, x_4 = a_2^{kp2}, x_5 = b_2^{kp2}, x_6 = c_{kp_2} \ldots, \quad (2.15)$$

де $x_1, x_2, x_3, \ldots$ – вхідні параметри НММ;

a_1^{kp1}, a_2^{kp2} ... – нормалізовані абциси координат ключових точок;

b_1^{kp1}, b_2^{kp2} ... – нормалізовані ординати координат ключових точок;

c_kp_1, c_kp_2 ... – закодоване значення кольору ключової точки.

У випадку аналізу тривимірного зображення множина вхідних параметрів розширюється за рахунок врахування глибини зображення:

$$x_i = a_i^{kp_i}, x_{i+1} = b_i^{kp_i}, x_{i+2} = c_i^{kp_i}, x_{i+3} = c_kp_i, i = 1 \ldots I\ , \quad (2.16)$$

де I – кількість ключових точок.

При формуванні вхідного поля ЗНМ, призначеної для аналізу КП, передбачено проведення експериментальних досліджень, спрямованих на обмеження переліку параметрів ДКП, що підлягають нейромережевому аналізу. У задачі розпізнавання особи і емоцій слухача на основі ГС передбачається мінімізувати кількість вхідних параметрів НММ за рахунок вилучення з цього сигналу незначущих спектральних складових. Для цього можливо використовувати зміну частоти основного тону ГС, який співвідноситься з першим мел-кепстральним коефіцієнтом кожного з квазістаціонарних фрагментів цього сигналу:

$$x_1 = mel_1^1, x_2 = mel_1^2, \ldots x_K = mel_1^K, \quad (2.17)$$

де mel_1^K – величина першого мел-кепстрального коефіцієнту для k–го квазістаціонарного фрагменту ГС;

K – кількість квазістаціонарних фрагментів ГС, що підлягають аналізу.

Крім того, для розпізнавання емоцій можливо використовувати декілька перших мел-кепстральних коефіцієнтів ГС, що характеризують декілька перших гармонік ГС. В цьому випадку вхідні параметри НММ визначаються так:

$$x_1 = mel_1^1, \ldots x_K = mel_1^K, x_{K+1} = mel_2^1, \ldots x_{2K} = mel_2^K. \quad (2.18)$$

Розрахунок мел-кепстральних коефіцієнтів доцільно проводити на так званих квазістаціонарних фрагментах ГС з використанням прямого перетворення Фур'є, відповідно загальноприйнятому алгоритму, наведеному в [236, 255, 268].

Крім того, для зменшення ресурсоємності навчання доцільно застосування НММ, які вже були навчені на прикладах з області розпізнавання емоцій або на прикладах з іншої прикладної області. Цей постулат співвідноситься з можливістю асоціативного переносу знань в НММ.

Таким чином, в результаті проведених досліджень розвинуто методологічний базис розробки НММ, призначених для нейромережевого аналізу БП слухачів СДН для розпізнавання їх особи та емоційного стану. Розвиток здійснено з позицій співвіднесення запропонованих підходів з основними факторами, які впливають на ефективність нейромережевого аналізу БП, що зумовлює можливість ефективного використання відповідних засобів в очікуваних умовах СДН вищих навчальних закладів. Розвиток методологічного базису реалізовано за рахунок обґрунтованого застосування ряду відомих та ряду запропонованих підходів до нейромережевого аналізу БП для розпізнавання емоційного стану. До запропонованих відносяться підхід до визначення конструктивних параметрів, навчання та ресурсоємності. Показано, що перспективним шляхом подальших досліджень є розширення зазначеного перелік підходів для впровадження найбільш ефективних рішень в області теорії НМ в засоби розпізнавання емоцій та особи слухачів СДН.

2.3. Моделі обробки біометричних параметрів, асоційованих з зображеннями

Проведені в першому розділі дослідження свідчать про те, що до БП, котрі асоціюються з зображеннями, можна віднести зображення обличчя, відбитки пальців, зображення вушної раковини, райдужну оболонка ока та сітківку ока. Прийнято, що аналіз зображення обличчя дозволяє розпізнати емоції та особу

слухача СДН, а аналіз інших зображень тільки особу. Зображення, які асоційовані з вказаними БП, показано на рис. 2.6.

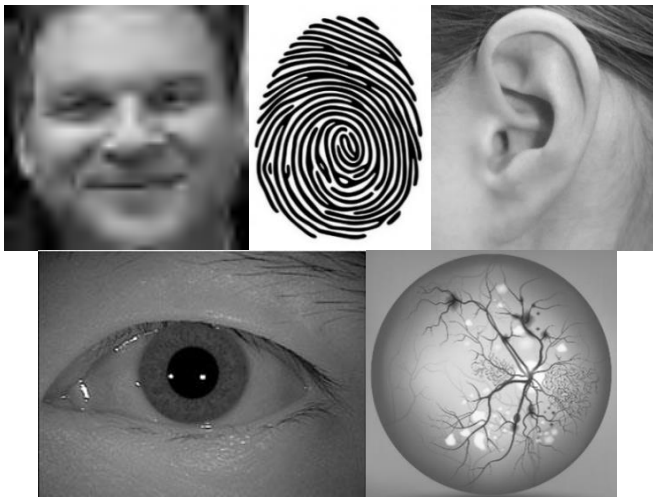

Рис. 2.6. Зображення, асоційовані з біометричними параметрами

Як свідчать результати першого розділу та дані [29, 30, 56, 114], враховуючи характеристики загальнодоступних комп'ютерних засобів відеореєстрації та можливості сучасних засобів нейромережевого аналізу зображень, доцільно встановити обмеження на максимально допустимі розміри зображень, асоційованих з вказаними БП. В базовому варіанті максимально допустимі розміри таких зображень в пікселях наведені в табл. 2.1.

Таблиця 2.1

Максимально допустимі розміри відсканованого зображення БП

Назва біометричного параметру	**Висота**	**Ширина**
Обличчя	1920	1080
Відбиток пальця	500	280
Зображення вушної раковини	480	270
Райдужна оболонка ока	320	280
Сітківка ока	320	280

Вказані зображення можуть бути зареєстровані у монохромному, напівтоновому або кольоровому вигляді. Оскільки вказані БП по суті представляють собою одно- або багатокольорові прямокутні зображення, то базуючись на результатах [80, 111], можна стверджувати, що для їх аналізу найбільш ефективним видом НММ являються ЗНМ.

Відзначимо, що у базовому випадку аналіз кольорового зображення $\boldsymbol{L}$ представляє собою матрицю, кожний елемент якої співвідноситься з чотирма коефіцієнтами, які і є вхідними параметрами ЗНМ:

$$\boldsymbol{l}_{x,y} = \{k_R, k_G, k_B, k_A\}. \tag{2.19}$$

Оскільки формат RGBA передбачає подання червоного, зеленого і синього кольору числами від 0 до 255, а альфа-каналу – числом від 0 до 1, то в початковому варіанті кодування вхідних параметрів ЗНМ можливо реалізувати за допомогою виразів виду:

$$k_z = {}^{q_z}/_{255}\ , z \in \{R, G, B\}\ , \tag{2.20}$$

$$k_A = q_A. \tag{2.21}$$

Зазначимо, що при реєстрації кольору зображення у форматі RGB кодування вхідних параметрів ЗНМ реалізується лише з використанням виразу (2.20).

У спрощеному випадку при аналізі напівтонового зображення кожен елемент матриці $\boldsymbol{L}$ співвідноситься тільки з одним коефіцієнтом, що є вхідним параметром ЗНМ:

$$l_{x,y} = k_M. \tag{2.22}$$

В цьому випадку кодування вхідних параметрів ЗНМ деталізується так:

$$k_M = {}^{q_M}/_{255}\,. \tag{2.23}$$

Оскільки серед технологій комп’ютерного зору, що використовуються для аналізу зображень, асоційованих з БП, найбільш апробованою є технологія розпізнавання обличчя, то саме модель обробки зображення обличчя використана в якості прототипу інших моделей обробки БП, асоційованих з зображеннями.

Модель обробки зображення обличчя. Основним призначенням вказаної моделі є забезпечення теоретичного підґрунтя створення методу формування вхідного поля ЗНМ, що використовується для розпізнавання емоцій та особи слухача СДН. Розмір вхідного поля такої ЗНМ може змінюватись від 1920×1080 до 29x29 пікселів. При цьому розмір зареєстрованого зображення обличчя визначається роздільною здатністю відеокамери або розміром зображень, що знаходяться в БД навчальних прикладів. Особливості відеореєстрації слухачів СДН обумовлюють необхідність проведення центрування та масштабування зареєстрованого зображення. Базуючись на результатах [137, 282], прийнято, що центрування та масштабування зареєстрованого зображення доцільно реалізувати на основі виділення зовнішнього кордону зображення обличчя, котре в першому наближенні здійснюється за допомогою класичного алгоритму Віоли-Джонса [115]. Власне ітераційний пошук координат точки центру зареєстрованого зображення здійснюється до виконання умови, що може бути записана за допомогою виразів:

$$\sum_{m=1}^{M} \frac{v_m(x_m, y_m) - c(x_c, y_c)}{\|v_m(x_m, y_m) - c(x_c, y_c)\|} = 0, \tag{2.24}$$

$$F_1 = \sum_{m=1}^{M} \frac{v_m(x_m, y_m)}{\|v_m(x_m, y_m) - c(x_c, y_c)\|}, \tag{2.25}$$

$$F_2 = \sum_{m=1}^{M} \frac{1}{\|v_m(x_m, y_m) - c(x_c, y_c)\|}, \tag{2.26}$$

$$c(x_c, y_c) = \frac{F_1}{F_2}, \tag{2.27}$$

де v_m – яскравість m-ої точки, що входить до складу зображення;

M – кількість точок зображення;

x_m, y_m – координати m-ої точки зображення;

c – яскравість центру зображення;

x_c, y_c – координати точки центру зображення.

Зазначимо, що вираз (2.27) є модифікацією виразу (2.24), пристосованою для використання алгоритму Вайсфельда, що дозволяє забезпечити обчислювальну ефективність програмного забезпечення для пошуку координат центру.

В першому наближенні для спрощення розрахунків координати центру монохромного зображення можливо розрахувати так:

$$x_c = \frac{1}{M}\sum_{m=1}^{M} x_m, \tag{2.28}$$

$$y_c = \frac{1}{M}\sum_{m=1}^{M} y_m. \tag{2.29}$$

Для півтонового та кольорового зображення дані вирази дещо видозмінюються:

$$x_c = \frac{1}{M}\sum_{m=1}^{M} (v_m x_m), \tag{2.30}$$

$$y_c = \frac{1}{M}\sum_{m=1}^{M} (v_m y_m). \tag{2.31}$$

Визначення осі симетрії в базовому варіанті пропонується реалізувати на основі ітераційного підходу [21, 25, 29]. Гіпотеза, на основі якої побудовано даний підхід, полягає в тому, що вісь симетрії монохромного зображення проходить через його центр, визначений за допомогою (2.24-2.31). Знайдена точка приймається за початок відліку.

Після визначення центру зображення наступний етап алгоритму пов’язаний з побудовою прямих ліній, що описуються виразом $a_0 x + a_1 y + a_3 = 0$ та проходять через центр зображення і через кожну точку контуру зображення. Далі відносно цієї прямої проводиться дзеркальне відображення всіх точок контуру.

У випадку $a_0 x + a_1 y + a_3 < 0$ ці точки включаються до множини дзеркально

відображених точок $\{A\}$, в протилежному випадку – до множини $\{B\}$. Після цього розраховується ступінь схожості відзеркаленої та базової частин контуру:

$$\mu_i = \sum_i \min_j \left\{ \left(x_i^B - x_j^A\right)^2 + \left(y_i^B - y_j^A\right)^2 \right\}, \tag{2.32}$$

$$\mu_i = \sum_i \min_j \left\{ \left(v_i^B x_i^B - v_j^A x_j^A\right)^2 + \left(v_i^B y_i^B - v_j^A y_j^A\right)^2 \right\}, \tag{2.33}$$

де $x_i^B,\ y_i^B$ – координати i-ої точки, що входить до множини $\{B\}$;

$x_j^A,\ y_j^A$ – координати j-ої точки, що входить до множини $\{A\}$;

v_i^B – інтенсивність кольору i-ої точки, що входить до множини $\{B\}$;

v_j^A – інтенсивність кольору j-ої точки, що входить до множини $\{A\}$.

Вираз (2.32) використовується для монохромних зображень, а вираз (2.33) – для напівтонових або кольорових.

Для розрахунку коефіцієнта збігу осі симетрії з лінією, що проходить через точку із множини $\{B\}$, використовується вираз виду:

$$\delta_i = 1 - \frac{\mu_i}{\max_i \mu_i}. \tag{2.34}$$

При $\delta_i = 1$ дана лінія являється віссю симетрії, а при $\delta_i = 0$ збіг відсутній.

Після розрахунку (2.34) для кожної точки із множини $\{B\}$ обираються точки-кандидати, через які може проходити вісь симетрії зображення. Умова віднесення i-ої точки до множини точок-кандидатів описується наступним виразом:

$$\delta_i > \min_i \delta_i + 0{,}02. \tag{2.35}$$

Завершальною стадією розрахунків є вибір серед точок-кандидатів точок, що лежать на шуканих осях симетрії. Вибір реалізується за рахунок видалення із визначеного околу кожної точки-кандидата точок, для яких

$$\delta_w < \delta_i, \tag{2.36}$$

де δ_w - коефіцієнт збігу для w-ої точки, що знаходиться в околі точки-кандидата.

В загальному випадку радіус околу видалення залежить від числа точок в

контурі. В даному випадку прийнято:

$$k = 0{,}1M + 1. \tag{2.37}$$

Після визначення координат центру та осі симетрії за необхідності до зображення обличчя можуть застосовуватись описані в [21, 25, 29] операції центрування та масштабування.

В залежності від характеристик доступних технічних засобів реєстрації та від особливостей наявних навчальних прикладів, нейромережевому аналізу підлягають або монохромні, або напівтонові, або кольорові зображення.

Первинна обробка цих зображень, як правило, полягає у застосуванні лінійного фільтру, призначеного для корекції яскравості пікселів:

$$R(x, y) = \sum_{i=-a}^{a} \sum_{j=-b}^{b} \big(B(x + i, y + j) \times M(x + i, y + j)\big), \tag{2.38}$$

$$a = \frac{l}{2}, \tag{2.39}$$

$$b = \frac{h}{2}, \tag{2.40}$$

де R – відфільтрований фрагмент зображення;

B – початковий фрагмент зображення;

M – фільтр;

l – ширина фільтру;

h – висота фільтру.

Алгоритм розрахунку параметрів фільтру наведено в [21, 29, 321] та визначається особливостями відеореєстрації. Якщо первинна обробка зображень реалізується для підвищення контрастності, то використовуються фільтри, в яких значення в точці (0, 0) більше ніж 1, а сума всіх значень дорівнює 1. Наприклад,

$$M_{k1} = \begin{pmatrix} -1 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & -1 \end{pmatrix}, \tag{2.41}$$

$$M_{k2} = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{pmatrix}, \quad (2.42)$$

де M_k – фільтр для підвищення контрасту зображення.

Для виправлення так званих крайових ефектів зображення доповнюється пікселями, колір яких визначається шляхом екстраполяції. Також слід зазначити, що застосування фільтрів виду (2.41, 2.42) не змінює середню інтенсивність зображення, оскільки сума елементів фільтру дорівнює одиниці. Для остаточного виділення кордонів зображення обличчя та кордонів об'єктів, що знаходяться на такому зображенні, використано алгоритм Кенні, що базується на виявленні перепадів яскравості. Для такого виявлення застосовується процедура фільтрації Собеля. Застосування фільтру Собеля пояснюється простотою обчислень та достатньої якістю апроксимації, що є результатом використання маски невеликого розміру 3x3 [21, 242]. Математичне забезпечення виявлення перепадів яскравості зображення в точці з координатами (*i,j*) за Собелем складають вирази:

$$D_x = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad (2.43)$$

$$D_y = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \quad (2.44)$$

$$S_x = D_x * f, \quad (2.45)$$

$$S_y = D_y * f, \quad (2.46)$$

$$S = \sqrt{S_x^2 + S_y^2}, \quad (2.47)$$

де f – початкове зображення;

D_x, D_y – маски фільтру Собеля;

S – відфільтроване зображення з виділеними кордонами об'єктів;

$*$ – операція згортки.

Розрахунок результату згортки матриці, що відповідає початковому зображенню *f*, з матрицями, що відповідають маскам фільтру Собеля D_x та D_y , реалізується за допомогою виразів виду:

$$S_x(i,j) = (-1)\cdot f(i-1,j-1) + (-2)\cdot f(i-1,j) + (-1)\cdot f(i-1,j+1) + (0)\cdot f(i,j-1) + (0)\cdot f(i,j) + (0)\cdot f(i,j+1) + (1)\cdot f(i+1,j-1) + (2)\cdot f(i+1,j) + (1)\cdot f(i+1,j+1) = -f(i-1,j-1) - 2\cdot f(i-1,j) - f(i-1,j+1) + f(i+1,j-1) + 2\cdot f(i+1,j) + f(i+1,j+1) \quad (2.48)$$

$$S_y(i,j) = (-1)\cdot f(i-1,j-1) + (0)\cdot f(i-1,j) + (1)\cdot f(i-1,j+1) + (-2)\cdot f(i,j-1) + (0)\cdot f(i,j) + (2)\cdot f(i,j+1) + (-1)\cdot f(i+1,j-1) + (0)\cdot f(i+1,j) + (1)\cdot f(i+1,j+1) = -f(i-1,j-1) - 2\cdot f(i,j-1) - f(i+1,j-1) + f(i-1,j+1) + 2\cdot f(i,j+1) + f(i+1,j+1) \quad (2.49)$$

Напрямок градієнта $\nabla f = (S_x, S_y)^T$ визначається кутом між його напрямком і віссю абсцис:

$$\vartheta = arctg\left(\frac{S_y}{S_x}\right). \quad (2.50)$$

Величина градієнта визначається нормою Гьольдера:

$$\|x\|_p = \sqrt[p]{\sum_i x_i^p}. \quad (2.51)$$

Довжина вектору градієнта в точці, розрахована за допомогою (2.48, 2.49, 2.51), буде впливати на те, чи увійде піксель до складу кордону.

Величина кута, розрахована за допомогою (2.48-2.50), використовується для визначення напрямку контуру в точці. Тобто кут ϑ використовується для зменшення товщини кордону контуру – світлі пікселі результуючого зображення, які не лежать в напрямку збільшення градієнту, ігноруються.

Останній крок алгоритму – двохпорогове відсікання з гістерезисом. Якщо вибрати занадто високий поріг, то можливі помилки першого роду – втрата точок,

які лежать на кордоні.

Якщо вибрати дуже низький поріг, то почастішають помилки другого роду – зайві точки приймаються за фрагменти межі. Тому використовують два порога. Спочатку застосовують високий поріг, який виділяє пікселі, що достовірно належать кордону. Потім обходять кордон і, використовуючи інформацію про напрямки, добудовують його, застосовуючи низький поріг.

Така процедура дозволяє залишати на зображені кордони, що відповідають безперервним кривим, і дозволяє відсікати дрібні зміни інтенсивності. В результаті обробки на зображенні залишаються лише монохромні контури обличчя. Зазначимо, що для приведення кольорового зображення у напівтонове, а потім у монохромне (чорно-біле), в базовому варіанті використовуються вирази виду:

$$N(x,y) = 0{,}2125R(x,y) + 0{,}7154G(x,y) + 0{,}0721B(x,y), \qquad (2.52)$$

$$Ib(x,y) = \begin{cases} 1, if\ N(x,y) > \delta \\ 0, if\ N(x,y) \le \delta \end{cases}, \qquad (2.53)$$

де $N(x,y)$ – результат перетворення кольорового пікселя з координатами x, y в напівтоновий формат;

$R(x,y)$, $G(x,y)$, B(x,y) – значення, що відповідають червоному, зеленому та синьому каналу кольорового пікселя з координатами x, y;

$Ib(x,y)$ – результат перетворення напівтонового пікселя в монохромний;

δ – порогове значення.

Зазначимо, що труднощі застосування виразу (2.53) стосуються визначення порогового значення, котре залежить від кольору фону. Для подолання вказаних труднощів в математичному забезпеченні даної моделі використано алгоритм Бредлі-Рота, що базується на співвіднесенні порогового значення δ із середнім значенням яскравості в межах заданої ділянки зображення. На першому кроці алгоритму Бредлі-Рота для кожного із пікселів напівтонового зображення розраховується інтенсивність його кольору:

$$Q(x,y) = N(x,y) + Q(x-1,y) + Q(x,y-1) - Q(x-1,y-1). \qquad (2.54)$$

На наступних кроках алгоритму обробка зображення розділяється на вісім частин. Розраховується середня яскравість пікселів в кожній із цих частин. Порогове значення розраховується окремо для кожної із вказаних частин зображення як 15% від середньої яскравості даної частини [217, 274]. Типові завади, що заважають розпізнавати емоції та особу слухача по зображенню його обличчя, як правило, викликані ракурсом відеореєстрації та наявністю предметів, що закривають частину зображення. Відповідно до теоретичних положень [217], нівелювання негативного впливу ракурсу доцільно реалізовувати за рахунок адаптації параметрів НММ до розпізнавання повернутих зображень та/або за рахунок внесення у навчальну вибірку відповідних прикладів. Тому в даній моделі обробки БП, асоційованих з зображеннями, типові завади, викликані ракурсом відеореєстрації, не враховуються. Інший клас типових завад доцільно нівелювати запропонованим методом застосування вейвлет-перетворень для фільтрації вхідного поля. В загальному випадку параметри шестидесяти чотирьох частково показаних на рис. 2.7 ключових точок обличчя, що відображають емоції слухача СДН, визначаються за допомогою методів, наведених в [218, 278].

Рис. 2.7. Приклад часткового відображення 64 ключових точок обличчя при розпізнаванні емоцій

За допомогою комп'ютерних експериментів встановлено, що для випадку розпізнавання емоцій слухача СДН може бути достатньо 14 контрольних точок, схема розміщення яких показана на рис. 2.8.

Особливістю наведеної схеми являється наявність двох позначених на рис. 2.8 точок, що відповідають внутрішнім кутам очей слухача СДН. Вказана особливість дозволяє уніфікувати механізм визначення початку відліку при нормалізації координат контрольних точок, що використовуються для розпізнавання емоцій та для розпізнавання особи.

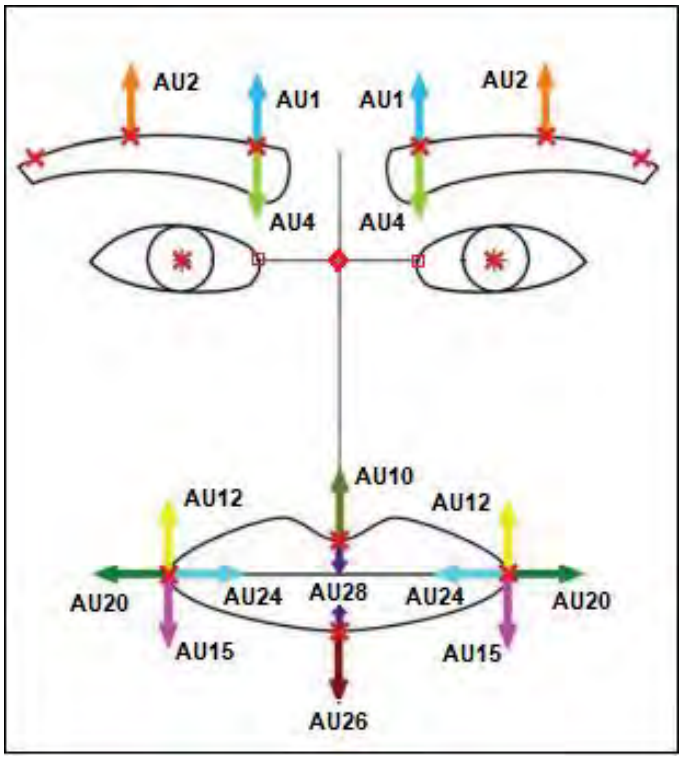

Рис. 2.8. Схема розміщення 14 ключових точок обличчя при розпізнаванні емоцій

Крім 14 ключових точок, на рис. 2.8 також показано рухові одиниці областей брів і губ, позначені згідно з системою кодування Пола Екмана [12]:

- AU1 – підняття верхньої частини брови;

- AU2 – підняття зовнішньої частини брови;

- AU4 – опускання брів;
- AU10 – підняття верхньої губи;
- AU12 – підняття куточків губ;
- AU15 – опускання куточків губ;
- AU20 – відтягування губ;
- AU24 – звуження губ;
- AU26 – підняття/опускання нижньої щелепи;
- AU28 – стиснення губ.

В базовому випадку саме рухові одиниці, показані на рис. 2.8, доцільно використовувати для розпізнавання емоцій слухача СДН.

Визначення параметрів контрольних точок обличчя для розпізнавання особи реалізується на основі результатів [141, 312]. Для цього окреслюються контури очей, брів, губ, носа і інших значимих елементів обличчя. На визначених місцях контурів виставляються контрольні точки. Після цього розраховуються відстані між цими точками.

Кількість контрольних точок обличчя, яку можливо визначити на фронтальному та повернутому зображенні, залежить від розміру зареєстрованого зображення та знаходиться в межах від 12-92. Відповідно нормативної документації [196-198], з урахуванням особливостей автоматизованого розпізнавання та особливостей реєстрації зображення обличчя, передбачено визначення 23 антропометричних контрольних точок на фронтальному зображенні. Розташування вказаних точок показане на рис. 2.9. На цьому рисунку цифрами позначено антропометричні контрольні точки: 1 - надпідборідна, 2 - губна нижня, 3 - ротова, 4 - губна верхня, 5 - підносова, 6 - передньоносова, 7 - точка у внутрішньому куті ока, 8 - точка у зовнішньому куті ока, 9 - контур лінії росту волосся, 10 - надбрівна лінія, 11 - точка на перетині лінії, яка з'єднує точки лобних бугрів, що найбільш виступають, 12 - надглабелярна, 13 - точка в ділянці, де сходяться надбрівні дуги, 14 - найбільш глибока точка перенісся, 15 - зовнішньоочна, 16 - нижньоочна,

17 - козелок вуха, 18 - вилицева, 19 - носокрилова, 20 - губнокутова, 21 - кутова щелепна, 22 - щелепна, 23 - зона підборіддя.

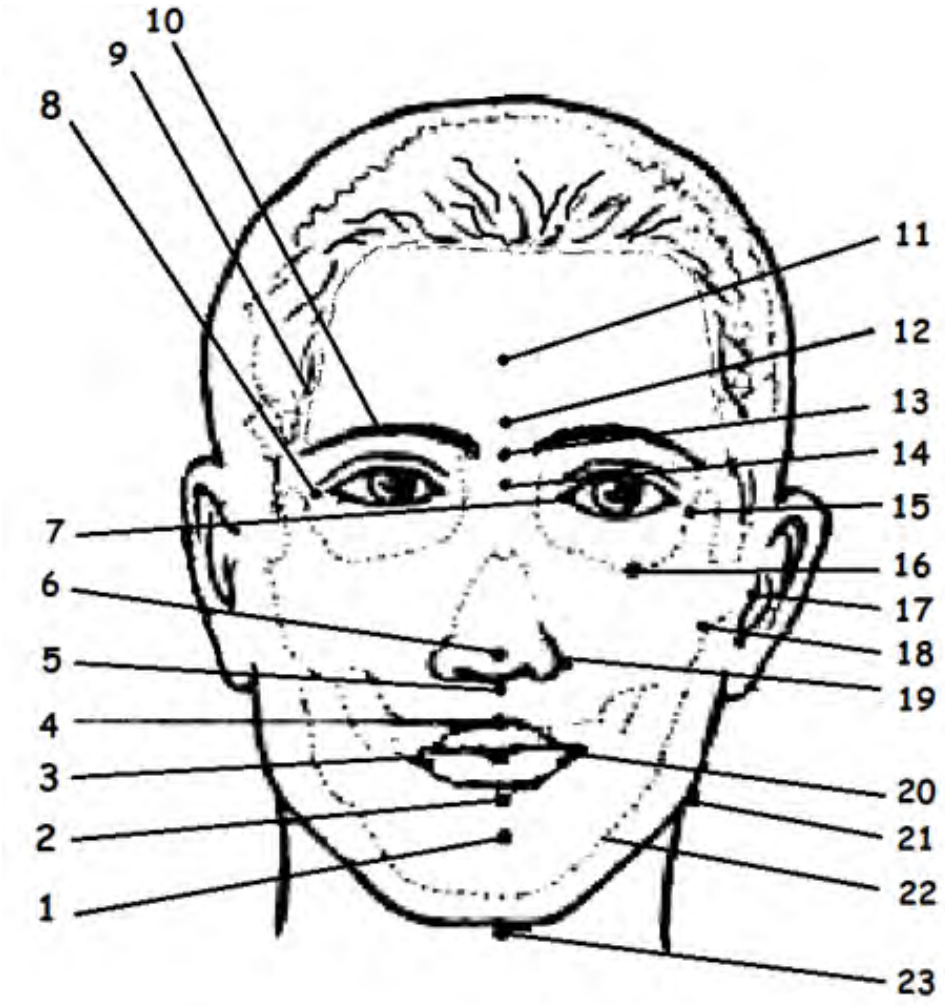

Рис. 2.9. Антропометричні контрольні точки на двовимірному зображенні обличчя людини

Зазначимо, що особливості реєстрації зображення обличчя та особливості доступних навчальних БД зумовлюють необхідність обчислення нормованих координат контрольних точок обличчя як у випадку розпізнавання емоцій, так і у випадку розпізнавання особи. Нормалізація координат здійснюється на основі розрахунку відношення відстаней між контрольними точками. Ілюстрацією процесу визначення відстаней являється рис. 2.10.

В базовому варіанті в якості точки початку відліку можливо використати

точку, що відповідає середині прямої лінії, яка з'єднує внутрішні кути очей. На рис. 2.8, 2.10 ця точка позначена ромбом.

Рис. 2.10. Виділення антропометричних точок на двовимірному зображенні обличчя людини з подальшим обчисленням відстаней між ними

Нормалізовані координати контрольних точок можливо розраховувати за допомогою виразів виду:

$$\bar{x}(i) = x(i)/l_x, \tag{2.55}$$

$$\bar{y}(i) = y(i)/l_y, \tag{2.56}$$

де i – номер контрольної точки;

$\bar{x}(i)$, $\bar{y}(i)$ – нормалізовані координати і-ої контрольної точки;

$x(i)$, $y(i)$ – координати і-ої контрольної точки;

l_x, l_y – базис нормалізації по осі абцис та по осі ординат.

Для спрощення розрахунків можливо прийняти, що $l_x = l_y$. В якості базису нормалізації доцільно використовувати відстань між контрольними точками, які не складно визначити на зображенні обличчя слухача СДН. Наприклад, це може бути відстань між зіницями очей.

Нормалізацію параметрів кольору пікселя зображення або кольору

контрольної точки можливо здійснити за допомогою виразу виду:

$$\bar{c} = (c - min_c)/(max_c - min_c), \quad (2.57)$$

де $\bar{c}$ – нормалізоване значення кольорового каналу пікселя;

c – зареєстроване значення кольорового каналу пікселя;

max_c, min_c – максимально та мінімально можливі значення кольорового каналу пікселя;

Кодування кольорових параметрів пікселя зображення або контрольної точки полягає у співставленні їх кольорових каналів з кольоровими каналами вхідного поля ЗНМ. Операція кодування геометричних параметрів пікселів зображення полягає у співставленні координат кожного пікселя з координатами вхідного поля ЗНМ.

Особливістю операції кодування координат контрольних точок являється необхідність перенесення початку відліку із визначеної контрольної точки у лівий нижній кут обробленого зображення обличчя. Вказана особливість пояснюється тим, що прийнято початок координат вхідного поля ЗНМ встановлювати у нижній лівий кут зображення.

Вираз для кодування координат контрольних точок може бути записано у вигляді:

$$\tilde{x} = \bar{x} + a, \quad (2.58)$$

$$\tilde{y} = \bar{y} + b, \quad (2.59)$$

де $\tilde{x}$, $\tilde{y}$ – закодовані значення координат контрольних точок;

$\bar{x}$, $\bar{y}$ – нормалізовані значення координат контрольних точок;

a, b – координати точки, що прийнята за початок відліку в операції нормалізації координат контрольних точок.

Слід відзначити, що описані операції забезпечують адаптацію процедури формування вхідного поля ЗНМ до можливої варіації умов застосування в СДН в розрізі підходу до мінімізації ресурсоємності НМЗ.

Розроблена модель обробки зображення обличчя використана в якості прототипу при розробці подібних моделей, що базуються на використанні допоміжних БП, асоційованих із зображеннями. При цьому враховано, що зображення, асоційовані з цими БП, в першу чергу відрізняються від зображення обличчя розмірами, а для вушної раковини та відбитків пальців ще й іншою кількістю та іншими координатами контрольних точок. Крім того, допоміжні БП не використовуються для розпізнавання емоцій слухача СДН. Тому вушна раковина та відбитки пальців характеризуються не двома, як зображення обличчя, а одним набором контрольних точок. Відзначимо, що підхід на основі контрольних точок для аналізу райдужної оболонки та сітківки ока не використовується. Розглянемо особливості моделей обробки зображень, асоційованих з допоміжними БП.

Модель обробки зображення вушної раковини. Індивідуальні особливості формування вхідного поля ЗНМ, призначеної для аналізу вушної раковини визначаються тим, що нейромережевий аналіз передбачає використання тільки монохромного фронтального зображення окремо лівої та окремо правої вушної раковини. При цьому зображення вушної раковини в умовах СДН може бути зареєстроване в кольоровому або напівтоновому форматі.

Таким чином, особливістю обробки зображення вушної раковини являється обов’язкове застосування виразів (2.59-2.61) для приведення такого зображення у монохромний формат.

Типові завади реєстрації вушної раковини, як і при реєстрації обличчя, викликані поворотом голови слухача СДН та наявністю об’єктів, що закривають частину зображення (волосся, головний убір). Врахувати вказані особливості можливо, по аналогії з обробкою зображення обличчя, за рахунок використання капсульної НМ та фільтру на основі вейвлет-перетворень.

В якості точки початку відліку, що застосовується для нормалізації контрольних точок, прийнято використовувати козелок, а в якості базису нормалізації по осі абцис та по осі ординат використано відстань між козелком та

ніжкою завитку.

На рис. 2.11 козелок позначено цифрою 1, ніжка завитку – цифрою 2, а базис нормалізації - літерою *L*. Використання для розрахунку базису козелка та ніжки завитку пояснюється тим, що серед анатомічних точок ці точки - одні з найбільш чітко виражених та на відміну від зовнішніх кордонів вушної раковини в більшості випадків не перекриваються поширеними типовими завадами.

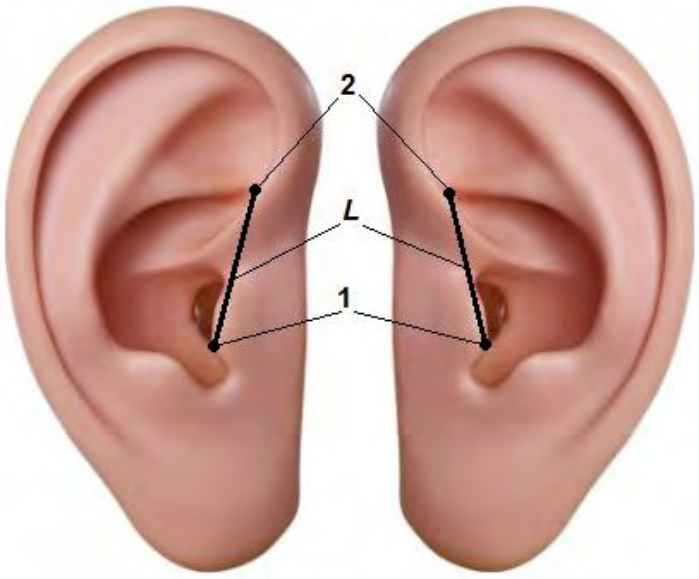

Рис. 2.11. Визначення точки початку відліку та базису нормалізації контрольних точок на двовимірному зображенні вушної раковини

Модель обробки відбитків пальців. Базуючись на вітчизняних та закордонних нормах в області захисту інформації, визначено, що нейромережевий аналіз відбитків пальців повинен передбачати використання напівтонового або монохромного двовимірного зображення. Кожен піксель відбитка напівтонового зображення характеризується градаціями сірого кольору від 0 до 255.

Мінімальний рівень яскравості точки, відповідний чорному кольору, має дорівнювати нулю. Максимальний рівень яскравості точки відповідає білому кольору. У випадку необхідності подання на вхід ЗНМ монохромного зображення зареєстроване напівтонове зображення підлягає бінаризації. Таким чином,

особливістю обробки зображення відбитка пальця є обов'язкове застосування виразів (2.53, 2.54) для приведення такого зображення у монохромний або напівтоновий формат.

Скелетизація зображення відбитка пальця полягає у зменшенні товщини папілярних ліній до одного пікселя і реалізується з використанням шаблонів, наведених в [53, 206].

Максимальні розміри відбитків пальців слухачів, які повинні зберігатись в біометричних БД, дорівнюють 40,6×38,1 мм (1,6×1,5 дюйма). При рекомендованій роздільній здатності в 500 dpi вказаний розмір відбитка характеризується за допомогою 800×750 точок. Разом з тим, в загальнодоступних БД зберігаються відбитки розміром 512×512, 300×300, 256×364, 448×478, 240×320 пікселів, кожен з яких також описується градаціями сірого кольору від 0 до 255 [229].

Відзначимо, що поширені сканери відбитків дозволяють реконструювати зображення відбитка з розміром приблизно 25×14 мм, що відповідає зображенню розміром 500×280 пікселів. Також аналіз технології сканування відбитків пальців вказує на те, що кут повороту відбитка знаходиться в межах $\pm 15°$, а його розмір варіюється в межах $\pm 10\%$.

Оскільки такі варіації зображення ефективно нівелюються ЗНМ, це дозволяє істотно знизити вимоги до попередньої обробки зображення відбитка пальця. При цьому базис нормалізації по осі абцис та по осі ординат доцільно співвіднести з розміром зареєстрованого відбитку пальця.

Враховуючи результати першого розділу та дані [53, 180], визначена доцільність використання двох типів контрольних точок відбитків пальців. Це показані на рис. 2.12 точки, що відповідають початку (закінченню) папілярних ліній або точки, що визначають їх злиття (розгалуження) папілярних ліній.

На рис. 2.12 початок (закінчення) папілярної лінії позначена літерою *A*, а розгалуження папілярних ліній – літерою *B*.

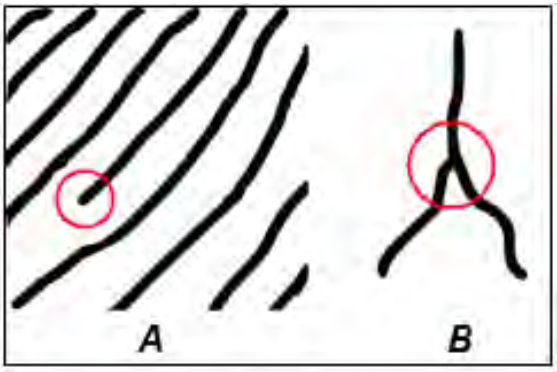

Рис. 2.12. Контрольні точки (мінуції) відбитку пальця

Використання описаних ключових точок пояснюється їх достатньою кількістю (від 30 до 70 на одному відбитку) для розпізнавання особи слухача СДН та наявністю апробованої технології виявлення таких точок. Тому, на відміну від визначення ключових точок обличчя, модель зображення відбитку пальця передбачає локальну обробку всього зображення відбитка пальця за допомогою маски (матриці $|M|$) розміром 9×9 пікселів.

В процесі обробки підраховується число пікселів, що знаходяться навколо центру маски і мають ненульові значення. Центральний піксель приймається за контрольну точку, якщо він сам має одиничне значення, а число сусідніх ненульових пікселів дорівнює 1 (закінчення лінії) або 3 (роздвоєння лінії). Математичне забезпечення визначення контрольної точки в межах маски складають вирази:

$$if\ S = 1 \rightarrow m_{2,2} = k_{end}, \tag{2.60}$$

$$if\ S = 3 \rightarrow m_{2,2} = k_{fork}, \tag{2.61}$$

$$S = \sum_{i=1, i\neq 2}^{9} \sum_{j=1, j\neq 2}^{9} m_{i,j}, \tag{2.62}$$

де $m_{i,j}$ – елемент маски;

k_{end} – контрольна точка закінчення лінії;

k_{fork} – контрольна точка роздвоєння лінії.

Модель обробки райдужної оболонки ока. Характер вхідних даних та результати першого розділу даної дисертаційної роботи дозволяють інтерпретувати розпізнавання особи слухача на основі райдужної оболонки ока у вигляді завдання нейромережевої класифікації круглих зображень обмеженого розміру у відтінках сірого кольору. Тобто на вхід ЗНМ зображення райдужної оболонки ока доцільно подати у напівтоновому форматі. Вказане твердження підкріплюється тим, що у вільнодоступних БД, котрі застосовуються в апробованих засобах біометричної аутентифікації, містяться напівтонові зображення ока людини.

На відміну від зображення обличчя, проведення нейромережевого аналізу зображення райдужної оболонки ока зумовлює необхідність підвищення деталізації вхідного зображення. Ще однією відмінністю являється складність виділення кордонів зображення райдужної оболонки [108]. При цьому в умовах СДН зображення ока може бути зареєстроване в кольоровому форматі, недостатня якість якого може бути викликана нерівномірним освітленням та позиціонуванням відеокамери. Означені особливості зумовили застосування в моделі обробки зображення райдужної оболонки ока запропонованих для первинної обробки зображення обличчя операцій, що пов'язані з перетворенням зареєстрованого зображення у напівтоновий формат та підвищенням його контрастності.

Крім того, передбачається використання операції виділення зображення райдужної оболонки на основі виразу:

$$grad = \max_{(r,x_0,y_0)} \left| G_\sigma(r) * \frac{\mathrm{d}}{\mathrm{d}r} \oint_{r,x_0,y_0} \frac{I(x,y)}{2\pi r} \mathrm{d}s \right|, \tag{2.63}$$

$$r \in [R_{min}, R_{max}, \Delta], \tag{2.64}$$

де $I(x,y)$ – зображення ока, визначене в декартовій системі координат;

r – радіус пошуку кордонів райдужної оболонки ока;

R_{min}, R_{max} – мінімальне та максимальне можливе значення радіусу;

Δ – крок пошуку радіусу зображення райдужної оболонки ока;

(x_0, y_0) – координати центру зіниці ока;

* – символ, що позначає операцію згортки матриць;

$G_\sigma(r)$ – функція Гауса;

$\mathrm{d}s$ – елементарна дуга;

σ – радіус функції Гауса.

Нівелювання типових завад, пов'язаних з відблисками та частковими оклюзіями вій і повік, передбачено реалізувати за рахунок застосування вейвлет-перетворень для фільтрації вхідного поля.

Приклад обробленого зображення райдужної оболонки наведено на рис. 2.13.

Рис. 2.13. Оброблене зображення райдужної оболонки ока

Оскільки умови застосування НМЗ аналізу БП слухача СДН передбачають проведення відеореєстрації зображення ока на достатньо близькій відстані, то необхідності нівелювати можливі геометричні спотворення зображення райдужної оболонки ока не виникає.

Визначення контрольних точок на зображенні радужної оболонки, по аналогії з [120, 155, 253], доцільно реалізувати за рахунок згортки матриці, що відповідає обробленому зображенню радужної оболонки, з функцією перетворення Ерміта:

$$f_n(x,\sigma) = \frac{(-1)^n e^{-\left(\frac{x}{\sigma}\right)^2}}{\sqrt{2^n n!\,\pi}} H_n\left(\frac{x}{\sigma}\right), \qquad n = 0,1,2,\dots N, \tag{2.65}$$

$$H_n(y) = 2yH_{n-1}(y) - 2(n-1)H_{n-2}(y), \tag{2.66}$$

де $H_n(y)$ – поліном Ерміта n-го порядку;

σ – визначений коефіцієнт.

В процесі обробки зображень райдужної оболонки ока при розрахунках (2.65) доцільно прийняти $n = 10$, а $\sigma = 3$. Також у випадку $n = 0$ прийнято $H_0(y) = 1$, а у випадку $n = 1$ $H_1(y) = 2y$.

Зазначимо, що на відміну від відомих, в розробленій моделі обробки зображення райдужної оболонки ока подання зображення у вигляді, розгорнутому в полярних координатах, не передбачається. Це пояснюється доведеною можливістю ефективного аналізу круглих та еліпсовидних зображень за допомогою ЗНМ [241, 303]. Перед подачею в ЗНМ кругле зображення райдужної оболонки ока доповнюється до квадрату пікселями фонового (білого) кольору.

Модель обробки зображення сітківки ока. Відповідно результатів досліджень, описаних в першому розділі даної дисертаційної роботи, першочерговою задачею, що виникає в процесі формування вхідного поля НММ, що призначена для розпізнавання особи слухача СДН на основі сітківки ока, є визначення кордонів зіниці ока. Першочерговість цієї задачі викликана тим, що саме через зіницю ока спрямовується інфрачервоний промінь, який повинен освітити сітківку ока.

В даній моделі для визначення кордону зіниці ока, як і у випадку виділення контурів зображення обличчя людини, передбачено використання алгоритму Кенні, заданого виразами (2.43-2.51). Результат застосування алгоритму Кенні для обробки зображення ока для визначення контурів зіниці ока наведено на рис. 2.14.

Можливий недолік застосування алгоритму Кенні пов’язаний з тим, що отриманий кордон зіниці ока є незамкнутим. Для нівелювання такого недоліку пропонується використовувати описану в [22, 49, 61] технологію фільтрації, що полягає у розширенні знайдених кордонів. Розширення здійснюється до тих пір, доки отримані окремі блоки не будуть об’єднані у великі лінійні сегменти. Як показують результати досліджень, за рахунок такого об’єднання периметр зіниці прийме форму замкнутого кола.

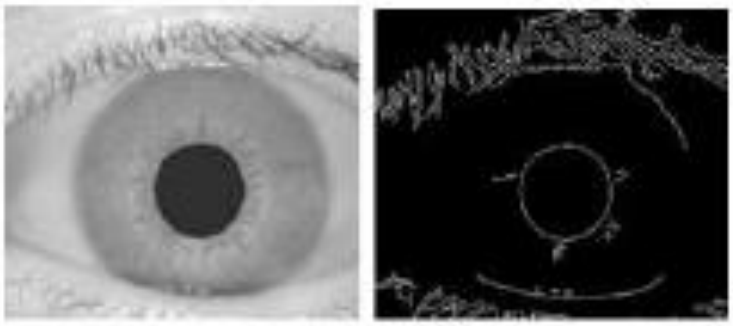

Рис. 2.14. Результат застосування алгоритму Кенні для визначення контурів зіниці ока

Після визначення зовнішнього кордону область зображення, що співвідноситься з зіницею ока, заливається чорним кольором. Можливі недоліки на зображені зіниці, що викликані відблисками або іншими завадами, відображаються у вигляді світлих областей. Передбачено нівелювання таких недоліків за рахунок заливки світлих областей, інтенсивність кольору яких нижча, ніж визначене порогове значення. Визначення центру зіниці базується на побудові двох хорд, що проходять через центр зіниці. Центр визначається як точка перетину двох перпендикулярів, проведених через середини хорд.

Виділена зіниця ока використовується як область, через яку пропускається інфрачервоний промінь, що використовується для підсвічування сітківки ока. Зазначимо, що наведені операції виділення кордонів та центру зіниці доцільно застосовувати в засобах відеореєстрації для їх адаптації до реєстрації зображення сітківки ока слухача. Зареєстроване в інфрачервоному діапазоні зображення сітківки ока являється джерелом інформації для нейромережевого аналізу.

На відміну від процедури обробки зображення райдужної оболонки ока, при обробці зображення сітківки ока завдання нівелювання спотворення зображення за рахунок повороту зіниці ока залишається актуальним. Це пояснюється тим, що поворот зіниці відбувається через підсвідому зміну погляду і величину повороту складно обмежити. При цьому через незначні розміри і складний характер поверхні

сітківки ока навіть незначний поворот зіниці викликає ускладнення процесу розпізнавання.

Оскільки при відновленні повернутого зображення сітківки ока величина повороту може бути розрахована на основі аналізу зіниці ока, то функція спотворення зображення сітківки може бути апріорно визначена. Це дозволяє провести відновлення повернутого зображення сітківки ока без застосування капсульної НМ за допомогою фільтру Вінера [21, 61, 155]:

$$\widehat{Im}(u,v) = \hat{F}(u,v)Im(u,v), \tag{2.67}$$

$$\hat{F}(u,v) = \left(\frac{H^*(u,v)}{|H(u,v)|^2 + Z(u,v)}\right) G(u,v), \tag{2.68}$$

$$Z(u,v) = \frac{S_\eta(u,v)}{S_f(u,v)}, \tag{2.69}$$

де $\widehat{Im}(u,v)$ – відновлене зображення;

$Im(u,v)$ – зашумлене зображення;

$\hat{F}(u,v)$ – фільтр Вінера в частотній області;

$Z(u,v)$ – відношення сигнал/шум;

$G(u,v)$ – спектр спотвореного зображення;

$H(u,v)$ – функція спотворення;

$H^*(u,v)$ – функція, що є комплексно-спряженою до $H(u,v)$.

Також перед визначенням контрольних точок для покращення якості зображення сітківки ока передбачено реалізувати операцію стиснення та операцію відновлення контурів неякісно відсканованих судин сітківки ока.

Особливості вказаних операцій пов’язані із специфікою сканування сітківки ока та специфікою розпізнавання.

В основі операції стиснення покладена гіпотеза про можливість представлення судин сітківки ока у вигляді контурних напівтонових ліній товщиною в один піксель. Тому стиснення відбувається до тих пір, доки товщина

судини на зображенні не буде дорівнювати одному пікселю. При цьому спочатку виконується пошук пікселя, що є частиною судини. Як правило, пошук починається з верхнього лівого кута зображення сітківки ока і полягає у визначенні пікселя сірого кольору. Знайдений піксель визначається як центр маски розміром 3x3 пікселя. Зазначимо, що в цьому випадку верхній ряд пікселів та перший зліва піксель центрального ряду будуть замальовані білим кольором.

Оскільки передбачається, що піксель судини може мати не більше двох сусідніх пікселів, то далі, як це показано на рис. 2.15, аналізуються 4 сусідніх пікселі знайденої точки.

На рис. 2.15 пікселі, що підлягають аналізу, позначені символом «?».

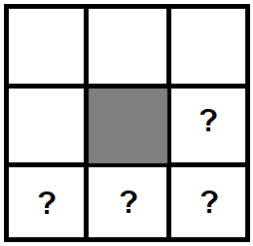

Рис. 2.15. Ілюстрація операції пошуку пікселів судин

Якщо піксель в середині маски не має сусідів сірого кольору, то він відкидається і продовжується пошук інших пікселів кровоносних судин. В інших випадках це або кінцева, або внутрішня точка судини. На кожному кроці пошуку сусід сірого кольору останнього пікселя позначається як пройдений і вибирається наступним центральним пікселем маски 3x3. Вибір такого сусіда проводиться за принципом: найкращий сусід той, у якого найбільша кількість непомічених сірих сусідів. Інші пікселі, що входять до маски, замальовуються білим кольором. Використання описаної операції зменшення товщини судин, а також неякісне сканування зумовлюють можливість спотворення зображення сітківки ока, що полягає у хибних розривах контурів судин. Для відновлення контуру між двома прилеглими кінцевими точками визначаються кути між прямою, що об’єднує дані

кінцеві точки, та лініями контуру судин. Якщо величини даних кутів менші, ніж визначена порогова величина λ, то кінцеві точки об'єднуються. Так, рис. 2.16 ілюструє операцію відновлення контуру між судинами, що відповідають дугам AB та DC. Лінії, що продовжують ці дуги, позначені як AI та CJ, а лінію, що поєднує кінцеві точки – як AC. Кути, величини яких порівнюються з пороговою величиною, позначені як ρ та ψ.

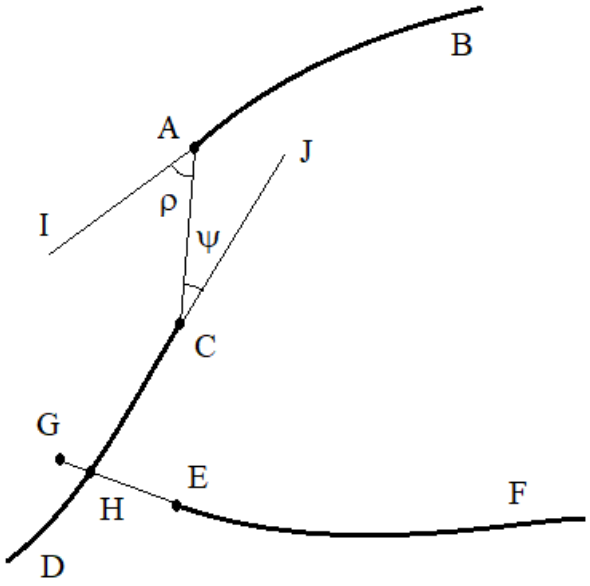

Рис. 2.16. Ілюстрація операцій відновлення контуру

Таким чином, умова об’єднання контурів AB та DC описується виразом

$$\rho \leq \lambda \text{ and } \psi \leq \lambda. \tag{2.70}$$

Щоб відновити точки розгалуження для кожної кінцевої точки обчислюється напрямок продовження контуру судини, після чого проводиться розширення контуру на радіус фіксованої довжини. Якщо це розширення перетинається з іншим сегментом, то знайдена точка являється точкою розгалуження. На рис. 2.16 точка розгалуження позначена як H, лінія продовження контуру судини – як FG, а відрізок, що відповідає розширенню контуру, як GE. Таким чином, умова відновлення точки перетину записується у вигляді:

$$[\,HE\,] \leq [GE]. \tag{2.71}$$

Операцію розрахунку контрольних точок на зображенні сітківки ока доцільно реалізувати, як і у випадку обробки зображення райдужної оболонки ока, з використанням виразів (2.65, 2.66).

В результаті проведених досліджень розроблено моделі обробки зображень БП, асоційованих з зображеннями, що за рахунок врахування особливостей реєстрації вказаних зображень в умовах СДН, вимог до підготовки вхідних даних ЗНМ та за рахунок обґрунтованого застосування технологій комп'ютерного зору, забезпечили теоретичний базис розробки методів формування вхідного поля НММ розпізнавання емоцій та особи слухачів СДН.

Перспективи подальшого розвитку моделей обробки співвідносяться з інтеграцією в них процедур фільтрації на основі методів вейвлет-перетворень, що дозволить підвищити ефективність обробки зашумлених зображень. Крім того, перспективи подальших досліджень пов'язані з адаптацією моделей обробки зображення райдужної оболонки ока до умов задачі створення НМЗ розпізнавання емоцій.

2.4. Нейромережева модель розпізнавання емоцій слухача, адаптована до нівелювання завад, пов'язаних з поворотом обличчя на зображенні

Проведемо деяке уточнення постановки задачі розпізнавання емоцій слухача СДН. Відповідно результатів досліджень першого розділу, в більшості випадків впродовж навчального процесу, пов'язаного із сприйняттям учбових матеріалів, виникає необхідність розпізнавання наперед визначеної множини базових емоцій на основі аналізу статичних зображень обличчя, зафіксованого за допомогою відеокамери із середніми характеристиками. При цьому питання розпізнавання особи, попередньої фільтрації зображення, впливу освітленості, виділення на зображенні окремих осіб і виправлення свідомого спотворення обличчя слухача СДН з метою приховування ним свого емоційного стану не розглядається.

Дані [95, 138, 217] та результати, отримані при розробці моделі обробки зображення обличчя, вказують на те, що в цьому випадку, як і для випадку реєстрації обличчя користувачів ІС загального призначення, основні перешкоди виникають в результаті повороту обличчя слухача СДН відносно жорстко зафіксованої відеокамери. Загальноприйнятий підхід до нівелювання вказаної перешкоди за рахунок внесення в навчальну вибірку відповідних прикладів, що містять повернуті зображення обличчя, потребує використання значних додаткових ресурсів.

Разом з тим, відповідно до [170, 185, 258], нівелювати таку перешкоду можливо за рахунок використання моделі на базі капсульної НМ (CapsNetEm), яка є однією із найбільш досконалих модифікацій ЗНМ, що пристосована до аналізу повернутих і зашумлених зображень. Структура мережі CapsNetEm, що адаптована до задачі розпізнавання емоцій на основі зображення обличчя, показана на рис. 2.17.

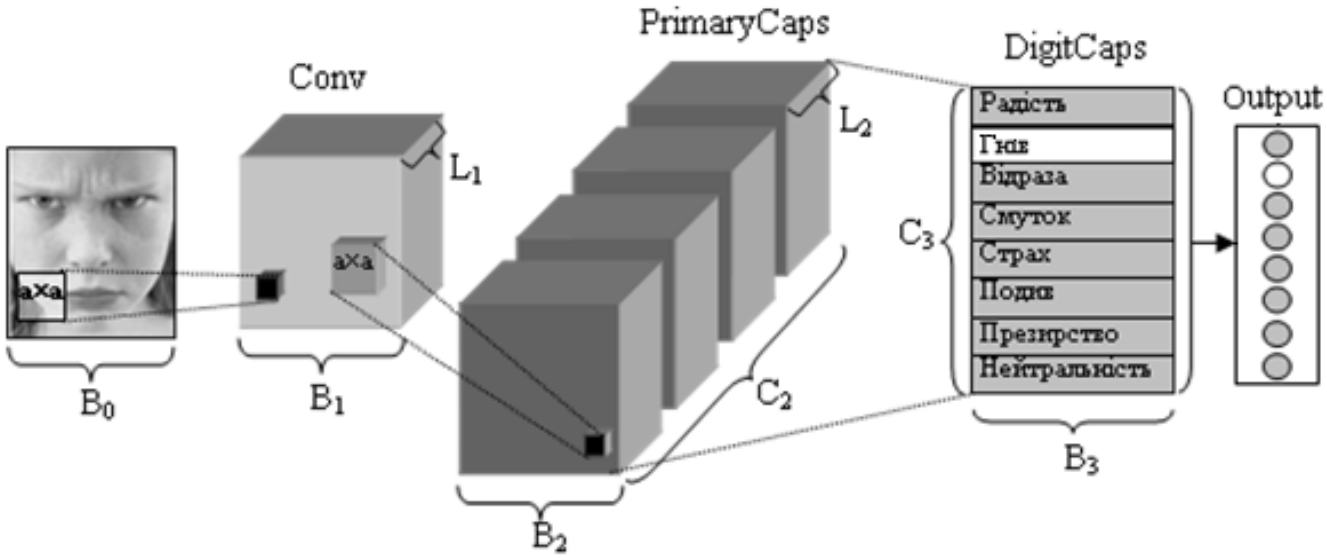

Рис. 2.17. Структура CapsNetEm, адаптована до розпізнавання емоцій на основі аналізу зображення обличчя

Основними структурними одиницями CapsNetEm є вхідний шар нейронів, згортковий шар, шар первинних капсул, шар згорткових капсул та вихідний шар. На

рис. 2.17 використані наступні позначення, які відповідають структурним параметрам мережі:

- Conv – згортковий шар;
- PrimaryCaps – шар первинних капсул;
- DigitCaps – шар згорткових капсул;
- B_0 – вертикальний і горизонтальний розмір аналізованого зображення;
- B_1 – розмір карт ознак в згортковому шарі;
- B_2 – розмір сітки в шарі первинних капсул;
- B_3 – кількість одиниць згортки в кожній з капсул в шарі згорткових капсул;
- C_2 – кількість каналів в шарі первинних капсул;
- C_3 – кількість капсул в шарі згорткових капсул;
- L_1 – кількість карт ознак в згортковому шарі;
- L_2 – кількість одиниць згортки в кожному каналі шару первинних капсул;
- a – розмір ядра згортки.

Вхідний шар нейронів CapsNetEm співвідноситься з вхідним зображенням, а шар згорткових капсул – з емоціями, що мають бути розпізнані. Зазначимо, що на відміну від класичної капсульної мережі [258, 292], в CapsNetEm, адаптованій до використання в СДН для розпізнавання емоцій слухача за зображенням обличчя, блок декодування відсутній. Це пояснюється тим, що потреби у відновленні повернутого зображення обличчя слухача СДН не виникає. Порівняння з класичною капсульною НМ дозволяє стверджувати, що відсутність блоку декодування дозволяє приблизно на 15% зменшити ресурсоємність мережі. Розрахунок вхідних і вихідних сигналів для нейронів в згортковому шарі здійснюється так само, як і в ЗНМ. При цьому використовується класична функція активації ReLU виду:

$$y_k^{(i,j)} = max\left(0, x_k^{(i,j)}\right), \tag{2.72}$$

де $y_k^{(i,j)}$ – вихідний сигнал (*i,j*)-го нейрону *k*-ої карти ознак;

$x_k^{(i,j)}$ – сумарний вхідний сигнал (*i,j*)-го нейрону *k*-ої карти ознак.

При цьому сумарний вхідний сигнал нейрону згорткового шару розраховується так:

$$x_k^{(i,j)} = x_{0,k} + \sum_{s=1}^{a}\sum_{t=1}^{a}\left(w_{k,s,t}x_k^{(i+s,j+t)}\right), \quad (2.73)$$

де $x_{0,k}$ – зсув нейронів k-ої карти ознак;

a – розмір ядра згортки;

$w_{k,s,t}$ – ваговий коефіцієнт (s,t)-го зв'язку нейрону k-ої карти ознак.

Розрахунок вхідних і вихідних сигналів капсул реалізується так:

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2}\frac{s_j}{\|s_j\|}, \quad (2.74)$$

$$s_j = \sum_{i=1}^{C_3}\left(c_{i,j}\hat{u}_{j|i}\right), \quad (2.75)$$

$$\hat{u}_{j|i} = W_{i,j}u_i, \quad (2.76)$$

$$c_{i,j} = exp\left(b_{i,j}\right)\Big/\sum_{k=1}^{C_2} exp\left(b_{i,j,k}\right), \quad (2.77)$$

$$\Delta b_{i,j} = v_j\hat{u}_{j|i}, \quad (2.78)$$

$$b_{i,j} = b_{i,j} + \Delta b_{i,j}, \quad (2.79)$$

де v_j – вхідний вектор j-ої капсули в шарі згорткових капсул;

s_j – складова j-ої капсули шару згорткових капсул у вихідному сигналі мережі;

$c_{i,j}$ – ваговий коефіцієнт узгодженості між і-ою капсулою в шарі первинних капсул та j-ою капсулою в шарі згорткових капсул;

$\hat{u}_{j|i}$ – прогнозована величина вихідного сигналу і-ої капсули в шарі первинних капсул;

$W_{i,j}$ – матриця вагових коефіцієнтів зв'язків між і-ою капсулою в шарі первинних капсул та j-ою капсулою в шарі згорткових капсул;

$b_{i,j}$ – логарифм ймовірності зв'язку між і-ою капсулою в шарі первинних капсул та j-ою капсулою в шарі згорткових капсул;

u_i – вихідний сигнал і-ої капсули в шарі первинних капсул;

$\Delta b_{i,j}$ – коефіцієнт корекції при ітеративному розрахунку $b_{i,j}$.

Відзначимо, що вираз виду (2.74) представляє собою так звану squash-функцію.

Процес навчання CapsNetEm полягає у мінімізації функціоналу:

$$\sum_{k=1}^{K} E_k \to min, \tag{2.80}$$

$$E_k = T_k \max(0, m^+ - \|v_k\|)^2 + \lambda(1 - T_k) \max(0, \|v_k\| - m^-)^2, \tag{2.81}$$

де $K = C_2$ – кількість капсул в шарі згорткових капсул.

При цьому для капсули, яка відповідає k-ій емоції, $T_k = 1$ тільки тоді, якщо на зображенні обличчя ця емоція відображається. В протилежному випадку $T_k = 0$. Значення інших коефіцієнтів $m^+ = 0{,}9,\ m^- = 0{,}1,\ \lambda = 0{,}5$. Реалізувати режими навчання та розпізнавання CapsNetEm пропонується з використанням алгоритму, наведеному в [258]. Особливістю навчання CapsNetEm є застосування алгоритму динамічної маршрутизації між капсулами для розрахунку вагових коефіцієнтів $c_{i,j}$. Псевдокод процедури маршрутизації має наступний вигляд:

procedure Routing($\hat{u}_{j|i}$; r; l)

for all capsule i in layer l and capsule j in layer (l + 1): $b_{ij} \leftarrow 0$

for r iterations do

for all capsule i in layer l: $c_i \leftarrow \text{softmax}(b_i)$

for all capsule j in layer (l + 1): $s_j \leftarrow \sum_i c_{ij} \hat{u}_{j|i}$

for all capsule j in layer (l + 1): $v_j \leftarrow \text{squash}(s_j)$

for all capsule i in layer l and capsule j in layer (l + 1): $b_{ij} \leftarrow b_{ij} + \hat{u}_{j|i} \cdot v_j$

return v_j

По аналогії з класичною капсульною мережею, в базовому варіанті CapsNetEm визначено доцільність встановлення значень наступних конструктивних параметрів: кількість карт ознак L_1=256; кількість каналів в шарі первинних капсул C_2=32; кількість одиниць згортки в шарі первинних капсул L_2=8; кількість одиниць згортки в кожній із капсул шару згорткових капсул B_3=8; кількість капсул в шарі згорткових капсул повинна дорівнювати кількості визначених класів емоцій.

Таким чином, в результаті проведених досліджень розроблена НММ CapsNetEm адаптована до розпізнавання базових емоцій слухача СДН з урахуванням можливого повороту обличчя на зображенні. Розробка базується на капсульній НМ, що довела свою ефективність при розпізнаванні завад, викликаних поворотом зображення, що аналізується. На відміну від використаного прототипу, конструктивні параметри розробленої CapsNetEm адаптовані до умов СДН, що забезпечує можливість приблизно на 15% зменшити ресурсоємність мережі.

Подальші дослідження в області вдосконалення CapsNetEm доцільно співвіднести із збільшенням в ній кількості схованих шарів нейронів, що відповідає гіпотезі про можливість підвищення точності та зменшення обчислювальної реурсоємності НМЗ за рахунок збільшення рівнів ієрархії розпізнавання.

2.5. Висновки до другого розділу

В даному розділі вирішувалось наукове завдання розвитку методологічного забезпечення технології автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання на основі нейромережевого аналізу множини допустимих БП, реєстрацію яких можливо здійснити за допомогою поширеного АПЗ. Основні результати розділу наступні:

1. Вперше розроблено концептуальну модель розпізнавання емоційного стану слухачів системи дистанційного навчання, яка за рахунок конкретизації операцій, що повинні виконуватись для здійснення нейромережевого аналізу

біометричних параметрів, а також за рахунок обґрунтування переліку параметрів оцінки ефективності розпізнавання, дозволяє деталізувати напрямки досліджень з розробки відповідних засобів автоматизованого розпізнавання. Також визначено, що напрямки подальших досліджень пов'язані з удосконаленням методологічної бази розробки та використання нейромережевих засобів розпізнавання емоцій слухачів системи дистанційного навчання.

2. Отримав подальший розвиток методологічний базис розробки нейромережевих моделей аналізу біометричних параметрів слухачів системи дистанційного навчання для розпізнавання їх емоцій та особи. Розвиток методологічного базису здійснено з позицій співвіднесення підходів до нейромережевого розпізнавання емоцій та особи з основними факторами, які впливають на ефективність аналізу біометричних параметрів в очікуваних умовах функціонування системи дистанційного навчання. Розвиток методологічного базису полягає у доведенні принципової можливості застосування нейромережевих моделей для розпізнавання емоцій та обґрунтованому застосуванні при розробці таких моделей ряду відомих та запропонованих підходів.

3. Удосконалено підходи до визначення конструктивних параметрів, навчання та мінімізації ресурсоємності нейромережевих засобів за рахунок теоретичного обґрунтування діапазону значень конструктивних параметрів, кодування очікуваного вихідного сигналу навчальних прикладів із можливістю врахування близькості еталонів, навчання нейромережевої моделі з використанням експертних даних, обґрунтованої мінімізації кількості вхідних параметрів, що забезпечило можливість розробки ефективних моделей і методів обробки біометричних параметрів та методів нейромережевого розпізнавання емоційного стану та особи слухача системи дистанційного навчання.

4. Отримали подальший розвиток моделі обробки біометричних параметрів, асоційованих з зображеннями, що за рахунок врахування особливостей реєстрації вказаних зображень в умовах системи дистанційного навчання, вимог до підготовки

вхідних даних згорткової нейронної мережі та за рахунок обґрунтованого застосування технологій комп'ютерного зору дозволяють нівелювати типові завади і викривлення та забезпечують теоретичний базис розробки методів формування вхідного поля нейромережевої моделі розпізнавання емоцій слухачів. На відміну від відомих розроблені моделі формалізують процеси центрування, масштабування, бінаризації, виділення контуру, фільтрації та виділення контрольних точок в комплексі для кожного із біометричних параметрів. Перспективи подальшого розвитку моделей обробки співвідносяться з інтеграцією в них процедур фільтрації на основі методів вейвлет-перетворень, що дозволить підвищити ефективність обробки зашумлених зображень. Крім того, перспективним шляхом вдосконалення моделі обробки райдужної оболонки ока є її адаптація до підготовки вхідних даних для нейромережевої моделі розпізнавання емоцій.

5. Отримала подальший розвиток побудована на базі капсульної нейронної мережі нейромережева модель CapsNetEm, адаптована до розпізнавання базових емоцій слухача системи дистанційного навчання з урахуванням можливого повороту обличчя на зображенні. На відміну від використаного прототипу конструктивні параметри розробленої CapsNetEm адаптовані до умов системи дистанційного навчання, пов'язаних з відсутністю необхідності відновлення повернутого зображення обличчя, що забезпечує можливість приблизно на 15% зменшити ресурсоємність мережі. Подальше вдосконалення CapsNetEm доцільно співвіднести із збільшенням в ній кількості схованих шарів нейронів, що співвідноситься з гіпотезою про можливість за рахунок збільшення рівнів ієрархії розпізнавання підвищити точність та зменшити обчислювальну ресурсоємність нейромережевих засобів.

РОЗДІЛ 3
МЕТОДИ ФОРМУВАННЯ ВХІДНОГО ПОЛЯ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ РОЗПІЗНАВАННЯ ЕМОЦІЙ СЛУХАЧІВ СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ

3.1. Метод застосування вейвлет-перетворень для фільтрації вхідного поля

Як вже було відзначено в розділі 2, вхідне поле НММ розпізнавання емоцій слухачів СДН на основі зображення обличчя слухача може бути асоційоване з БП, котрий характеризується геометричними показниками. По суті, такий БП представляє собою цифрове зображення, що є дискретним двовимірним сигналом. Типові завади, котрі значно ускладнюють процес розпізнавання таких БП, в основному викликані:

- наявністю предметів, що закривають частину зображення;
- нерівномірністю освітлення, що призводить до відблисків на зображенні.

При цьому в багатьох випадках локалізація та інші параметри типових завад мають нестаціонарний динамічний характер.

Для прикладу, на рис. 3.1 показано спотворені під впливом типових завад зображення обличчя та зображення райдужної оболонки ока. На рис. 3.1 показані типові завади, до яких відносяться шапка, окуляри та відблиски світла. Зазначимо, що не зважаючи на природу завад, їх негативний вплив полягає у тому, що на зображенні з’являються ділянки, котрі власне і спотворюють зображення. При цьому в умовах СДН характер та локалізація спотворень можуть змінюватись в часі.

Відповідно рекомендацій [29, 53, 148], представити БП, що асоціюється з напівтоновим або кольоровим зображенням та характеризується за допомогою геометричних показників, доцільно за допомогою двовимірної прямокутної або квадратної матриці. Тому для аналізу таких сигналів прийнято використовувати

двовимірне дискретне вейвлет-перетворення, яке є композицією одновимірних дискретних вейвлет-перетворень [67, 274, 283].

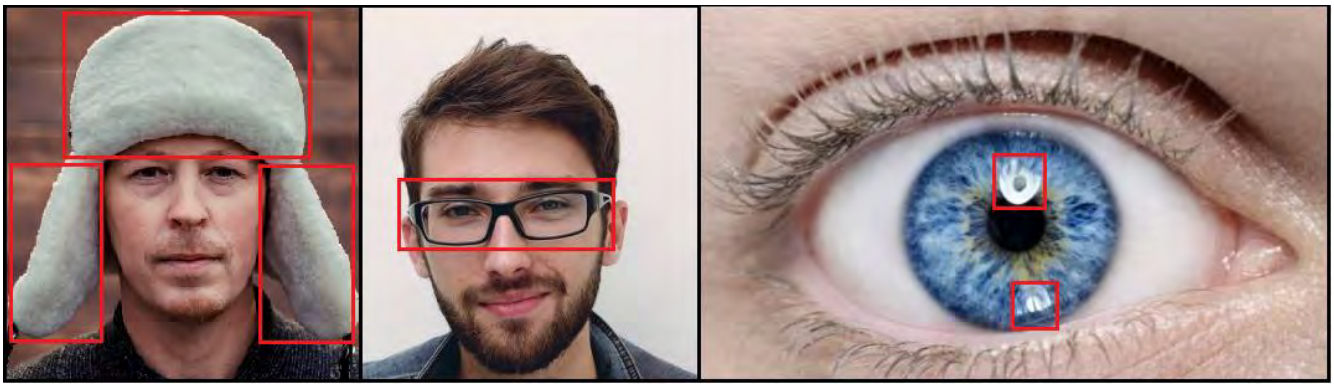

Рис. 3.1. Спотворені під впливом типових завад зображення обличчя та зображення райдужної оболонки ока

На кожному етапі деталізації спочатку одновимірні дискретні вейвлет-перетворення застосовуються для кожного рядка двовимірної матриці вхідних даних. Після цього виконуються одновимірні вейвлет-перетворення для кожного стовпця отриманої матриці [27, 35].

Математичний апарат для розрахунку матриці вейвлет-коефіцієнтів одного рядка пікселів кольорового зображення визначається виразами виду:

$$\begin{cases} \boldsymbol{W}_{m,k} = \dfrac{1}{\sqrt{a_0^m}} \displaystyle\sum_{n=0}^{N-1} \left(q(x_n)\varphi^*(a_0^m x_n - k)\right), \\ 1 \le m, k \le N \end{cases} \tag{3.1}$$

$$q = 0{,}299R + 0{,}587G + 0{,}114B, \tag{3.2}$$

де a_0 – початковий масштаб;

k – параметр зсуву;

m – параметр масштабу;

x_n – координата n-ої точки зображення;

* – операція комплексного спряження;

N – розмір зображення;

$q(x_n)$ – величина закодованого кольору зображення в точці x_n;

φ – базисний вейвлет;

R,G, B – компоненти кожного із каналів кольору в форматі RGB для окремого пікселя.

Зазначимо, що вирази (3.1, 3.2) використовуються у випадку, коли при обробці зображення його кольором можливо знехтувати. У випадку, коли це реалізувати не доцільно, вейвлет-коефіцієнти розраховуються окремо для кожного із кольорових каналів. Відповідний математичний апарат можливо записати так:

$$\begin{cases} \boldsymbol{W}_{m,k}(i) = \dfrac{1}{\sqrt{a_0^m}} \displaystyle\sum_{n=0}^{N-1} \left(q(x_n, i)\varphi^*(a_0^m x_n - k)\right), \\ 1 \le m, k \le N \end{cases} \tag{3.3}$$

де $q(x_n, i)$ – інтенсивність і-го каналу кольору в точці x_n.

При використанні RGB-моделі канали співвідносяться з червоним, зеленим та блакитним кольором.

Власне процедура фільтрації полягає у вилученні визначених компонент матриці $\boldsymbol{W}_{m,k}$. В більшості випадків вилучаються ті компоненти матриці $\boldsymbol{W}_{m,k}$, для яких параметр зсуву та/або параметр масштабу виходять за межі наперед визначеного діапазону.

Після цього компоненти відфільтрованої матриці $\boldsymbol{W}_{m,k}$ використовуються в процедурі оберненого вейвлет-перетворення для відновлення очищеного зображення.

Для напівтонового зображення використовується вираз (3.4), для багатокольорового зображення – вираз (3.5).

$$q(x_n) = \frac{\pi}{ln(a_0)} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} (\varphi^*(x_n)\boldsymbol{W}_{m.k}), \tag{3.4}$$

$$q(x_n, i) = \frac{\pi}{ln(a_0)} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} \left(\varphi^*(x_n) \boldsymbol{W}_{m.k}(i)\right), \tag{3.5}$$

де i – номер кольорового каналу.

Базуючись на сформованих в [154, 188, 288, 318] вимогах, які забезпечують ефективність вейвлет-перетворень, визначено доцільність використання діадного дискретного вейвлет-перетворення, що за рахунок зменшення кількості копій базисного вейвлету дозволяє зменшити обчислювальну складність процедури обчислення вейвлет-коефіцієнтів.

Застосування діадного дискретного вейвлет-перетворення призводить до модифікації виразів (3.1, 3.3-3.5):

$$\begin{cases} \boldsymbol{W}_{m,k} = \dfrac{1}{\sqrt{2^m}} \displaystyle\sum_{n=0}^{N-1} \left(q(x_n) \varphi^*(2^m x_n - k)\right), \\ 1 \le m, k \le N \end{cases} \tag{3.6}$$

$$\begin{cases} \boldsymbol{W}_{m,k}(i) = \dfrac{1}{\sqrt{2^m}} \displaystyle\sum_{n=0}^{N-1} \left(q(x_n, i) \varphi^*(2^m x_n - k)\right), \\ 1 \le m, k \le N \end{cases} \tag{3.7}$$

$$q(x_n) = \frac{\pi}{ln(2)} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} (\varphi^*(x_n) \boldsymbol{W}_{m.k}), \tag{3.8}$$

$$q(x_n, i) = \frac{\pi}{ln(2)} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} \left(\varphi^*(x_n) \boldsymbol{W}_{m.k}(i)\right). \tag{3.9}$$

Труднощі створення ефективних фільтрів багато в чому пов’язані зі складністю розрахунку меж діапазонів, які визначають перелік компонентів матриці $\boldsymbol{W}_{m,k}$, котрі підлягають вилученню. Це пояснюється динамічним

характером завад для одного і того ж об’єкту в різні моменти часу реєстрації біометричних параметрів.

Наприклад, як це показано на рис. 3.2, в момент часу t_1 в лівому нижньому куті райдужної оболонки ока виникає завада типу «відблиск», пов’язана з нерівномірністю освітлення. В момент часу t_2 через зміну освітлення подібна завада виникає вже у правому верхньому куті райдужної оболонки ока. Тому координати місцеположення завад змінюється з плином часу.

Рис. 3.2. Ілюстрація динамічного характеру локалізації завад

Разом з тим, динамічний характер локалізації однієї або декількох завад дозволяє стверджувати, що в кожен окремий момент часу на зареєстрованому зображенні можуть бути як викривлені, так і невикривлені частини зображення, котре асоційоване з біометричним параметром. При цьому місцеположення цих частин змінюється. Це дозволяє стверджувати, що за рахунок порівняння зображень, зареєстрованих у різні моменти часу, можливо відтворити незашумлене зображення.

Вказаний постулат дозволяє запропонувати підхід до розробки методу фільтрації БП. В базовому варіанті підхід передбачає:

- Розрахунок значень вейвлет-коефіцієнтів кожного із зареєстрованих зображень.

- Попарну інтеграцію розрахованих значень вейвлет-коефіцієнтів між собою.

- Реалізацію зворотнього вейвлет-перетворення, що призводить до отримання відфільтрованого невикривленого зображення.

Зазначимо, що попарна інтеграція вейвлет-коефіцієнтів, здійснена з урахуванням необхідності нівелювання завад, може бути реалізована за допомогою одного із принципів:

• 'max' - вибір максимального з двох коефіцієнтів;

• 'min' - вибір мінімального з двох коефіцієнтів;

• 'mean' - вибір середнього значення з двох коефіцієнтів;

• 'img1' - вибір коефіцієнта першого зображення;

• 'img2' - вибір коефіцієнта другого зображення;

• 'rand' - випадковий вибір з двох коефіцієнтів.

Також при створені методу фільтрації використано результати роботи [186, 188], в якій наведено метод вибору найбільш ефективного базисного вейвлету, що відповідає необхідності досягнення достатньої точності спектрального аналізу при мінімальних обчислювальних витратах. Приклад фільтрації зображення обличчя, з нестаціонарними завадами, викликаними нерівномірністю освітлення, реалізований відповідно наведеного підходу за допомогою засобів програмного комплексу MATLAB, показано на рис. 3.3.

Базуючись на запропонованому підході та результатах [186, 279, 319, 326], розроблено метод фільтрації БП, що асоціюються з двовимірними кольоровими або напівтоновими зображеннями. В базовому випадку реалізація методу полягає у виконанні наступних етапів:

1. Провести реєстрацію відеоряду, що відповідає послідовному відображенню зображень, котрі асоціюються з певним БП.

2. Визначити тип зображення, що підлягає фільтрації. Мається на увазі необхідність аналізу кольорового, напівтонового (з одним кольоровим каналом) або чорно-білого (монохромного) зображення.

3. Визначити перелік підконтрольних зображень відеоряду, що відповідають зображенням, у яких викривлення локалізовані у різних частинах.

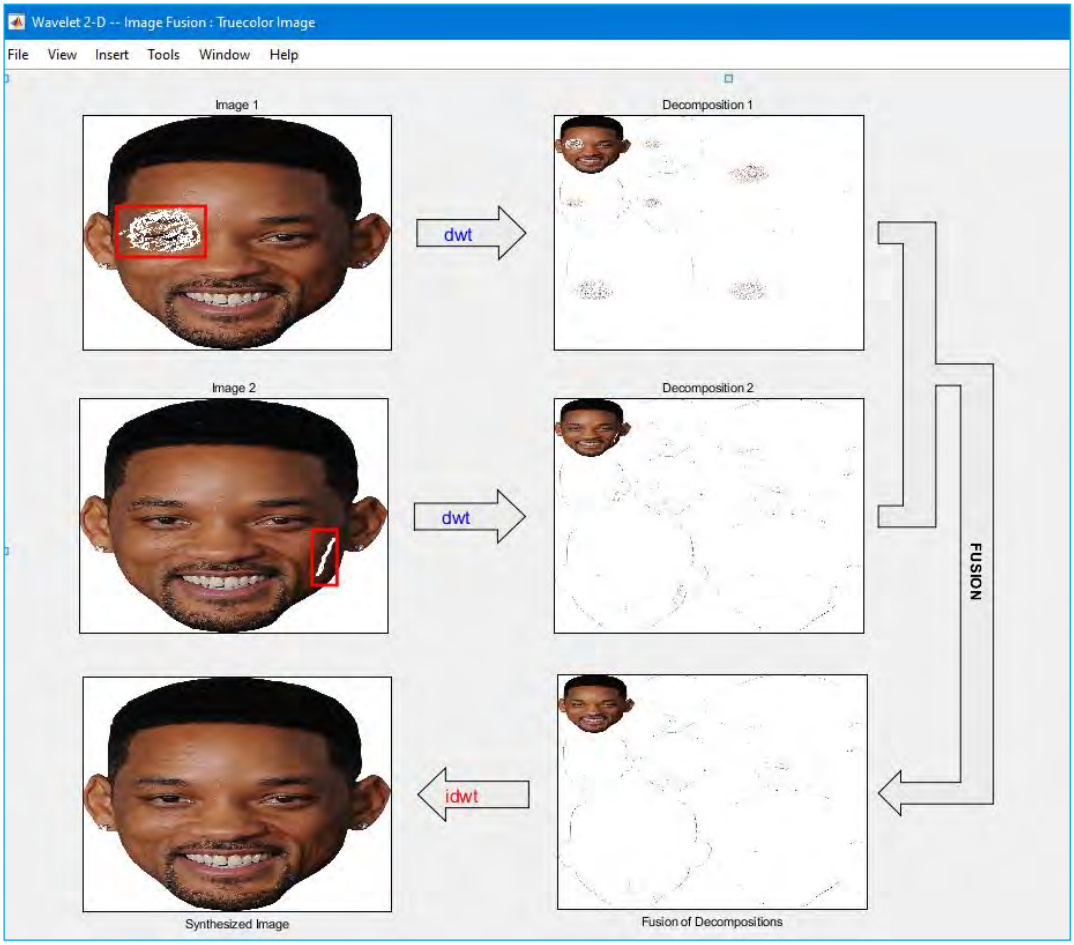

Рис. 3.3. Приклад фільтрації зображення обличчя засобами MATLAB

4. Визначити перелік типових завад, що притаманні певним БП.

5. Реалізувати вибір найбільш ефективного базисного вейвлету.

6. Визначити необхідний рівень деталізації вейвлет-перетворень.

7. Розрахувати значення вейвлет-коефіцієнтів кожного із підконтрольних зображень.

8. Обрати принцип інтеграції вейвлет-коефіцієнтів.

9. Попарно інтегрувати розраховані значення вейвлет-коефіцієнтів між собою.

10. Використовуючи отримані значення вейвлет-коефіцієнтів провести реалізацію зворотнього вейвлет-перетворення, що призводить до отримання незашумленого зображення.

Особливості наведеного методу пов’язані з тим, що:

- При реалізації другого, третього та шостого етапів слід враховувати особливості поставленої задачі розпізнавання БП.

- При реалізації сьомого та десятого етапів використовуються вирази (3.6, 3.7) та (3.8, 3.9).

- Для визначення найбільш ефективного типу базисного вейвлету можливо застосувати наведений в роботах [90, 186] підхід, в якому обґрунтовано доцільність використання базисного вейвлету, геометрична форма якого найбільш подібна до геометрії процесу, що аналізується. При цьому забезпечується мінімізація обсягу обчислень при вейвлет-перетвореннях та досягається максимальна точність відтворення зображення після зворотнього вейвлет-перетворення.

За аналогією з відомими методами розробки засобів прогнозування параметрів комп'ютерних систем [26, 188, 288, 326], запропонований п’ятий етап методу фільтрації БП на основі вейвлет-перетворень базується на наступних положеннях:

- Критерії вибору найбільш ефективного типу базисного вейвлету повинні відображати міру его пристосованості до поставленої задачі фільтрації.

- Під k-им критерієм визначення найбільш ефективного типу базисного вейвлету будемо розуміти міру забезпечення в даному типі вейвлету k-ої вимоги задачі фільтрації;

- Розрахунок ефективності і-го типу базисного вейвлету можливо представити у вигляді:

$$R_i = \sum_{k=1}^{K} \alpha_k \, r_k(i), \tag{3.13}$$

де R_i- показник інтегральної ефективності і-го типу базисного вейвлету;

$r_k(i)$- k-ий критерій ефективності і-го типу материнського вейвлету;

k- ваговий коефіцієнт k-го критерію ефективності;

K - кількість критеріїв ефективності.

- Найбільш ефективний тип базисного вейвлету визначається виразом вигляду:

$$R_{eff} = max\{R\}_I, \tag{3.14}$$

де $\{R\}_I$- множина показників інтегральної ефективності типів базисних вейвлетів;

I - потужність множини.

Використовуючи аналогію з загальноприйнятою методологією обробки зображень [11, 78, 113], для звуження номенклатури варіантів пошуку найбільш ефективного типу базисного вейвлету запропоновано обмежитись дослідженням вейвлетів HAAR, MHAT, WAVE, FHAT, LP, Daubechies wavelet, Morlet, які визначаються виразами виду (3.15-3.20).

$$\varphi(x) = \begin{cases} 1, x \in [0; 0{,}5] \\ -1, x \in [0{,}5; 1], \\ 0, x \in [0; 1] \end{cases} \tag{3.15}$$

$$\varphi(x) = \frac{2}{\sqrt{3}}\pi^{-0{,}25}(1 - x^2)e^{-0{,}5x^2} \tag{3.16}$$

$$\varphi(x) = xe^{-0{,}5x^2}, \tag{3.17}$$

$$\varphi(x) = \begin{cases} 1, x \in \left]-\infty, \frac{1}{3}\right] \\ -0{,}5, x \in \left]\frac{1}{3}, 1\right], \\ 0, x \in]1, \infty[\end{cases} \tag{3.18}$$

$$\varphi(x) = (\pi x)^{-1}\big(sin(2\pi x) - sin(\pi x)\big), \tag{3.19}$$

$$\varphi(x) = \sqrt{2}\sum_{k} g_k\, \phi(2x - k), \tag{3.20}$$

$$\phi(x) = \sqrt{2}\sum_{k} h_k\, \phi(2x - k), \tag{3.21}$$

$$\varphi(x) = e^{ik_0 x} e^{-0{,}5x^2}, \tag{3.22}$$

де φ – функція, що визначає базисний вейвлет;

ϕ – рівняння, що визначає розтягування Daubechies wavelet;

g_k, h_k – коефіцієнти, що використовуються при побудові Daubechies wavelet;

k_0 – коефіцієнт, що використовується при побудові Morlet-вейвлету.

В результаті аналізу [26, 62, 90, 188] сформовано множину критеріїв ефективності базисного вейвлету. Перелік цих критеріїв наведено в табл. 3.1.

Таблиця 3.1

Критерії ефективності типу базового вейвлету

Критерій	**Опис**
1	2
r_1	Наявність у вейвлет-коефіцієнтах надлишкової інформації
r_2	Відсутність у вейвлет-коефіцієнтах надлишкової інформації
r_3	Можливість реалізації швидкого вейвлет-перетворення
r_4	Наявність нескінченної регулярності
r_5	Наявність випадкової регулярності
r_6	Симетричність базисної функції
r_7	Асиметричність базисної функції

Таблиця 3.1 (продовження)

1	2
r_8	Ортогональність базисної функції
r_9	Наявність функції масштабування
r_{10}	Можливість повного відновлення сигналу
r_{11}	Компактність базисної функції
r_{12}	Подібність геометрії базисної функції з геометрією процесу, що підлягає аналізу

За аналогією з [26, 186, 188], прийнято, що в першому наближенні значення запропонованих критеріїв від r_1 до r_{11} можуть бути оцінені по двобальній дискретній шкалі. При цьому для *i*-го типу базисного вейвлету значення *k*-го критерію дорівнює 1, якщо відповідна *k*-та вимога повністю забезпечується, та дорівнює 0, якщо не забезпечується. Розраховані значення критеріїв для апробованих типів базисних вейвлетів наведені в табл. 3.2.

Таблиця 3.2

Значення критеріїв ефективності для апробованих типів базисних вейвлетів

Тип базисного вейвлету	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}	r_{11}
WAVE-вейвлет	0	1	0	1	0	1	0	0	0	1	0
MHAT	0	1	0	1	0	1	0	0	0	1	0
LP-вейвлет	0	1	0	0	1	0	1	0	0	0	0
Daubechies wavelet	1	0	1	0	1	0	1	1	1	1	1
HAAR	1	0	1	0	1	0	1	1	1	1	1
FHAT	1	0	1	0	1	1	0	0	0	1	0
Morlet	0	1	0	1	0	1	0	0	0	1	0

Слід зазначити, що апріорно визначити значення критерію ефективності r_{12}, який визначає подібність геометрії базисної функції з геометрією процесу, що підлягає аналізу, досить важко. Тому в табл. 3.2 його значення відсутнє.

Використання розроблених базових положень визначення найбільш ефективного виду базисного вейвлету і результати робіт, присвячених побудові ефективних нейромережевих моделей оцінки параметрів, дозволили розділити п'ятий етап методу фільтрації БП, асоційованих з напівтоновим або кольоровим зображенням, на наступні кроки:

1. Формалізація умов завдання з метою визначення виду і значень параметрів, що визначають завади на БП, асоційованому з зображенням. При цьому слід реалізувати подання кожного рядка/стовпця пікселів зображення, асоційованого з БП, у вигляді графіка у декартовій системі координат.

2. Визначення значення критерію r_{12} для кожного з апробованих типів базисних вейвлетів, заданих виразами (3.16-3.20). Визначення полягає в оцінці близькості геометрії форми базисного вейвлету і геометрії графіків, асоційованих з рядками пікселів зображення БП.

3. Визначення значущості кожного із критеріїв ефективності, представлених в табл. 3.1. Визначення можливо реалізувати шляхом експертного оцінювання.

4. Розрахунок ефективності кожного типу базисного вейвлету. Для цього слід використати вираз (3.13).

5. Визначення за допомогою виразу (3.14), найбільш ефективного типу базисного вейвлету.

Загальна схема розробленого методу фільтрації БП, що можуть бути представлені у вигляді напівтонового або кольорового зображення, наведена на рис. 3.4.

На базі даного методу з використанням засобів MATLAB розроблене відповідне інструментальне забезпечення. Проведені з урахуванням даних [50, 113, 274] дослідження дозволяють стверджувати, що застосування запропонованого

методу в умовах СДН дозволяє приблизно на 7-10% підвищити суб'єктивно оцінену якість фільтрації БП, асоційованих з зображеннями.

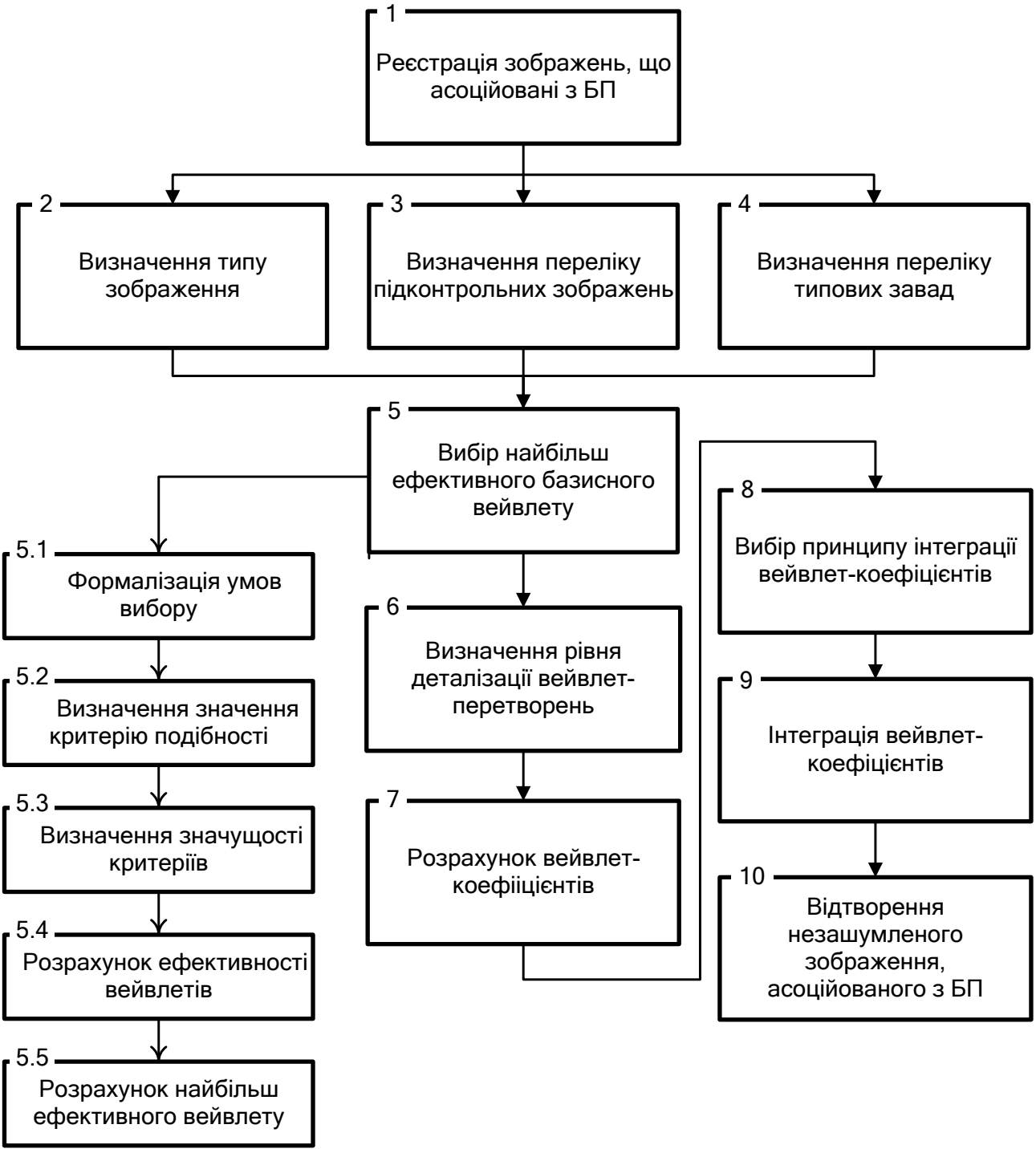

Рис. 3.4. Структурна схема методу фільтрації біометричних параметрів

3.2. Метод формування вхідного поля нейромережевої моделі аналізу біометричних параметрів, асоційованих з зображеннями

Практичний досвід і дані літературного аналізу [52, 53, 63] свідчать про те, що модернізація сучасних НМЗ розпізнавання емоцій та особи користувача ІС різноманітного призначення в значній мірі корелюється з вирішенням завдання підвищення ефективності нейромережевих технологій аналізу БП, які характеризують біологічні, психофізіологічні і поведінкові якості людини. Оскільки в останні 5-10 років розвиток цих технологій пов'язують з комп'ютерним зором, то саме такі технології використані в якості базису для розробки методу формування вхідного поля НММ аналізу БП, асоційованих з зображеннями.

Оскільки в більшості випадків для представлення кольору пікселів таких зображень використовується формат RGBA або формат RGB, то в початковому випадку аналітичну модель методу формування вхідного поля НММ для аналізу БП, асоційованих з зображеннями, можливо записати так:

$$\langle \boldsymbol{q_x}, \boldsymbol{q_y}, \boldsymbol{q_R}, \boldsymbol{q_G}, \boldsymbol{q_B}, \boldsymbol{q_A} \rangle \rightarrow \boldsymbol{\Psi}^{Im}, \tag{3.23}$$

$$\boldsymbol{\Psi}^{Im} = \begin{vmatrix} \boldsymbol{\psi}_{1,1} & \dots & \boldsymbol{\psi}_{1,N} \\ \dots & \dots & \dots \\ \boldsymbol{\psi}_{N,1} & \dots & \boldsymbol{\psi}_{N,N} \end{vmatrix}, \tag{3.24}$$

де $\boldsymbol{q_x}$, $\boldsymbol{q_y}$ – множина показників, що характеризують координати пікселів зареєстрованого кольорового зображення;

$\boldsymbol{q_R}$, $\boldsymbol{q_G}$, $\boldsymbol{q_B}$, $\boldsymbol{q_A}$ – множина показників, що характеризують канали кольору кожного із пікселів зареєстрованого зображення;

$\boldsymbol{\Psi}^{Im}$ – множина вхідних параметрів ЗНМ, що призначена для аналізу зображень;

N – розмір вхідного поля ЗНМ.

У випадку напівтонового або монохромного зображення вираз (3.23) спрощується так:

$$\langle \boldsymbol{q_x}, \boldsymbol{q_y}, \boldsymbol{q_M} \rangle \rightarrow \boldsymbol{\Psi}^{Im}, \quad (3.25)$$

де $\boldsymbol{q_M}$ – множина показників, що характеризують колір кожного із пікселів зареєстрованого напівтонового або монохромного зображення.

При цьому, враховуючи особливості реєстрації зображень та запропонований підхід до мінімізації ресурсоємності, підготовка значень зареєстрованих параметрів до виду, який є прийнятним до подання в якості вхідних даних ЗНМ, повинна включати в себе наступні етапи:

1. Центрування зареєстрованого зображення, що спричинене неможливістю на практиці достатньо точно зорієнтувати засоби реєстрації. Центрування полягає у застосуванні афінних перетворень для зіставлення центра та осей зображення БП з центром та осями вхідного поля ЗНМ.

2. Обмеження розміру зображення за рахунок обрізки малозначимих частин масштабованого зображення.

3. Масштабування обмеженого зображення до розміру, що відповідає вхідному полю ЗНМ.

4. Первинна обробка масштабованого зображення, що у загальному випадку полягає у застосуванні набору фільтрів для згладжування кольорової гами, перетворенні кольорового зображення у монохромне (бінаризація) та скелетизації для виділення значимих ознак.

5. Нівелювання негативного впливу типових завад.

6. Визначення параметрів ключових точок, що використовуються для розпізнавання емоцій.

7. Визначення параметрів контрольних точок, що використовуються для розпізнавання особи.

8. Нормалізація параметрів зображення та нормалізація параметрів ключових/контрольних точок.

9. Кодування параметрів зображення та параметрів ключових/контрольних

точок до виду, придатному до подачі в ЗНМ.

Узагальнена схема методу формування вхідного поля ЗНМ аналізу БП, асоційованих з зображеннями показана на рис. 3.5.

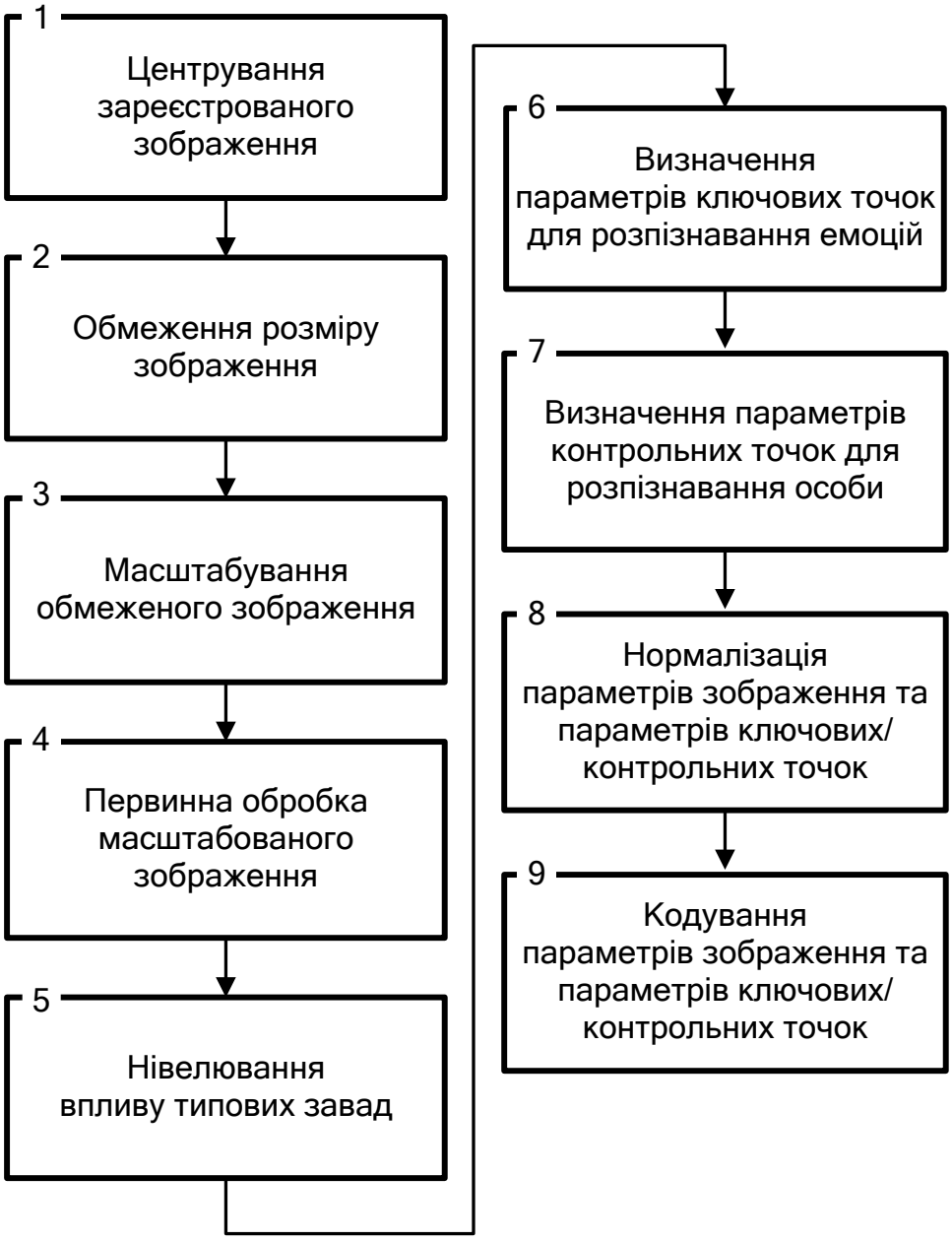

Рис. 3.5. Узагальнена схема методу формування вхідного поля НММ аналізу БП, асоційованих з зображеннями

Зазначимо, що відповідно результатів першого розділу, деталізацію етапів

методу доцільно співвіднести з особливістю відображення кожного із перерахованих БП: зображення обличчя, вушної раковини, райдужної оболонки ока, сітківки ока та відбитку пальця.

Тому аналітичну модель (3.23) доцільно записати у вигляді:

$$\langle \boldsymbol{B}, \{f_{reg}\}, \boldsymbol{U}_{img}, \boldsymbol{U}_{CNN}, H_{CNN}, L_{CNN}, \boldsymbol{q_x}, \boldsymbol{q_y}, \boldsymbol{q_R}, \boldsymbol{q_G}, \boldsymbol{q_B}, \boldsymbol{q_A} \rangle \rightarrow \boldsymbol{\Psi}^{Im}, \quad (3.26)$$

$$\boldsymbol{B} = \{b_1, b_2, b_3, b_4, b_5\}, \quad (3.27)$$

де $\boldsymbol{B}$ – множина типів БП, асоційованих з зображеннями;

b_1 – зображення обличчя;

b_2 – зображення вушної раковини;

b_3 – відбиток пальця;

b_4 – райдужна оболонка ока;

b_5 – сітківка ока;

$\{f_{reg}\}$ – множина значень параметрів зареєстрованого зображення;

$\boldsymbol{U}_{img}$ – множина вимог до зображень, асоційованих з БП;

$\boldsymbol{U}_{CNN}$ – множина вимог до вхідних параметрів ЗНМ;

H_{CNN}, L_{CNN} – висота та ширина вхідного поля ЗНМ.

Проведені дослідження дозволили деталізувати етапи методу формування вхідного поля НММ аналізу БП, асоційованих з зображеннями.

Етап 1. Центрування зареєстрованого зображення. Обробка зображення, що відбувається на даному етапі полягає у розміщенні зареєстрованого зображення в центрі вхідного поля та зміни просторової орієнтації в необхідному напрямку. Входом етапу являються $b \in \boldsymbol{B}, \boldsymbol{U}_{img}, \boldsymbol{U}_{CNN}, \{f_{reg}\}$. Реалізація етапу розділяється на наступні кроки.

Крок 1.1. Визначення зовнішніх кордонів зображення, що реалізується за допомогою методу Віоли-Джонса. Якщо зовнішні кордони зображення визначити не вдалось, то вважається, що якість зареєстрованого зображення незадовільна і

формування вхідного поля ЗНМ неможливе.

Крок 1.2. Визначення центру зображення. Реалізується з використанням ітераційного алгоритму з використанням (2.24-2.27) або у випадку спрощених розрахунків за допомогою виразів (2.28-2.31).

Крок 1.3. Розрахунок величини лінійного зсуву, що здійснюється шляхом співставлення центру зображення з центром вхідного поля:

$$\Delta_x = x_c - x_0, \tag{3.28}$$

$$\Delta_y = y_c - y_0, \tag{3.29}$$

де Δ_x, Δ_y – лінійний зсув по осі абцис та по осі ординат;

x_c, y_c – координати центру зареєстрованого зображення;

x_0, y_0 – координати центру вхідного поля ЗНМ.

Зазначимо, що центром вхідного поля ЗНМ є точка, що знаходиться на перетині діагоналей прямокутника, розміри якого являються конструктивними параметрами такої НММ.

Крок 1.4. Перенесення зображення. Для кожної точки зображення процедура перенесення визначається виразами:

$$x_{shift}(i) = x_{initial}(i) - \Delta_x, \tag{3.30}$$

$$y_{shift}(i) = y_{initial}(i) - \Delta_y, \tag{3.31}$$

де $x_{initial}(i)$, $y_{initial}(i)$ – координати і-ої точки зареєстрованого зображення;

$x_{shift}(i)$, $y_{shift}(i)$ – координати і-ої точки перенесеного зображення.

Крок 1.5. Визначення осі симетрії зображення, що асоціюється з БП. Реалізується за допомогою ітераційного підходу, математичний апарат якого визначається виразами (2.32-2.37).

Крок 1.6. Розрахунок кута повороту. Оскільки вісь симетрії зображення асоційованого з БП та вертикальна вісь вхідного поля ЗНМ це прямі лінії, що задані рівняннями типу $ax + by + c = 0$ та перетинаються в центрі вхідного поля ЗНМ, котре співпадає з центром зображення, то для розрахунку величини кута між

вказаними осями використовується вираз:

$$\varphi = \cos^{-1}\left(\frac{a_1 a_2 + b_1 b_2}{\sqrt{a_1^2 + b_1^2}\sqrt{a_2^2 + b_2^2}}\right), \quad (3.32)$$

де φ – шуканий кут повороту;

a_1, b_1 – коефіцієнти рівняння, що описує вісь симетрії зображення, асоційованого з БП;

a_2, b_2 – коефіцієнти рівняння, що описує вертикальна вісь вхідного поля ЗНМ;

$\cos^{-1} z$ – функція арккосинуса числа z.

Крок 1.7. Поворот зображення. Розрахунок координат точок повернутого зображення здійснюється за допомогою виразів:

$$x_{\varphi,i} = x_i \cos\varphi - y_i \sin\varphi, \quad (3.33)$$

$$y_{\varphi,i} = x_i \cos\varphi + y_i \sin\varphi, \quad (3.34)$$

де x_i, y_i – координати і-ої точки початкового зображення;

$x_{\varphi,i}$, $y_{\varphi,i}$ – координати і-ої точки повернутого зображення.

Виходом першого етапу являється $\{f_c\}$ – множина точок відцентрованого та повернутого зображення.

Етап 2. Обмеження розміру зображення. Для цього підлягають видаленню частини зображення, що виходять за наведені в табл. 2.1 максимально допустимі розміри зареєстрованого зображення асоційованого з БП. На вхід етапу подаються множини $\{f_c\}$ та $\boldsymbol{U}_{img}$.

Враховуючи, що початок координат вхідного поля співпадає з центром цього поля, математичний апарат формування множини точок обрізки задано виразами:

$$if\ \ x_\tau < -0{,}5L_{reg} \rightarrow \tau(x,y) \in \{Cut\}, \quad (3.35)$$

$$if\ \ x_\tau > 0{,}5L_{reg} \rightarrow \tau(x,y) \in \{Cut\}, \quad (3.36)$$

$$if\ \ y_\tau < -0{,}5H_{reg} \rightarrow \tau(x,y) \in \{Cut\}, \tag{3.37}$$

$$if\ \ y_\tau > 0{,}5H_{reg} \rightarrow \tau(x,y) \in \{Cut\}, \tag{3.38}$$

де $\tau(x,y)$ – точка відцентрованого зображення з координатами x, y;

$\{Cut\}$ – множина точок відцентрованого зображення, що мають бути видалені;

L_{reg}, H_{reg} – максимального допустимі значення ширини та висоти зареєстрованого зображення, асоційованого з БП.

Передбачається, що значення L_{reg}, H_{reg} задані в множині вимог $\boldsymbol{U}_{img}$.

Етап розділяється на два кроки.

Крок 2.1. Формування множини точок, що мають бути видалені. Формування реалізується з використанням виразів (3.35-3.38). Результатом даного кроку є множина $\{Cut\}$.

Крок 2.2. Формування обмеженого зображення. Реалізація даного кроку може бути записана за допомогою виразу:

$$\{f_c\} - \{Cut\} = \{f_{cut}\}, \tag{3.39}$$

де $\{f_{cut}\}$ – множина точок обмеженого зображення.

Виходом етапу є $\{f_{cut}\}$.

Етап 3. Масштабування обмеженого зображення. Цей етап реалізується у випадку, коли розміри зображення, отриманого на вхід цього етапу, перевищують розмір вхідного поля ЗНМ, або коли розміри такого зображення менші від допустимих. На вхід етапу подаються $\{f_{cut}\}$, H_{CNN}, L_{CNN}, $\boldsymbol{U}_{CNN}$. Відповідно до означеного завдання, даний етап розділено на 5 кроків.

Крок 3.1. Визначення розміру обмеженого зображення. Для цього спочатку слід визначити максимальні та мінімальні координати точок, що входять до множини $\{f_{cut}(x,y)\}$:

$$x^{max} = \max_x(\{f_{cut}(x,y)\}), \tag{3.40}$$

$$y^{max} = \max_{y}(\{f_{cut}(x,y)\}), \tag{3.41}$$

$$x^{min} = \min_{x}(\{f_{cut}(x,y)\}), \tag{3.42}$$

$$y^{min} = \min_{y}(\{f_{cut}(x,y)\}), \tag{3.43}$$

Після цього розраховуються розміри обмеженого зображення:

$$L_{cut} = \left|x^{max} - x^{min}\right|, \tag{3.44}$$

$$H_{cut} = \left|y^{max} - y^{min}\right|, \tag{3.45}$$

де L_{cut}, H_{cut}- ширина та висота обмеженого зображення.

Крок 3.2. Перевірка мінімального розміру. Крок призначено для перевірки того, чи перевищують розміри зображення мінімально допустимий розмір об'єкта, що подається на вхід ЗНМ, призначеної для аналізу БП. Процес перевірки визначається виразом:

$$if\ (L_{cat} < L_{min}) \vee (H_{cat} < H_{min}) \Rightarrow stop, \tag{3.46}$$

де L_{min}, H_{min} – мінімально допустимі ширина та висота зображення.

Мінімально допустимі значення ширини та висоти зображення, задані в $\boldsymbol{U}_{CNN}$, залежать від архітектурних параметрів ЗНМ та необхідної точності розпізнавання. В першому наближенні прийнято $L_{min} = H_{min}$=32 пікселя.

Якщо права частина виразу (3.46) істинна, то формувати вхідне поле ЗНМ на основі поданого зображення недоцільно.

Крок 3.3. Розрахунок коефіцієнтів зміни розміру. Вказані коефіцієнти розраховуються за допомогою виразів:

$$m_x = \frac{L_{CNN}}{L_{cut}}, \tag{3.47}$$

$$m_y = \frac{H_{CNN}}{H_{cut}}, \tag{3.48}$$

де m_x, m_y – коефіцієнти зміни розміру по осі абцис та по осі ординат;

L_{CNN}, H_{CNN} – ширина та висота вхідного поля ЗНМ.

Крок 3.4. Перевірка рівня геометричних спотворень. На даному кроці визначається рівень геометричних спотворень обмеженого зображення та приймається рішення про можливість чи неможливість подальшого формування вхідного поля ЗНМ. Функціонал даного кроку можливо записати у вигляді:

$$if\ \left(m_x > (1 \pm \delta)\right) \wedge \left(m_y < (1 \pm \delta)\right) \Longrightarrow stop, \tag{3.49}$$

$$if\ \left(m_x < (1 \pm \delta)\right) \wedge \left(m_y > (1 \pm \delta)\right) \Longrightarrow stop, \tag{3.50}$$

де δ – коефіцієнт допустимого викривлення.

В першому наближенні $\delta = 0{,}1$.

Істинність правих частин виразів (3.49, 3.50) свідчить про неможливість формування вхідного поля ЗНМ на основі зображення, поданого на вхід методу.

Крок 3.5. Масштабування зображення. Даний крок передбачає, що зображення підлягає пропорційному масштабуванню. Тобто масштабний коефіцієнт по осі абцис дорівнює масштабному коефіцієнту по осі ординат. При цьому у випадку необхідності збільшення зображення, тобто при $m_x > 1$, в якості масштабного коефіцієнту обирається менший із коефіцієнтів m_x та m_y:

$$if\ (m_x > 1) \wedge \left(m_x \geq m_y\right) \Longrightarrow m = m_y\ \ else\ m = m_x. \tag{3.51}$$

У випадку необхідності зменшення зображення, тобто при $m_x < 1$, обирається більший із коефіцієнтів:

$$if\ (m_x < 1) \wedge \left(m_x \geq m_y\right) \Longrightarrow m = m_x\ \ else\ m = m_y. \tag{3.52}$$

Базуючись на результатах, визначено, що, з позицій досягнення допустимої якості обробки та спрощення розрахунків, величину масштабного коефіцієнту m, визначену за допомогою (3.51, 3.52), слід округлити до меншого числа, яке повинно дорівнювати 2^r, де r – ціле число. Тобто:

$$m = RoundMin(m, 2^r). \tag{3.53}$$

Після визначення масштабного коефіцієнту проводиться масштабування зображення, що реалізується за допомогою апарату бікубічної інтерполяції, що

наведений в [21, 114]. В результаті виконання даного кроку визначається множина точок масштабованого зображення $\{f_{scale}\}$.

Виходом третього етапу являється $\{f_{scale}\}$.

Етап 4. Первинна обробка масштабованого зображення. На вхід даного етапу подаються b, $\{f_{scale}\}$, $\boldsymbol{U}_{CNN}$. Використавши розроблені в другому розділі моделі обробки БП, асоційованих з зображеннями визначено, що виконання даного етапу поділяється на 3 кроки.

Крок 4.1. Корекція яскравості пікселів, що реалізується за допомогою виразів (2.38-2.42).

Крок 4.2. Визначення контурів на зображенні, що реалізується за допомогою виразів (2.43-2.51). Особливістю обробки зображення райдужної оболонки являється використання виразів (2.70, 2.71).

Крок 4.3. Приведення кольорового зображення, що полягає у перетворенні кольорового зображення у напівтонове або монохромне та реалізується з використанням математичного апарату (2.52-2.54). Даний крок обов'язково виконується при обробці зображень сітківки ока, відбитків пальців та вушної раковини. Для інших БП, зображення обличчя, райдужної оболонки ока, виконання даного кроку не є обов'язковим та залежить від умов задачі нейромережевого аналізу.

Виходом четвертого етапу є множина значень параметрів зображення, що пройшло первинну обробку - $\{f_{pr}\}$.

Етап 5. Нівелювання впливу типових завад. Входом етапу є b, $\{f_{pr}\}$. Виконання етапу розділено на 2 кроки.

Крок 5.1. Нівелювання завад, що закривають частину зображення. Даний крок співвідноситься з розробленим методом застосування вейвлет-перетворень для фільтрації вхідного поля, який описується за допомогою виразів (3.1-3.22).

Крок 5.2. Відновлення повернутого зображення. Даний крок виконується

виключно для відновлення повернутого зображення сітківки ока. Відновлення здійснюється за допомогою виразів (2.74-2.76).

Виходом етапу є множина значень параметрів відновленого зображення - $\{f_{filter}\}$.

Етап 6. Визначення параметрів ключових точок для розпізнавання емоцій. На вхід даного етапу подаються $b, \{f_{filter}\}, \boldsymbol{U}_{CNN}$. Відповідно до розробленої в другому розділі моделі обробки зображення обличчя, виконання даного етапу розділяється на 2 кроки.

Крок 6.1. Визначення кількості ключових точок, що здійснюється з використанням [152, 166, 167], у відповідності вимогам до вхідних параметрів ЗНМ, відображених у вхідній множині параметрів $\boldsymbol{U}_{CNN}$ даного методу. Передбачено, що мінімальна кількість таких точок (див. рис. 2.8) дорівнює 14, максимальна кількість (див. рис. 2.7) - 64.

Крок 6.2. Визначення координат ключових точок. Для цього використано апробовані алгоритми, описані в [167, 152, 212, 218].

Виходом шостого етапу являється множина значень параметрів ключових точок зображення, асоційованого з БП типу b - $\{f_{key_p}\}$.

Етап 7. Визначення параметрів контрольних точок для розпізнавання особи. Доцільність виконання даного етапу визначається у вимогах $\boldsymbol{U}_{CNN}$ і залежить від того, чи використовуються контрольні точки для розпізнавання особи, чи розпізнавання реалізується на основі аналізу зображення в цілому, тобто чи подаються параметри контрольних точок на вхід ЗНМ, чи на вхід ЗНМ подається зображення, асоційоване з БП. Подібно до шостого етапу, визначення параметрів контрольних точок регламентується вимогами. заданими у множині $\boldsymbol{U}_{CNN}$. Відмінністю від попереднього етапу являється залежність параметрів контрольних точок для розпізнавання особи від типу БП. Контрольні точки для зображення обличчя показані на рис. 2.9, для вушної раковини – на рис. 2.11, а для відбитків

пальців – на рис. 2.12. На вхід даного етапу подаються b, $\boldsymbol{U}_{CNN}$, $\{f_{filter}\}$. Етап розділяється на два кроки.

Крок 7.1. Визначення кількості контрольних точок. Для зображення обличчя використовуються підходи, наведені в [198], для відбитків пальців - [196], для вушної раковини - [214], для райдужної оболонки - [197], а для сітківки ока - [320].

Крок 7.2. Визначення координат контрольних точок, що реалізується в залежності від виду БП. Крім того, у випадку визначення координат контрольних точок на зображенні відбитків пальців додатково використовуються вирази (2.67-2.69), для райдужної оболонки ока - вирази (2.72, 2.73), а для сітківки ока – вирази (2.72, 2.73, 2.76, 2.77).

Виходом сьомого етапу являється множина значень параметрів контрольних точок зображення, асоційованого з БП типу b - $\{f_{check_p}\}$.

Етап 8. Нормалізація параметрів зображення. На вхід етапу подаються b, $\boldsymbol{U}_{CNN}$, H_{CNN}, L_{CNN} та одна із множин: $\{f_{filter}\}$, $\{f_{key_p}\}$, $\{f_{check_p}\}$. Процес нормалізації розділяється на 3 кроки.

Крок 8.1. Визначення точки початку відліку. У випадку нейромережевого аналізу всього зображення в якості точки початку відліку використовується нижній лівий кут прямокутника, що описує дане зображення. У випадку нейромережевого аналізу за контрольними та ключовими точками для обличчя використовується точка, що відповідає середині прямої лінії, яка з'єднує внутрішні кути очей (див. рис. 2.10), а для вушної раковини – козелок, позначений цифрою 11 на рис. 2.11, для відбитків пальців, райдужної оболонки ока та сітківки ока – нижній лівий кут прямокутника, що описує зображення вказаних БП.

Крок 8.2. Нормалізація координат ключових та контрольних точок, що здійснюється за допомогою виразів виду (2.62, 2.63).

Крок 8.3. Нормалізація кольору, що реалізується за допомогою виразу (2.64).

Виходом восьмого етапу являється множина нормалізованих значень $\{f_{norm}\}$.

Етап 9. Кодування параметрів зображення. На вхід даного етапу подаються $\{f_{norm}\}$, $\boldsymbol{U}_{CNN}$, H_{CNN}, L_{CNN} . Даний етап розділяється на два кроки.

Крок 9.1. Кодування геометричних параметрів, що полягає у співставленні координат кожного пікселя з координатами вхідного поля ЗНМ. При кодуванні координат ключових та контрольних точок зображення обличчя та зображення вушної раковини додатково застосовуються вирази (2.65, 2.66).

Крок 9.2. Кодування кольору. Процес кодування полягає у співставленні кольорових каналів кожного пікселя зображення з кольоровими каналами вхідного поля ЗНМ.

Виходом етапу являється $\boldsymbol{\Psi}^{Im}$ – множина параметрів, що описують вхідне поле ЗНМ, призначеної для розпізнавання емоцій або особи слухача СДН на основі аналізу одного із БП, асоційованих з зображенням.

Деталізована схема перших трьох етапів методу показана на рис. 3.6, а деталізована схема етапів 4-9 – на рис. 3.7. Математичне забезпечення методу складають вирази (3.23-3.53). Зазначимо, що перспективним напрямком вдосконалення запропонованого методу являється використання в ньому процедури кодування параметрів кожного із каналів зображення за допомогою вейвлет-перетворень, що задається виразом виду (3.54).

$$W_{i,j} = a_i^{-\frac{1}{2}} \sum_{k=1}^{p} f(x_k)\, \psi\left(\frac{x_k - b_j}{a_i}\right), \tag{3.54}$$

де ψ – породжуючий вейвлет;

p – кількість точок зображення;

x_k – номер точки зображення;

$f(x_k)$ – нормалізоване значення інтенсивності кольору пікселя, розраховане за допомогою (2.17-2.20);

a, b – масштаб та зсув вейвлету.

Така модифікація дозволить асоціювати вхідне поле ЗНМ із зображенням

вейвлет-спектрограми, що, відповідно даних [113, 186], дозволить підвищити точність та зменшити ресурсоємність НММ.

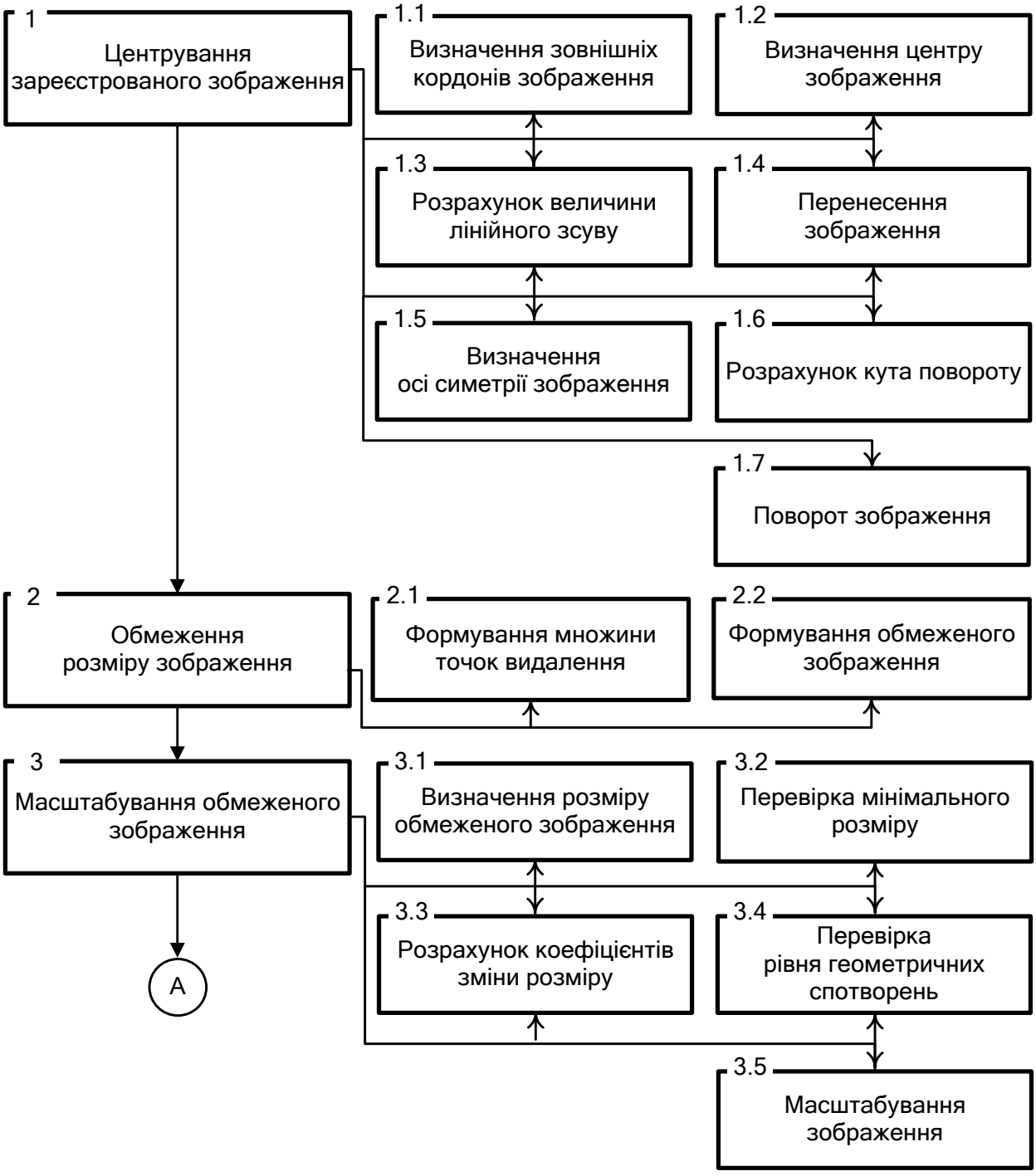

Рис. 3.6. Деталізована схема 1-3 етапів методу формування вхідного поля НММ аналізу БП, асоційованих з зображеннями

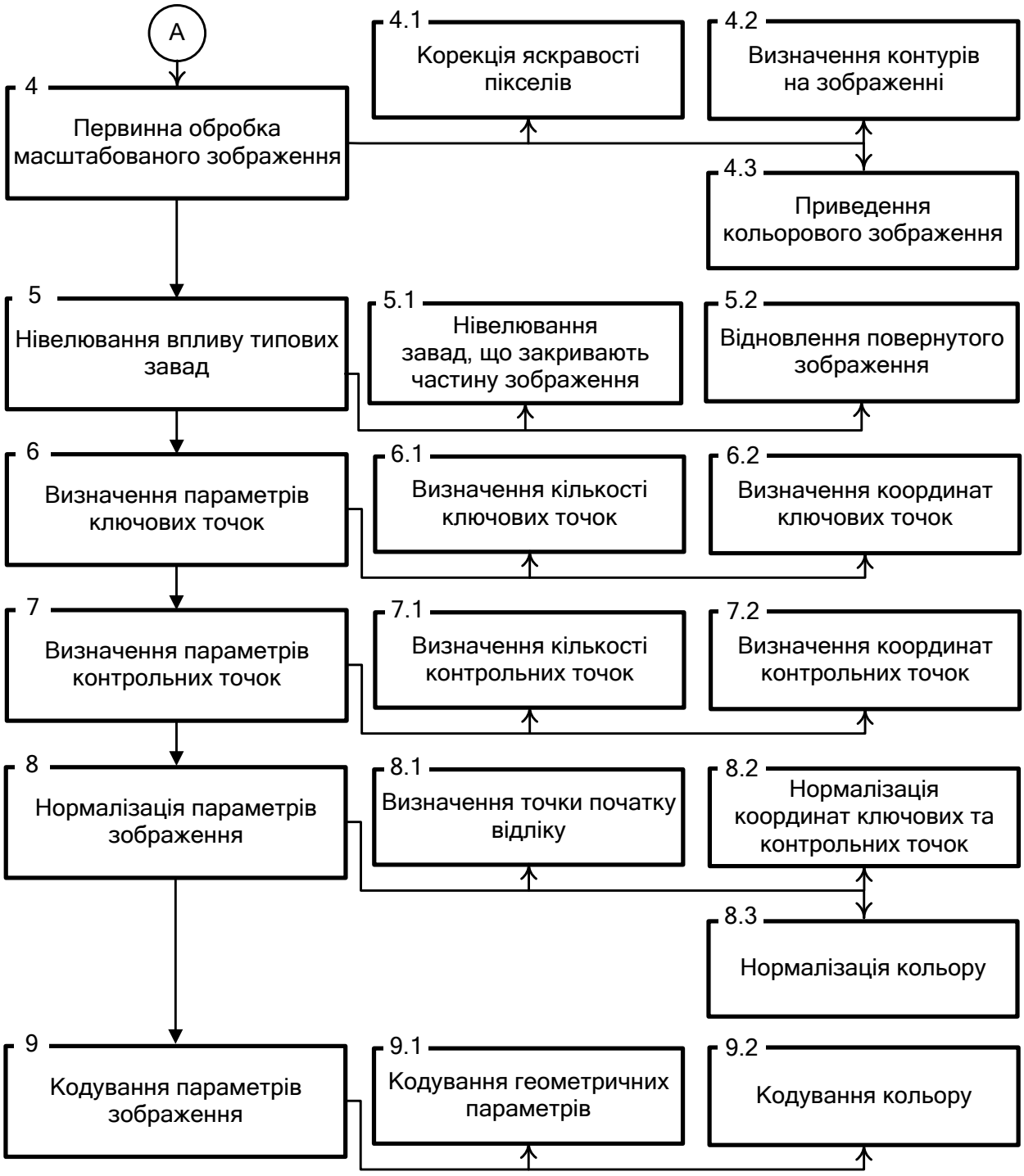

Рис. 3.7. Деталізована схема 4-9 етапів методу формування вхідного поля НММ аналізу БП, асоційованих з зображеннями

Таким чином, в результаті проведених досліджень отримав подальший розвиток метод формування вхідного поля НММ аналізу БП, асоційованих з

зображеннями, що на відміну від відомих, за рахунок адаптації процедур, пов’язаних із обробкою зображень, нівелюванням впливу типових завад та процедур визначення параметрів ключових та контрольних точок, до вимог нейромережевого аналізу зображень обличчя, вушної раковини, райдужної оболонки ока, сітківки ока та відбитку пальця, забезпечує можливість ефективного розпізнавання емоцій та особи слухача СДН.

3.3. Метод формування вхідного поля нейромережевої моделі аналізу клавіатурного почерку

Як свідчать результати досліджень, проведених в другому розділі даної дисертаційної роботи, найбільш ефективним типом НММ, що призначена для аналізу КП для розпізнавання емоцій та особи слухача СДН, є ЗНМ.

Особливістю класичного варіанту ЗНМ є необхідність представлення вхідної інформації у вигляді квадратного монохромного/кольорового двовимірного зображення.

У базовому варіанті використовується чорно-біле зображення. Більш складні варіанти передбачають використання тривимірного сірого і кольорового зображення. Зазначена особливість накладає істотне обмеження на використання ЗНМ – можливість аналізу КП на текстових фрагментах з фіксованою кількістю символів. При цьому загальна постановка задачі розпізнавання емоцій користувача ІС передбачає необхідність аналізу як заздалегідь визначеного фрагмента тексту, так і фрагмента тексту довільного змісту.

Перший випадок можна співвіднести з моніторингом емоційного стану користувача при введенні ним парольних даних.

Другий випадок співвідноситься з поточним моніторингом підтвердження особи/емоційного стану користувача при введенні ним з клавіатури текстової інформації. Відповідно до даних [45, 91, 210], такий моніторинг можливо здійснити

за рахунок аналізу КП при введенні тексту фіксованої довжини навіть при використанні відносно спрощених статистичних моделей.

Таким чином, наведене обмеження щодо фіксованої кількості символів практично не здійснює негативного впливу на функціональність ЗНМ. У першому випадку кількість символів в парольних даних і так є фіксованою величиною. У другому випадку обсяг фрагмента тексту довільного змісту, що підлягає аналізу, обмежується заздалегідь визначеною кількістю символів. Зазначимо, що відповідно результатів першого розділу, які стосуються відомих засобів аналізу КП, увагу акцентовано на розпізнавання емоцій та особи слухача СДН за рахунок аналізу ДКП. В якості параметрів, що характеризують вхідні сигнали відповідної системи нейромережевого аналізу ДКП, можуть бути використані:

- термін утримання клавіш (ТУК);

- термін між термін між послідовним відпусканням однієї клавіші та натисканням наступної клавіші (ТМК);

- динаміка терміну утримання клавіш (ДТУК);

- динаміка терміну між послідовним натисненням двох клавіш (ДТМК);

- відношення ТУК до ТМК (ВТ).

Розрахунок вказаних параметрів здійснено за допомогою виразів:

$$\tau_r(i) = t_u(i) - t_d(i)\,, \tag{3.54}$$

$$\tau_b(i, i-1) = t_u(i) - t_d(i-1)\,, \tag{3.55}$$

$$q_{br}(i, i-1) = {\tau_b(i, i-1)}\Big/{\tau_r(i)}\,, \tag{3.56}$$

$$\nu_r(i, i-1) = {\big(\tau_r(i) - \tau_r(i-1)\big)}\Big/{\tau_r\,(i)}\,, \tag{3.57}$$

$$\nu_b(i, i-1) = {\tau_b(i) - \tau_b(i-1)}\Big/{\tau_b(i)}\,, \tag{3.58}$$

де τ_r - термін утримання клавіш;

t_d - момент часу натиснення клавіші;

t_u - момент часу відпускання клавіші;

τ_b - термін між натисканням двох окремих клавіш;

q_{br} - відношення терміну утримання клавіші до терміну відпускання клавіші;

i - номер клавіші при наборі тексту;

ν_r - динаміка терміну утримання клавіш;

ν_b - динаміка терміну між послідовним натисненням двох клавіш.

При цьому в комп'ютерній системі при введені тестових даних реєструються такі параметри ДКП, як:

- ASCII-код або Unicode введеного символу;

- момент часу натиснення клавіші;

- момент часу відпускання клавіші.

Тому всі початкові дані, складові виразів (3.54-3.58), що застосовуються для розрахунку параметрів ДКП слухачів СДН, можуть бути зареєстровані за допомогою розповсюдженого універсального комп'ютерного забезпечення.

В базовому варіанті для перетворення параметрів ДКП, заданих виразами (3.54-3.58), до виду, що придатний для подачі в ЗНМ, пропонується вісь ординат співвіднести з розкладкою клавіатури - ASCII-кодами або Unicode клавіш, або символами, що відповідають клавішам. Вісь абцис пропонується співвіднести з введенням текстом. Таким чином, одній окремій точці зображення відповідатиме окремий символ введеного тексту. По осі абцис координата закодованого символу відповідає позиції (номеру) даного символу в тексті. Координата по осі ординат відповідає позиції символу на клавіатурі/ASCII-коду/Unicode попереднього символу в тексті. При цьому вважається, що перший символ на клавіатурі відповідає символу пробілу.

У випадку, коли кількість символів тексту буде більше, ніж кількість символів на клавіатурі/ASCII-кодів/Unicode, то для збереження квадратної форми малюнок зверху по осі ординат доповнюється рядками, які відповідають символу пробілу.

Якщо ж кількість символів тексту менша від кількості символів на клавіатурі, то для збереження форми малюнок справа доповнюється стовпцями з символами пробілу. Хоча доповнення до квадратної форми і збільшує кількість вхідних параметрів НММ, що призводить до збільшення обчислювальної ресурсоємності такої моделі, однак збереження квадратної форми забезпечує можливість використання апробованих типів ЗНМ. Це пояснюється тим, що більшість апробованих високоефективних типів таких НММ пристосовані до аналізу зображень виключно квадратної форми.

Ілюстрацією запропонованого підходу є рис. 3.8, на якому показано двомірне чорно-біле зображення закодованого тексту «AUTOMOBILE».

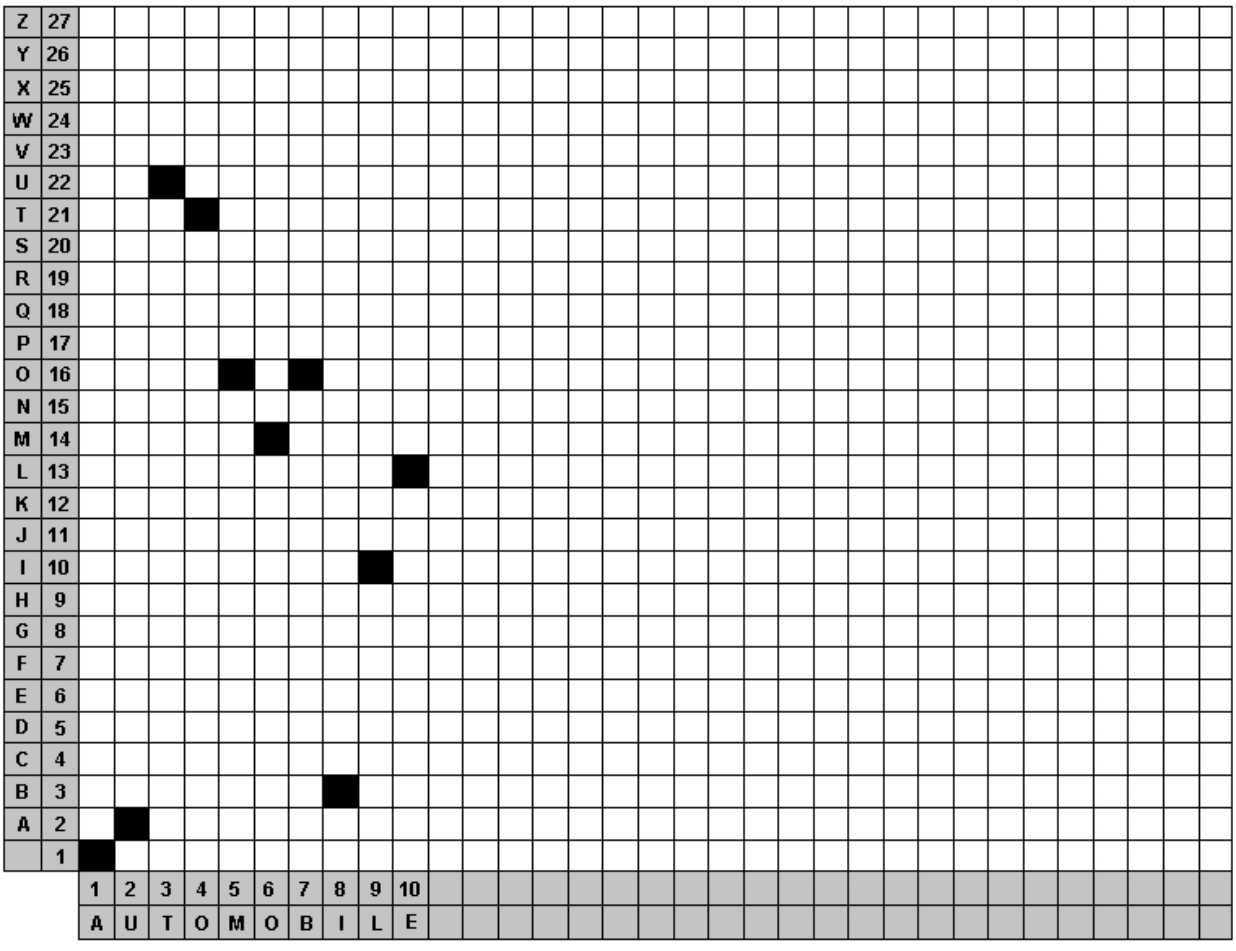

Рис. 3.8. Чорно-біле відображення закодованого слова «AUTOMOBILE»

Для спрощення демонстрації при кодуванні прийнято припущення про необхідність аналізу тексту, який складається виключно із великих літер англійського алфавіту і символу пробілу. На даному рисунку сірим кольором виділені допоміжні фрагменти, на яких по осі ординат відображаються символи/номер символу на клавіатурі, а по осі абцис номер символу/символ закодованого тексту. Для наочності окремі точки зображення відокремлюються прямими лініями. Кожна залита чорним кольором точка зображення відповідає закодованим значенням символу тексту. Наприклад, символу «T» відповідає залита чорним кольором точка зображення, що знаходиться на перетині вертикалі, проведеної з позиції №3 («T») по осі абцис, і горизонталі, проведеної з позиції №22 («U») осі ординат. Це пояснюється тим, що в слові «AUTOMOBILE» літера «T» слідує за літерою «U». У числовому вигляді даний рисунок представляє собою квадратну матрицю. Залиті чорним кольором точки рисунка відповідають елементам матриці, рівним 1, а світлі точки - елементам матриці, рівним 0. Кодування введеного тексту у вигляді чорно-білого зображення не дозволяє реалізувати розпізнавання особи та емоцій слухачів СДН, що передбачає аналіз як мінімум одного з основних параметрів КП. Тому в представленому методі формування вхідного поля передбачено подання введеного тексту у вигляді кольорового квадратного зображення з багатоканальним растром. У найпростішому випадку метод передбачає два растрових канали. Кожна точка такого зображення повинна характеризувати введений символ і один з параметрів КП, що відноситься до процесу введення цього символу. В першому наближенні це можуть бути два базових параметри КП – ТУК або ТМК.

Проведені дослідження виявили деякі труднощі використання вказаних параметрів, що пояснюється необхідністю їх реєстрації в універсальних комп’ютерних системах з точністю до однієї мілісекунди. Тому, відповідно до рекомендацій [73, 210], розроблена Windows-орієнтована програма KeyboardProgram, що дозволяє реєструвати значення ТУК та ТМК з помилкою, яка

дорівнює тривалості 50 тактів процесора. При частоті процесора $2{,}4\times10^{6}$ Гц помилка реєстрації ТУК та ТМК складає $1{,}2\times10^{-4}$ с, що достатньо для засобів аналізу КП [131, 208, 300]. Зазначимо, що вказана частота процесора відповідає розповсюдженому комп'ютерному забезпеченню СДН [38, 277]. Для прикладу, на рис. 3.9 та рис. 3.10 показано гістограми значень вказаних параметрів, отримані при введенні тексту «HEY THERE».

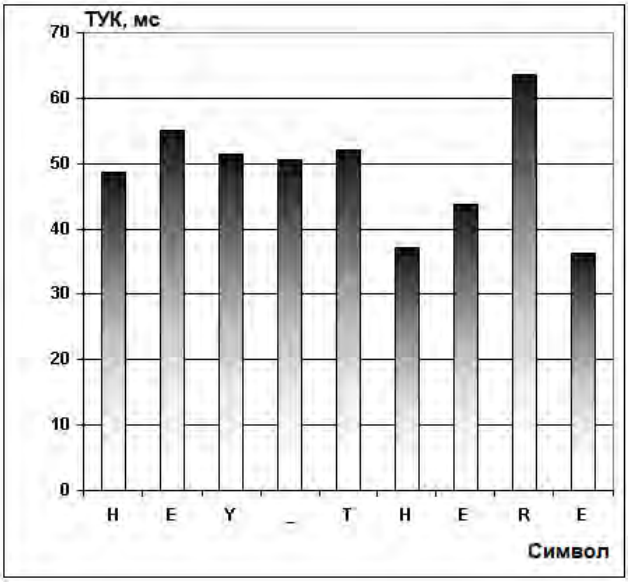

Рис. 3.9. Гістограма значень параметру ТУК для тексту «HEY THERE»

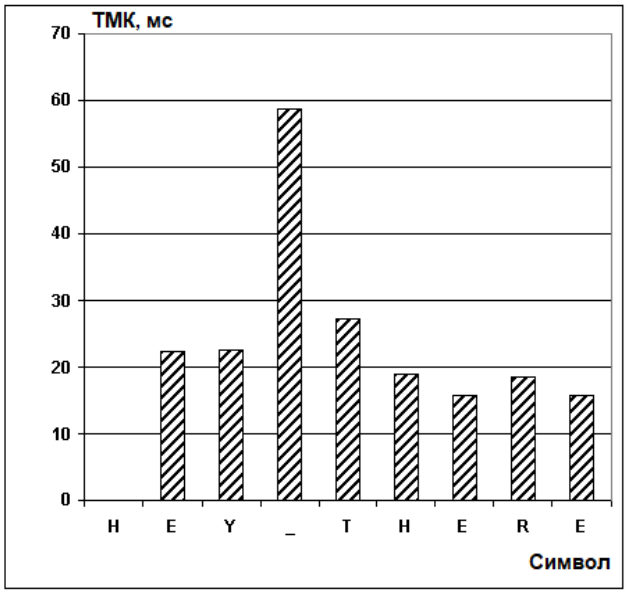

Рис. 3.10. Гістограма значень параметру ТМК для тексту «HEY THERE»

В якості ілюстрації результатів описаного підходу до кодування параметрів КП у вигляді зображення з двоканальним растром на рис. 3.11 фрагментарно представлено відображення тексту «HEY THERE», закодованого з використанням параметру ТУК.

Z	27												
Y	26				32;51								
X	25												
W	24												
V	23												
U	22												
T	21						104;37						
S	20												
R	19									101;36			
Q	18												
P	17												
O	16												
N	15												
M	14												
L	13												
K	12												
J	11												
I	10												
H	9		101;55					101;44					
G	8												
F	7												
E	6			121;51					114;64				
D	5												
C	4												
B	3												
A	2												
	1	104;49				116;52							
		1	2	3	4	5	6	7	8	9	10	11	12
		H	E	Y		T	H	E	R	E			

Рис. 3.11. Відображення тексту «HEY THERE», закодованого у вигляді зображення з двоканальним растром

На відміну від рис. 3.8, кожна точка зображення рис. 3.11, що відповідає закодованому значенню символу тексту, характеризується двома цифрами, записаними у відповідній клітинці. Перша цифра – це ASCII-код введеного символу, а друга цифра – ТУК. Наприклад, символу «Y» відповідає зображення, що знаходиться на перетині вертикалі, проведеної з позиції №3 («Y») по осі абцис, і

горизонталі, проведеної з позиції №6 («Е») осі ординат. Це пояснюється тим, що в тексті «HEY THERE» літера «Y» слідує за літерою «Е». У відповідній клітинці представлені значення 121 (ASCII-код) і 51 (ТУК). Ілюстрацією кодування параметрів КП у вигляді квадратного зображення з трьома кольоровими каналами є рис. 3.12.

я	34						
ю	33						
ь	32						
щ	31						
ш	30						
ч	29						
ц	28						
х	27						
ф	26						
у	25						
т	24						
с	23						
р	22			232;91;25			
п	21		240;93;32				
о	20						
н	19						
м	18						
л	17						
к	16						
й	15						
ї	14						
і	13						242;104;30
и	12				226;274;35		
з	11						
ж	10						
є	9						
е	8						
д	7						
г	6						
г	5						
в	4					179;106;27	
б	3						
а	2						
	1	239;105;54					
		1	2	3	4	5	6
		п	р	и	в	і	т

Рис. 3.12. Відображення тексту «привіт», закодованого у вигляді зображення з трьохканальним растром

Гістограми значень ТУК та ТКМ для цього випадку показано на рис. 3.13 та рис. 3.14.

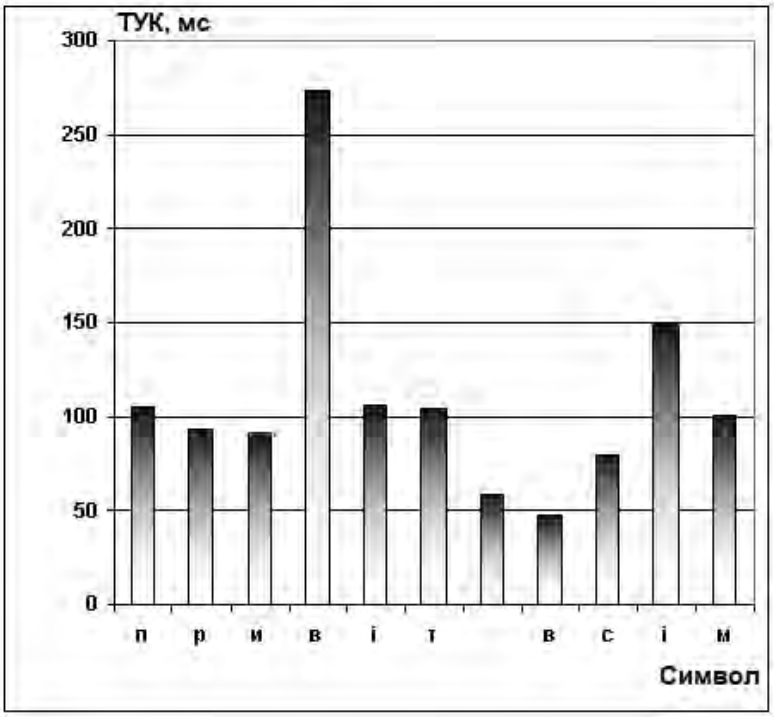

Рис. 3.13. Гістограма значень параметру ТУК для тексту «привіт всім»

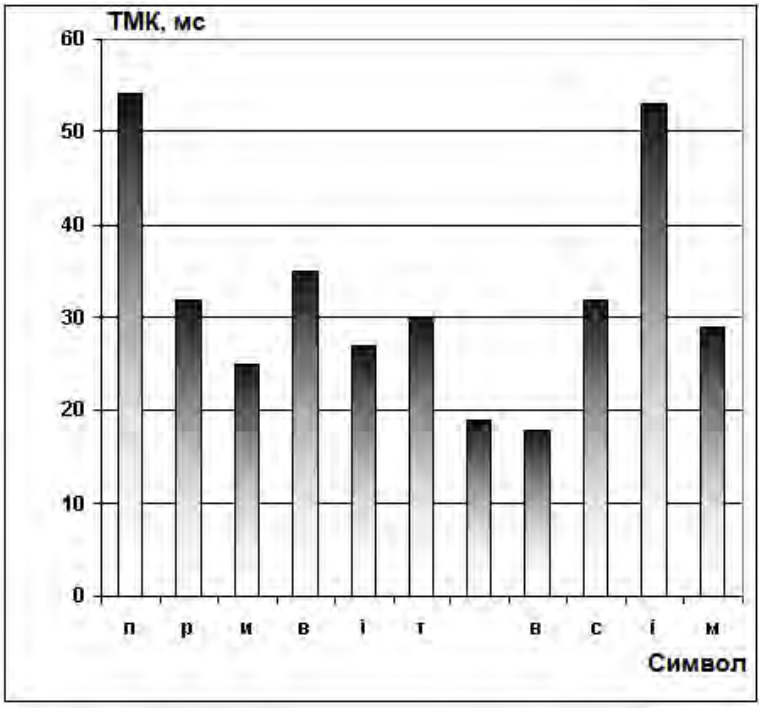

Рис. 3.14. Гістограма значень параметру ТМК для тексту «привіт всім»

Для спрощення демонстрації результатів кодування на рис. 3.12 прийнято припущення про необхідність аналізу тексту, який складається виключно з малих літер українського алфавіту і символу пробілу.

Кожна точка зображення, що відповідає закодованому значенню символу тексту, характеризується трьома цифрами, записаними у відповідній клітинці. Перша цифра – це ASCII-код введеного символу, друга цифра відповідає ТУК, а третя цифра – ТМК. Наприклад, символу «р» відповідає точка зображення, що знаходиться на перетині стовпчика позначеного символом «р» з рядком позначеним символом «п». У відповідній клітинці представлені значення 240 (ASCII-код), 93 (ТУК) та 32 (ТМК).

Таким чином, в цьому випадку при кодуванні параметрів КП використовується аналогія з кодуванням кольорового зображення з трьохканальним растром, що відповідає зображенню в форматі RGB. Зазначимо необхідність подальших досліджень, пов’язаних з визначенням кількості растрових каналів зображення, з яким асоціюються параметри КП.

З використанням розробленої Windows-орієнтованої програми KeyboardProgram проведені експерименти, результати яких дозволяють стверджувати, що при неперервному введенні текстових даних значення ТУК та ТМК, як правило, не перевищують 500 мс. Відповідно, значення ТУК та ТМК, що перевищують цю величину, слід відфільтрувати.

Крім того, результати досліджень виявили певну залежність значень ТУК та ТМК від типу клавіатури, які дещо відрізняються від наведених в першому розділі результатів аналізу літературних робіт, присвячених розпізнаванню КП. Літературні дані свідчать про те, що на клавіатурі з коротким ходом клавіш (клавіатура ноутбука) в середньому ТУК дорівнює 100 мс, а ТМК дорівнює 150 мс. При цьому для клавіатури з довгим ходом клавіш (стандартна клавіатура) ТУК дорівнює 150 мс, а ТМК дорівнює 200 мс. Середні величини ТУК і ТМК, отримані в процесі даного дослідження, приблизно в 1,5-2 рази менші.

Відповідно, при використанні клавіатури ноутбука в методі формування вхідного поля НММ доцільно передбачити можливість приведення параметрів ДКП до клавіатури стандартного типу.

В першому наближенні для цього можливо використати вирази:

$$\tau_r = 1{,}5\tau_{r,n}, \tag{3.59}$$

$$\tau_b = 1{,}33\tau_{b,n}, \tag{3.60}$$

де $\tau_{r,n}$, $\tau_{b,n}$ – термін утримання клавіші та термін між послідовним натисненням двох клавіш при використанні клавіатури ноутбука.

Інтеграція наведених результатів проведених досліджень, пов'язаних з даними [182, 251, 317], дозволила запропонувати метод формування вхідного поля НММ, призначеної для аналізу КП для розпізнавання емоцій та особи слухача СДН. В загальному вигляді перетворення інформації, що реалізується даним методом, можливо представити за допомогою виразу:

$$\langle\{SD\}, \{\Theta_S\}, TK, \{\tau_r\}\rangle \rightarrow \langle H, L, |X|\rangle, \tag{3.61}$$

де $\{SD\}$ – множина доступних символів, що можуть бути використані при введенні тексту;

$\{\Theta_S\}$ – множина обмежень на використання символів;

TK – тип клавіатури, що використовується для вводу тексту, котрий підлягає аналізу;

$\{\tau_r\}$ – параметри ДКП, що підлягають реєстрації;

H, L – висота та ширина вхідного поля ЗНМ;

$|X|$ – матриця значень, що відповідають вхідному полю ЗНМ.

Реалізація запропонованого методу складається з наступних етапів.

Етап 1. Визначення параметрів символьних даних. Вхідними даними етапу є множина доступних символів ($\{SD\}$), що можуть бути використані при введенні тексту, та множина обмежень на використання символів ($\{\Theta_S\}$). Кроки етапу 1:

Крок 1.1. Визначення множини допустимих символів ($\{S_u\}$), що можуть

входити до тексту, при введенні якого буде проводитись реєстрація первинних параметрів КП – введеного символу (s), часу натиснення клавіші (t_d) та часу відпускання клавіші (t_u).

Крок 1.2. Розрахунок кількості допустимих символів (C_S), що можуть входити до тексту, при введенні якого буде проводитись реєстрація первинних параметрів КП.

Крок 1.3. Визначення максимального обсягу символів (V_T), що можуть входити до тексту.

Крок 1.4. Упорядкування множини допустимих символів $\{S_u\}$ за певним критерієм. Надалі цю упорядковану множину будемо позначати як $\{S\}$. В базовому варіанті критерієм упорядкування є номер символу в алфавіті, визначеному $\{\Theta_S\}$.

Виходом етапу є значення C_S, V_T та множина $\{S\}$.

Етап 2. Визначення параметрів вхідного поля. Входом даного етапу є $\{S\}$, V_T, C_S. Кроки етапу 2:

Крок 2.1. Визначення розмірності осі ординат. Для цього слід співвіднести вісь ординат вхідного поля ЗНМ з символами, що входять до упорядкованої $\{S\}$. При цьому вісь ординат, як і упорядкована $\{S\}$, починається з символу пропуску.

Крок 2.2. Розрахунок розміру вхідного поля ЗНМ. Для цього слід використати вираз:

$$\begin{cases} if\ C_S == V_T \rightarrow H, L = C_S \\ C_S > V_T \rightarrow H, L = C_S \\ C_S < V_T \rightarrow H, L = V_T \end{cases}, \tag{3.62}$$

де H – висота вхідного поля;

L – ширина вхідного поля.

Крок 2.3. Розрахунок кількості доповнюючих рядків:

$$A_r = L - V_T, \tag{3.63}$$

A_r – кількість рядків, заповнених символом пропуску, що доповнюють вхідне поля зверху.

Крок 2.4. Розрахунок кількості доповнюючих стовбців:

$$A_c = H - C_s, \tag{3.64}$$

де A_c – кількість стовбців, заповнених символом пропуску, що доповнюють вхідне поля справа.

Виходом етапу є значення A_r, A_c, H, L.

Етап 3. Розрахунок параметрів растрів. На вхід етапу подаються множини $\{S\}$, $\{\Theta_S\}$ та $\{\tau_r\}$. Кроки етапу 3:

Крок 3.1. Визначення відповідності значень першого растру зображення, асоційованого з вхідним полем ЗНМ, закодованому значенню введеного символу:

$$R_1(i) = {}^{i}/_{C_s}\,, \tag{3.65}$$

де R_1 – перший растр;

$R_1(i)$ – значення і-го закодованого символу, що входить до $\{S\}$;

i – номер символу в упорядкованій $\{S\}$.

Крок 3.2. Визначення множини значень растрів. На даному кроці спочатку визначається перелік параметрів динаміки КП, що будуть використані для нейромережевого аналізу ($\{DKP\}$). Для цього реалізується співставлення $\{\tau_r\}$ та $\{\Theta_S\}$. Надалі розраховується кількісті растрів (K_R), що характеризують вхідне поле ЗНМ та співвіднесенні растрів з параметрами КП. Оскільки перший растр співставлено з ASCII-кодами символів, то співвіднесення починається з другого растру та визначається виразом:

$$R_{j+1} \to DKP_j \; j = 2,3, \dots K_R, \tag{3.66}$$

де R_{j+1} – растр з номером j+1.

Виходом етапу є K_R та кортеж $\langle \{R\}_J, \{DKP\}_J \rangle$.

Етап 4. Розрахунок параметрів клавіатурного почерку. На вхід етапу подаються тип клавіатури та значення параметрів КП, отримані при вводі тексту, що підлягає аналізу. Кроки етапу 4:

Крок 4.1. Визначення параметрів приведення початкових даних. Процедура

приведення в основному залежить від типу клавіатури, що використовується для введення тексту. В базовому варіанті це може бути стандартна клавіатура з довгим ходом (TK_l) чи клавіатура ноутбука з коротким ходом (TK_s).

Крок 4.2. Отримання початкових даних для розрахунку параметрів КП.

Крок 4.3. Приведення початкових даних. У випадку використання ноутбука застосування виразів (3.59, 3.60) для приведення значень параметрів ДКП до значень, що відповідають використанню клавіатурі стандартного типу.

Крок 4.4. Розрахунок ТУК та ТМК, що реалізується на основі приведених параметрів значень ДКП. Для цього слід використати вирази (3.54, 3.55).

Крок 4.5. Фільтрація ТУК та ТМК, що реалізується за допомогою виразів:

$$if\ \tau_{r,init} < t_{r,max} \rightarrow \tau_r = \tau_{r,init}\ else\ \tau_r = t_{r,max}, \quad (3.67)$$

$$if\ \tau_{b,init} < t_{b,max} \rightarrow \tau_b = \tau_{b,init}\ else\ \tau_b = t_{b,max}, \quad (3.68)$$

де $\tau_{r,init}$ – початкове невідфільтроване значення терміну утримання клавіші;

$\tau_{b,init}$- початкове невідфільтроване значення терміну між послідовним натисненням двох клавіш;

$t_{r,max}$ – граничне значення терміну утримання клавіші;

$t_{b,max}$ – граничне значення терміну між послідовним натисненням двох клавіш.

На основі експериментальних досліджень в першому наближенні визначено, що

$$t_{r,max} = t_{b,max} = 500 \text{ мс}. \quad (3.69)$$

Крок 4.6. Розрахунок ДТУК, ДТМК та ВТ. Для цього слід використати вирази (3.56-3.58) та відфільтровані значення ТУК та ТМК.

Виходом етапу є множина упорядкованих, відфільтрованих та приведених параметрів КП ($\{r\}$). В базовому випадку до складу цієї множини включені значення ASCII-кодів та відфільтровані та приведені значення ТУК, ТМК, ДТУК, ДТМК, ВТ, які були отримані при вводі проаналізованого тексту.

Етап 5. Формування вхідного поля ЗНМ. Входом етапу є K_R, кортеж $\langle\{R\}_J, \{DKP\}_J\rangle$ та множина $\{r\}$.

Для заповнення можливо використовувати вирази:

$$if\ j == 1 \rightarrow (r_1, r_2, \dots r_R), h = 1, l = 1, \tag{3.70}$$

$$if\ j > 1 \rightarrow (r_1, r_2, \dots r_R), h = h(s_{j-1}), l = j, \tag{3.71}$$

де j – номер введеного символу;

r_k – значення k-го параметру клавіатурного почерку, що відповідає k-му растру;

h - координата комірки вхідного поля згорткової нейронної мережі по осі ординат;

l - координата комірки вхідного поля згорткової нейронної мережі по осі абцис;

$h(s_j)$ – координата по осі ординат j-го символу, що входить до упорядкованої $\{S\}$.

Виходом п’ятого етапу є сформоване на основі параметрів КП вхідне поле ЗНМ, що може бути представлене за допомогою виразу:

$$\|X\| = \left\| \begin{matrix} \{x_{1,1}\} & \dots & \{x_{L,1}\} \\ \dots & \dots & \dots \\ \{x_{1,H}\} & \dots & \{x_{L,H}\} \end{matrix} \right\|, \tag{3.72}$$

де L – ширина вхідного поля ЗНМ;

H – висота вхідного поля ЗНМ;

$\{x_{i,j}\}$ – значення вхідного поля ЗНМ в комірці з координатами i, j.

Зазначимо, що перше значення вхідного поля ЗНМ в певній комірці відповідає ASCII-коду використаної клавіші, а інші значення – відфільтрованим та приведеним значенням параметрів динаміки КП.

Узагальнена схема запропонованого методу наведена на рис. 3.15. Зазначимо, що його математичне забезпечення складають вирази (3.54-3.72).

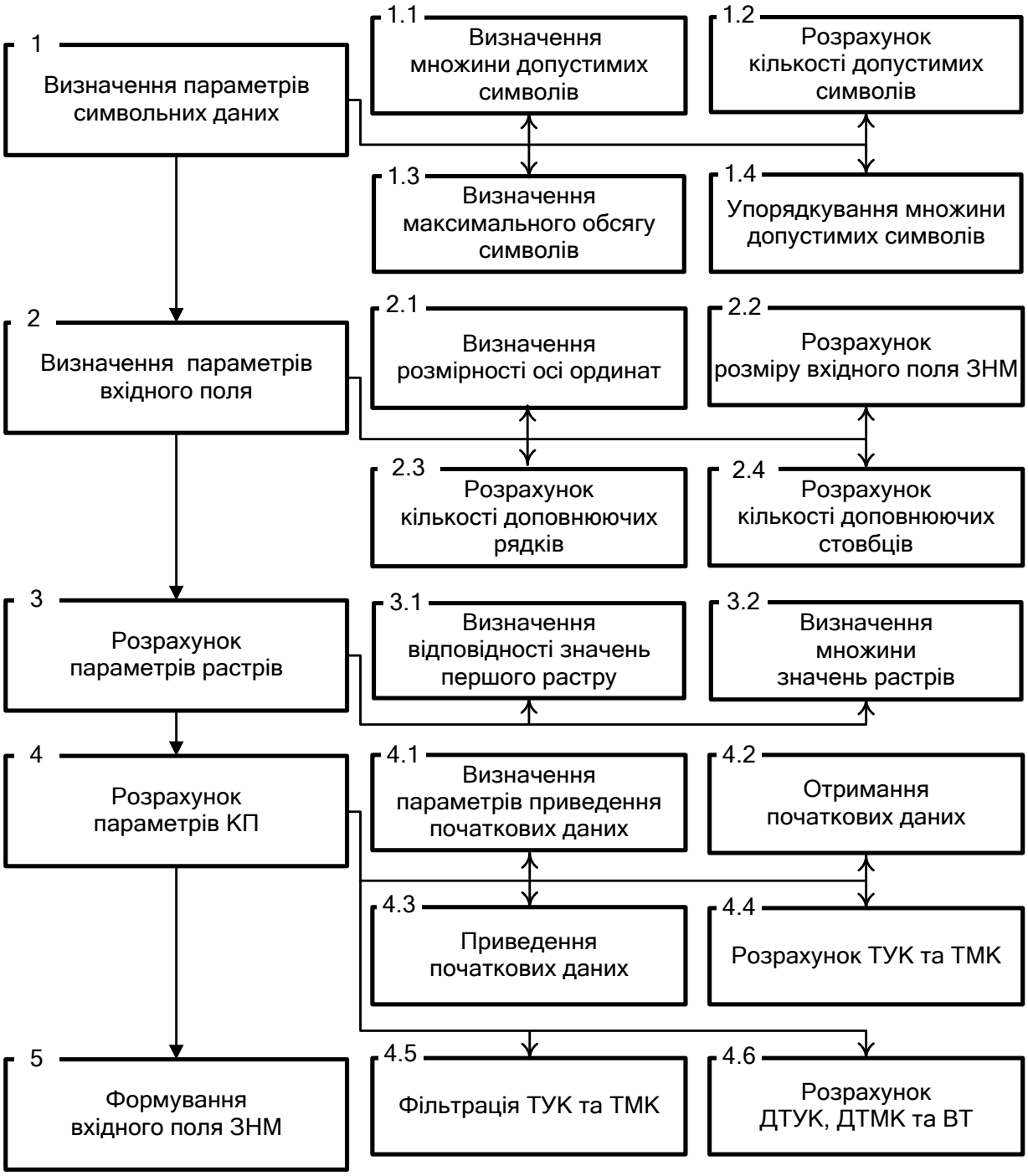

Рис. 3.15. Схема методу формування вхідного поля НММ аналізу КП

Особливістю розробленого методу є можливість паралельного виконання розрахунків. На другому етапі можливо паралельне виконання кроків 2.3 та 2.4, що співвідносяться з розрахунком кількості доповнюючих рядків та стовбців. На кроці

3.2 можливо розпаралелити розрахунки параметрів кожного із растрів. На четвертому етапі є можливість розпаралелювання розрахунків ТУК та ТМК (крок 4.4) та розпаралелювання розрахунків при фільтрації значень ТУК та ТМК (крок 4.5). Наукова новизна методу полягає у можливості подання параметрів КП у вигляді багатоканального кольорового рисунку, придатного для аналізу за допомогою ЗНМ, що забезпечує можливість підвищення ефективності розпізнавання емоцій та особи слухача СДН.

3.4. Метод формування вхідного поля нейромережевої моделі аналізу голосового сигналу

Відповідно до результатів першого та другого розділів даної дисертаційної роботи, визначення емоцій та особи слухача СДН доцільно реалізовувати на основі нейромережевого аналізу ГС фіксованої тривалості. В цьому випадку вираз, що описує перетворення інформації в методі формування вхідного поля НММ аналізу ГС, можливо записати так:

$$\langle \boldsymbol{U}_v, \langle V_{reg} \rangle, \boldsymbol{U}_{NN}, \boldsymbol{D}_V \rangle \rightarrow \boldsymbol{\Psi}^V, \tag{3.73}$$

де $\boldsymbol{U}_v$ – множина вимог до реєстрації та обробки ГС;

$\langle V_{reg} \rangle$ – кортеж параметрів зареєстрованого дискретного ГС;

$\boldsymbol{U}_{NN}$ – множина вимог до параметрів вхідного поля НММ;

$\boldsymbol{D}_V$ – множина параметрів клієнтського програмного-апаратного забезпечення;

$\boldsymbol{\Psi}^V$ – множина, що описує параметри вхідного поля НММ, яка призначена для аналізу ГС.

В свою чергу множина $\boldsymbol{U}_v$ складається із множин вимог до реєстрації ГС ($\boldsymbol{U}_{vr}$) та множини вимог до обробки ГС ($\boldsymbol{U}_{vp}$).

Враховуючи дані [86, 98, 164], з позицій можливостей використання слухачем

СДН поширеного апаратного забезпечення, визначено, що

$$\boldsymbol{U}_{vr} = \left\{f_{reg}^{min}, f_{reg}^{max}, d_{reg}^{max}, d_{reg}^{min}, \boldsymbol{tp}_V, l_{min}, l_{max}, \alpha, R_V^{min}, R_V^{max}, R_{ns}, q, \tau_{rv}, \tau_{min}\right\}, \quad (3.74)$$

де $f_{reg}^{min}, f_{reg}^{max}$ – мінімально допустима та максимально допустима частота дискретизації ГС;

$d_{reg}^{max}, d_{reg}^{min}$ – максимально допустима та мінімально допустима глибина дискретизації ГС;

$\boldsymbol{tp}_V$ – множина допустимих типів реєстрації ГС;

l_{min}, l_{max} – мінімально допустима та максимально допустима відстань від слухача до мікрофону;

α – допустима величина кута відхилення осі мікрофона;

R_V^{min}, R_V^{max} – мінімально допустимий та максимально допустимий рівень ГС, що підлягає реєстрації;

R_{ns} – максимально допустимий рівень шуму при реєстрації ГС;

q – відсутність багатоголосся при реєстрації ГС;

τ_{rv} – час реверберації;

τ_{min} – мінімально допустима тривалість реєстрації ГС.

При цьому

$$\boldsymbol{tp}_V = \{c_m, c_s, c_q\}, \quad (3.75)$$

де c_m – моноканальний ГС;

c_s – стереоканальний ГС;

c_q – квадроканальний ГС.

При виконанні вказаних вимог кортеж параметрів зареєстрованого дискретного ГС можливо представити у вигляді:

$$\langle V_{reg}\rangle = \langle f_{reg}, d_{reg}, tp_{V,reg}, \{x_{reg}\}, \{t_{reg}\}\rangle, \quad (3.76)$$

де f_{reg} – частота дискретизації ГС;

d_{reg} – глибина дискретизації ГС;

$tp_{V,reg}$ – тип ГС;

$\{x_{reg}\}$ – множина зареєстрованих параметрів ГС;

$\{t_{reg}\}$ – множина, що містить значення моментів реєстрації ГС.

Множину вимог до обробки ГС сформовано з врахуванням специфіки розрахунку вхідних параметрів НММ, призначеної для розпізнавання емоцій та особи слухача СДН:

$$\boldsymbol{U}_{vp} = \{tp_{V,d}, d_a, f_a, R_a, \tau_{em}, \tau_{qsf}\}, \quad (3.77)$$

де $tp_{V,d}$ – допустимий для подальшої обробки тип реєстрації ГС;

d_a, f_a – допустимі для подальшої обробки глибина та частота дискретизації ГС відповідно;

R_a – допустимий після обробки рівень шуму в ГС;

τ_{em} – тривалість ГС, що підлягає аналізу для розпізнавання емоцій;

τ_{qsf} – тривалість квазістаціонарного фрагменту.

З урахуванням вимог до реєстрації та обробки ГС множину параметрів клієнтського програмного-апаратного забезпечення сформовано у вигляді:

$$\boldsymbol{D}_V = \{f', d', tp'_V\}, \quad (3.78)$$

де f' - частота дискретизації аудіосистеми слухача;

d' - глибина дискретизації аудіосистеми слухача;

tp'_V - тип реєстрації ГС, встановлений в аудіосистемі слухача.

Множину вимог до параметрів вхідного поля НММ сформовано з позицій можливості їх застосування в апробованих НМЗ аналізу ГС:

$$\boldsymbol{U}_{NN} = \{\boldsymbol{\Omega}, tp_\Omega, \boldsymbol{tp}_{NN}, S_{NN}\}, \quad (3.79)$$

де $\boldsymbol{\Omega}$ – множина діагностичних параметрів ГС, що асоціюється з вхідними параметрами НММ;

tp_Ω – тип даних діагностичного параметру;

$\boldsymbol{tp}_{NN}$ – тип вхідного поля НММ;

S_{NN} – розмір вхідного поля НММ.

В базовому випадку передбачено, що вхідне поле НММ може бути або лінійного або прямокутного типу. Лінійний тип вхідного поля притаманний для НММ на основі БШМ, а прямокутний – для НММ на основі ЗНМ. Тобто

$$tp_{NN} = \{MLP, CNN\}. \quad (3.80)$$

Обмеження на розмір вхідного поля визначається з позицій використання в ЗНМ апробованих архітектурних рішень та/або застосування ЗНМ, які вже були попередньо навчені, що відповідає підходу до мінімізації ресурсоємності НМЗ.

В загальному випадку $\boldsymbol{\Omega}$ представляє собою один або декілька конкретних або інтегральних параметрів ГС. Наприклад, в якості $\boldsymbol{\Omega}$ може використовуватись амплітуда ГС, зареєстрованого в певний момент часу або коефіцієнти Фур'є.

В даному методі, базуючись на результатах першого розділу, в якості $\boldsymbol{\Omega}$ передбачено використовувати такі інтегральні параметри ГС, як мел-кепстральні коефіцієнти квазістаціонарного фрагменту.

З урахуванням наведених вимог (3.74-3.79), множина, що описує параметри вхідного поля НММ для аналізу ГС, може бути представлена за допомогою виразів:

$$\boldsymbol{\Psi}^{V}_{MLP} = |\psi_1, \psi_2, \dots, \psi_N|, \quad (3.81)$$

$$\boldsymbol{\Psi}^{V}_{CNN} = \begin{vmatrix} \psi_{1,1} & \dots & \psi_{1,N} \\ \dots & \dots & \dots \\ \psi_{M,1} & \dots & \psi_{M,N} \end{vmatrix}, \quad (3.82)$$

де $\boldsymbol{\Psi}_{V,MLP}$ - множина, що описує значення параметрів вхідного поля НММ на основі БШП;

$\boldsymbol{\Psi}_{V,CNN}$ - множина, що описує значення параметрів вхідного поля ЗНМ.

Враховуючи [204, 220], перетворення інформації, задане виразами (3.73-3.82), пропонується реалізувати за 7 етапів.

Етап 1. Базова оцінка значень зареєстрованих параметрів ГС. Даний етап полягає у перевірці дотримання тих вимог, перевірка яких можлива за рахунок співставлення значень зареєстрованих значень ГС з допустимими межами та не

вимагає обробки ГС. У випадку доступу СДН до зчитування параметрів апаратно-програмного забезпечення комп'ютера слухача можливо перевірити допустимість частоти дискретизації, глибини дискретизації, рівня потужності, тип реєстрації та тривалість реєстрації ГС. У випадку відсутності такого доступу тип реєстрації ГС не перевіряється, однак вважається, що він відповідає заданому.

Етап 2. Приведення параметрів ГС. На цьому етапі за необхідності реалізується приведення значень частоти та глибини дискретизації до заданих меж. Також за рахунок застосування фільтрів високих та низьких частот параметри зареєстрованого ГС приводяться в межі, притаманні для природньої мови.

Етап 3. Додаткова оцінка значень зареєстрованих параметрів ГС. Перевіряється рівень шуму при реєстрації ГС, відсутність багатоголосся та час реверберації ГС, що в базовому варіанті дозволяє оцінити акустичні характеристики приміщення в якому знаходиться слухач СДН.

Етап 4. Формування інтервалів розпізнавання. Реалізація даного етапу полягає у розділенні ГС на фіксовані фрагменти тривалістю τ_{em}. Після розділення кожен із фіксованих фрагментів розпізнавання розділяється на квазістаціонарні фрагменти тривалістю τ_{qsf}.

Етап 5. Спектральний аналіз. На даному етапі на кожному із квазістаціонарних фрагментів ГС розраховується амплітудно-частотна характеристика цього сигналу. Для цього використовується віконне перетворення Фур'є згладжених амплітуд ГС.

Етап 6. Розрахунок мел-кепстральних коефіцієнтів, що полягає у визначенні кількості мел-фільтрів, параметрів набору фільтрів, розрахунку енергії сигналу в межах кожного із фільтрів та застосуванні до результатів фільтрації дискретного косинусного перетворення. Визначення реалізується для кожного із квазістаціонарних фрагментів ГС. В даному методі, базуючись на сучасній психоакустичній моделі слуху, прийнято, що максимальна кількість мел-

кепстральних коефіцієнтів повинна дорівнювати 24. При цьому на вхід НММ для опису одного квазістаціонарного фрагменту можна використовувати меншу кількість мел-кепстральних коефіцієнтів. Мінімальна кількість мел-кепстральних коефіцієнтів для опису одного квазістаціонарного фрагменту дорівнює один.

Етап 7. Формування вхідного поля. Даний етап спрямований на співставлення отриманої множини мел-кепстральних коефіцієнтів кожного із фіксованих фрагментів розпізнавання ГС з вхідними параметрами НММ. У випадку використання НММ на базі БШП кожному мел-кепстральному коефіцієнту відповідає окремий вхідний параметр НММ. У випадку використання ЗНМ передбачена процедура подання множини мел-кепстральних коефіцієнтів у вигляді прямокутного напівтонового (у відтінках сірого) зображення.

Побудована на основі наведених етапів узагальнена схема методу показана на рис. 3.16.

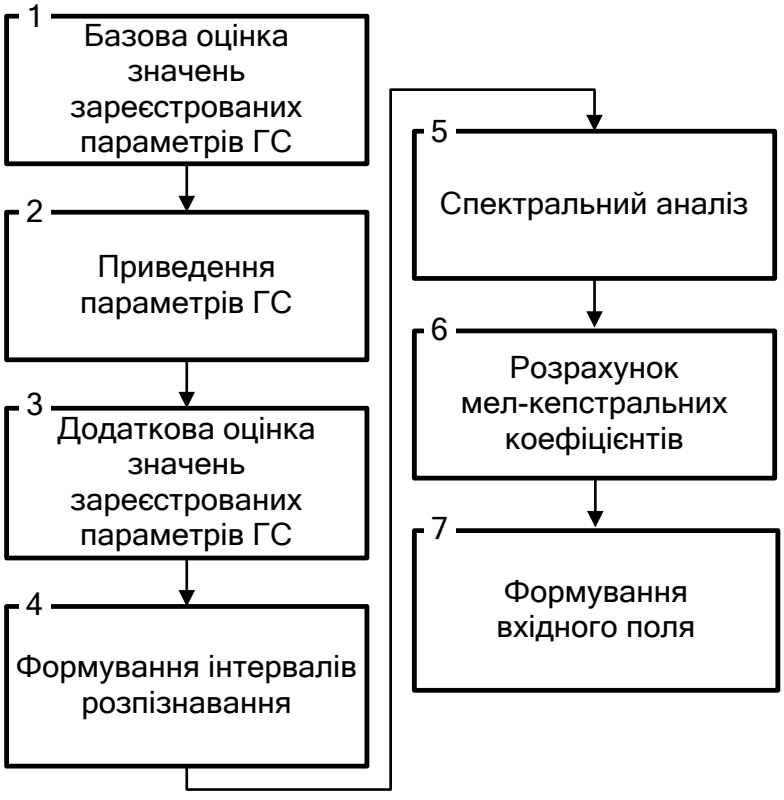

Рис. 3.16. Узагальнена схема методу формування вхідного поля НММ аналізу ГС

Запропонована процедура представлення множини мел-кепстральних коефіцієнтів у вигляді прямокутного напівтонового зображення базується на можливості співвіднесення однієї точки вхідного поля ЗНМ з одним окремим мел-кепстральним коефіцієнтом.

Кількість таких точок можливо розрахувати так:

$$Z = K_{qt} \times K_{mel}, \tag{3.83}$$

де Z – кількість мел-кепстральних коефіцієнтів;

K_{qt} – кількість квазістаціонарних фрагментів в межах піддослідного інтервалу ГС;

K_{mel} – кількість мел-кепстральних коефіцієнтів, що характеризують один квазістаціонарний фрагмент.

При відсутності перекриття квазістаціонарних фрагментів їх кількість на фрагменті ГС, що аналізується, можливо розрахувати так:

$$K_{qt} = Round\left(\frac{T}{\tau_{qsf}}\right), \tag{3.84}$$

де T – тривалість ГС;

$Round$ – функція округлення до найменшого цілого.

Зауважимо, що допустима тривалість t_{gt} варіюється в межах від 10 мс до 60 мс. Ще одне обмеження пов'язане з реалізацією швидкого перетворення Фур'є – кількість точок ГС повинна бути близькою до 2^r. Аналітично вказане обмеження можливо відобразити так:

$$\begin{cases} k_{qt} = \tau_{qsf} \times f \\ k_{qt} \cong 2^r \end{cases}, \tag{3.85}$$

де k_{qt}– кількість реєстрацій ГС на одному квазістаціонарному фрагменті;

f – частота дискретизації;

r – ціле додатне число.

Для прикладу, при використанні рекомендованої в [290, 297] кількості мел-кепстральних коефіцієнтів для одного квазістаціонарного фрагменту $K_{mel} = 10$. Враховуючи, що при частоті дискретизації ГС f = 24414 Гц, тривалості піддослідного інтервалу ГС T в межах від 1,2 с до 2 с та при тривалості квазістаціонарного фрагменту τ_{qsf} в межах від 10 мс до 60 мс, кількість мел-кепстральних коефіцієнтів, що характеризують аналізований ГС, Z = 1140. Таким чином, аналізований ГС можна співвіднести з вектором, 1140 компонент якого відповідають мел-кепстральним коефіцієнтам даного ГС.

Власне розрахунок значень мел-кепстральних коефіцієнтів можливо реалізувати за допомогою даних [84, 172]. Для ілюстрації результатів такого розрахунку на рис. 3.17, 3.18 показано графіки значень третього, четвертого та п’ятого мел-кепстральних коефіцієнтів на різних квазістаціонарних фрагментах ГС при озвучуванні текстів «Say the word back» та «Say the word bite».

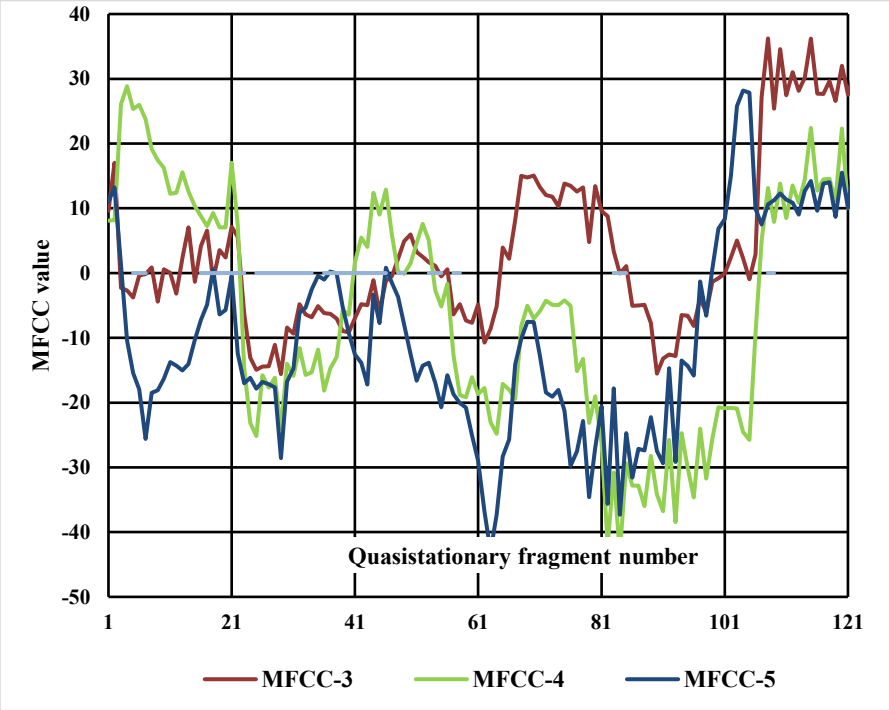

Рис. 3.17. Графіки значень мел-кепстральних коефіцієнтів при озвучуванні тексту «Say the word back»

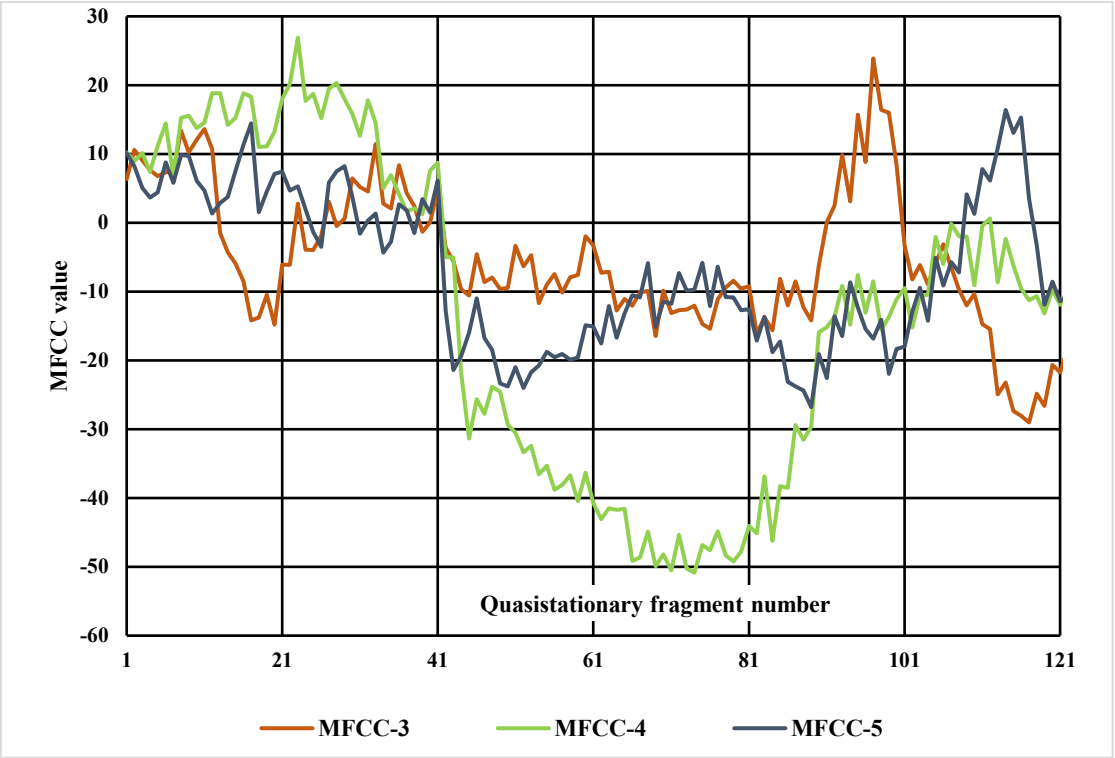

Рис. 3.18. Графіки значень мел-кепстральних коефіцієнтів при озвучуванні тексту «Say the word bite»

Розглянемо перетворення отриманого вектору мел-кепстральних коефіцієнтів в квадратну матрицю розмірністю $Z \times Z$. Вказану матрицю можливо представити у вигляді одноканального (напівтонового) квадратного зображення в декартовій системі координат.

Пропонується вимір осі ординат співвіднести з номером мел-кепстрального коефіцієнта в межах одного квазістаціонарного фрагменту, а вимір осі абцис - з номером квазістаціонарного фрагменту ГС. Колір окремої точки зображення з координатами (x_n, y_m) буде відповідати значенню m-го мел-кепстрального коефіцієнта для n-го квазістаціонарного фрагменту.

Оскільки кількість квазістаціонарних фрагментів аналізованого ГС більша, ніж кількість мел-кепстральних коефіцієнтів, то для збереження квадратної форми вхідного поля зображення, починаючи з одинадцятого, доповнюється рядками, заповненими нулями. В результаті вхідне поле ЗНМ асоціюється з матрицею

розмірністю 1140x1140, елементи якої відповідають значенням мел-кепстральних коефіцієнтів ГС. На рис. 3.19 фрагментарно показано двовимірне одноканальне зображення закодованого ГС при озвучуванні тексту «Say the word back». Рис. 3.20 відповідає закодованому ГС при озвучуванні тексту «Say the word bite».

		1	2	3	4	5
		MFCC-1	MFCC-2	MFCC-3	MFCC-4	MFCC-5
1	QF-1	-13,95	-18,62	9,62	8,09	11,08
2	QF-2	-14,13	-16,11	17,04	8,18	13,24
3	QF-3	-11,45	-22,95	-2,28	26,07	1,34
4	QF-4	-9,04	-27,60	-2,66	28,85	-10,41
5	QF-5	-7,43	-34,40	-3,78	25,34	-15,41

Рис. 3.19. Фрагментарне відображення закодованого ГС при озвучуванні тексту «Say the word back»

На рис. 3.19, 3.20 сірим кольором виділені допоміжні фрагменти. Горизонтальні фрагменти відображають номери мел-кепстральних коефіцієнтів (MFCC) і номери пікселів вхідного поля ЗНМ по осі абцис. Вертикальні допоміжні фрагменти зображення відображають номери квазістаціонарних фрагментів (QF) і номера пікселів вхідного поля ЗНМ по осі ординат.

		1	2	3	4	5
		MFCC-1	MFCC-2	MFCC-3	MFCC-4	MFCC-5
1	QF-1	15,41	15,39	6,31	10,24	10,12
2	QF-2	14,57	18,32	10,59	8,98	8,22
3	QF-3	13,24	15,83	9,05	10,13	5,06
4	QF-4	11,48	16,30	7,65	7,41	3,65
5	QF-5	13,64	19,28	6,77	11,14	4,43

Рис. 3.20. Фрагментарне відображення закодованого ГС при озвучуванні тексту «Say the word bite»

Для наочності окремі точки вхідного поля відокремлюються прямими лініями. Так, на рис. 3.19 точка вхідного поля з координатами x=2, y=3 і значенням -22,95 відповідає третьому квазістаціонарному фрагменту (QF-3), для якого значення другого мел-кепстрального коефіцієнту (MFCC-2) дорівнює -22,95.

Проведені дослідження дозволили деталізувати описані вище етапи методу формування вхідного поля HMM аналізу ГС.

Етап 1. Базова оцінка значень зареєстрованих параметрів ГС. На цьому етапі проводиться перевірка дотримання вимог реєстрації ГС. На вхід першого етапу подаються $\boldsymbol{U}_{vr}$, $\boldsymbol{D}_V$, $\langle V_{reg} \rangle$. Етап розділяється на кроки.

Крок 1.1. Валідація апаратно-програмного забезпечення. Даний крок виконується лише за умови доступу СДН до зчитування параметрів апаратно-програмного забезпечення комп'ютера слухача. Виконання даного кроку описується виразами:

$$if\ f' \in \left[f_{reg}^{min}, f_{reg}^{max}\right] \rightarrow RS_{\text{check}} = true\ else\ RS_{\text{check}} = false, \tag{3.86}$$

$$if\ d' \in \left[d_{reg}^{min}, d_{reg}^{max}\right] \rightarrow RS_{\text{check}} = true\ else\ RS_{\text{check}} = false, \tag{3.87}$$

$$if\ tp'_{\text{V}} \in \{c_m, c_s, c_q\} \rightarrow RS_{\text{check}} = true\ else\ RS_{\text{check}} = false, \tag{3.88}$$

де RS_{check} – результат перевірки.

На основі [59, 117] прийнято, що $f_{reg}^{min} = 20$ Гц, $f_{reg}^{max} = 8000$ Гц, $d_{reg}^{min} = 16$ біт, $d_{reg}^{max} = 64$ біт. У випадку відсутності доступу СДН до зчитування вказаних параметрів вважається, що $RS_{\text{check}} = false$.

Крок 1.2. Валідація результатів реєстрації, що полягає в розрахунку значень зареєстрованих параметрів ГС та порівняння отриманих значень з допустимими межами. Розрахунок значень параметрів описується виразами (3.89-3.94), а порівняння – виразами (3.95-3.98).

$$\bar{\tau} = \frac{\xi}{K-1} \sum_{k=1}^{K-1} \left(t_{x_{k+1}} - t_{x_k}\right), \tag{3.89}$$

$$\bar{f} = \frac{1}{\bar{\tau}} \tag{3.90}$$

$$\bar{d} = \frac{1}{K-1}\sum_{k=1}^{K-1}\left(d_{x_{k+1}} - d_{x_k}\right), \tag{3.91}$$

$$x^{max} = \max\{x\}_K, \tag{3.92}$$

$$x^{min} = \min\{x\}_K, \tag{3.93}$$

$$\tau_V = \xi\left(t_{x_K} - t_{x_1}\right), \tag{3.94}$$

де $\bar{\tau}$ – середній інтервал часу між двома послідовними реєстраціями ГС;

K – кількість реєстрацій, що використовуються для розрахунку середніх значень показників реєстрації ГС;

t_{x_k} – момент реєстрації k-го відліку амплітуди ГС;

ξ – коефіцієнт впливу каналу зв'язку;

$\bar{f}$ – середня частота дискретизації ГС;

$\bar{d}$ – середня глибина дискретизації ГС;

d_{x_k} – глибина дискретизації k-го відліку амплітуди ГС;

x^{max}, x^{min} – максимальне та мінімальне значення амплітуди ГС;

$\{x\}_K$ – множина зареєстрованих амплітуд ГС;

τ_V – тривалість зареєстрованого ГС.

В першому наближенні вважається, що $\xi = 1$, а K=1000.

$$if\ \bar{f} \in \left[f_{reg}^{min}, f_{reg}^{max}\right] \rightarrow RS_{\text{check}} = true\ else\ RS_{\text{check}} = false, \tag{3.95}$$

$$if\ \bar{d} \in \left[d_{reg}^{min}, d_{reg}^{max}\right] \rightarrow RS_{\text{check}} = true\ else\ RS_{\text{check}} = false, \tag{3.96}$$

$$if\ x^{max} \leq R_V^{max} \rightarrow RS_{\text{check}} = true\ else\ RS_{\text{check}} = false, \tag{3.97}$$

$$if\ x^{min} \geq R_V^{min} \rightarrow RS_{\text{check}} = true\ else\ RS_{\text{check}} = false. \tag{3.98}$$

Якщо результат перевірки негативний, тобто хоча б в одному із виразів (3.86-3.88, 3.95-3.98) $RS_{\text{check}} = false$, то вважається, що можливості апаратно-програмного забезпечення слухача незадовільні для аналізу ГС. У цьому випадку виконання методу припиняється.

Виходом етапу являється RS_{check}, значення якого свідчить про можливість подальшої реалізації методу.

Етап 2. Приведення параметрів ГС. Вхідними даними етапу являються $\langle V_{reg} \rangle$ та $\boldsymbol{U}_{vp}$. У відповідності до наведеного опису завдань етапу, його виконання розділяється на 3 кроки.

Крок 2.1. Приведення типу ГС. Відповідно до можливостей АПЗ, в першому наближенні вважається що $tp_{V,d} = c_m$. Тому на даному кроці передбачено приведення типу ГС до моноканалу. У випадку реєстрації стереоканального ГС приведення здійснюється за допомогою виразу (3.99), а у випадку квадроканального ГС – за допомогою виразу (3.100).

$$x_{reg,m} = \frac{x_{reg,1} + x_{reg,2}}{2}, \tag{3.99}$$

$$x_{reg,m} = \frac{x_{reg,1} + x_{reg,2} + x_{reg,3} + x_{reg,4}}{4}, \tag{3.100}$$

де $x_{reg,i}$ – значення амплітуди ГС, зафіксоване за допомогою і-го каналу реєстрації.

Крок 2.2. Розрахунок коефіцієнтів приведення, що описується виразами виду:

$$k_f = \frac{f_{reg}}{f_a}, \tag{3.101}$$

$$k_d = \frac{d_{reg}}{d_a}, \tag{3.102}$$

де k_f, k_d – коефіцієнти приведення частоти та глибини дискретизації ГС.

Прийнято припущення, що k_f – ціле число.

Крок 2.3. Децимація ГС. Крок виконується у випадку, коли частота дискретизації приведеного ГС повинна бути меншою ніж частота дискретизації зареєстрованого ГС, тобто при $k_f > 1$. Процедура децимації передбачає фільтрацію кожного із фрагментів ГС тривалістю τ_{em}. Фільтрація здійснюється для уникнення

явища аліасингу та полягає у видаленні високочастотних складових, що перевищують частоту Найквіста для результуючого ГС. Тобто відфільтровуються частоти, що перевищують величину $0{,}5f_a$. Після фільтрації проводиться проріджування відфільтрованого ГС.

Для фільтрації пропонується використовувати фільтр низьких частот, для якого амплітудно-частотна характеристика повинна відповідати обмеженням, що визначається виразом виду:

$$h(f_{reg}) = \begin{cases} 1, & if\ f_{reg} \le 0{,}5f_a \\ 0, & if\ f_{reg} > 0{,}5f_a \end{cases}. \tag{3.103}$$

Операція проріджування описується так:

$$\bar{x}(z) = x_{reg}\left(z \times k_f - 1\right), \quad z = 1{,}2 \dots, \left(K_{reg} - 1\right), \tag{3.104}$$

де $\bar{x}(z)$ – приведене значення амплітуди ГС для z-ої точки відліку;

K_{reg} –кількість зареєстрованих значень ГС.

Крок 2.4. Інтерполяція ГС. Крок виконується у випадку, коли частота дискретизації приведеного ГС повинна бути більшою ніж частота дискретизації зареєстрованого ГС, тобто при $k_f < 1$.

На даному кроці для кожного інтервалу ГС τ_{em} тривалістю спочатку формується проміжний сигнал виду:

$$\tau_a = \frac{1}{f_a}, \tag{3.105}$$

$$\tau_{reg} = \frac{1}{f_{reg}}, \tag{3.106}$$

$$x_{in}(z\tau_a) = \begin{cases} x_{reg}\left(z\tau_{reg}k_f\right), & z = 0, \frac{1}{k_f}, \frac{2}{k_f}, \dots K_{reg} \\ 0 & \text{для інших } z \end{cases}. \tag{3.107}$$

Після цього, отриманий проміжний сигнал оброблюється за допомогою фільтру низьких частот, для якого амплітудно-частотна характеристика повинна

відповідати обмеженням, що визначається виразом виду:

$$h(f_{in}) = \begin{cases} \frac{1}{k_f}, & if\ f_{in} \le 2k_f f_a \\ 0, & if\ f_{in} > 2k_f f_a \end{cases}. \quad (3.108)$$

Крок 2.5. Приведення глибини дискретизації. На цьому кроці значення передискретизованого ГС перезаписуються в змінні з кількістю розрядів, що дорівнює d_{reg}/k_d.

Крок 2.6. Застосування фільтра низьких частот, що зводиться до вилучення із зареєстрованого ГС частот вище f_V^{max}. Вимоги до амплітудно-частотної характеристики такого фільтру низьких частот визначається виразом виду:

$$h(f) = \begin{cases} 1, & if\ f \le f_V^{max} \\ 0, & if\ f > f_V^{max} \end{cases}. \quad (3.109)$$

Крок 2.7. Застосування фільтра високих частот, що зводиться до вилучення із зареєстрованого ГС частот нижчих ніж f_V^{min}. Вимоги до амплітудно-частотної характеристики такого фільтру високих частот визначається виразом виду:

$$h(f) = \begin{cases} 1, & if\ f \ge f_V^{min} \\ 0, & if\ f < f_V^{min}, \end{cases} \quad (3.110)$$

Виходом даного етапу є кортеж приведених параметрів реєстрації ГС, заданий виразом:

$$\langle \bar{V}_{reg} \rangle = \langle \bar{f}_{reg}, \bar{d}_{reg}, \{\bar{x}_{reg}\} \rangle, \quad (3.111)$$

де $\bar{f}_{reg}$ – частота дискретизації ГС;

$\bar{d}_{reg}$ – глибина дискретизації ГС;

$\{\bar{x}_{reg}\}$ – множина зареєстрованих параметрів ГС.

Етап 3. Додаткова оцінка значень зареєстрованих параметрів ГС. На вхід даного етапу подається $\langle \bar{V}_{reg} \rangle$, $\boldsymbol{U}_{vr}$ та $\boldsymbol{U}_{vp}$. Етап розділяється на 4 кроки.

Крок 3.1. Перевірка рівня шуму при реєстрації ГС. Рівень шуму перевіряється

на протязі перших 0,1 с запису ГС. Вважається, що на протязі перших 0,1 с слухач тільки готується для відповіді, а в цей час записується тільки фоновий шум. Виконання даної операції описується виразом:

$$if\ R_{reg} \le R_a \rightarrow RS_{\text{check}} = true \quad else\ RS_{\text{check}} = false. \tag{3.112}$$

Крім того, рівень шуму визначається на кожному із інтервалів τ_{em} за рахунок співставлення частоти основного тону ГС з допустимими межами. Вважається, що вихід частоти основного тону ГС за вказані межі свідчить про незадовільну якість такого сигналу, викликану шумом. Для пошуку частоти основного тону, відповідно до результатів [260, 267], використано вирази:

$$f_{mt} = \max_{A}\big(\mathbf{\Phi}(f, A)\big), \tag{3.113}$$

$$\mathbf{\Phi} = \big\{\big(f_{V_{min}}, A_{V_{min}}\big), \big(f_{V_{min}+1}, A_{V_{min}+1}\big), \dots \big(f_{V_{max}}, A_{V_{max}}\big)\big\}, \tag{3.114}$$

де f_{mt} – частота основного тону ГС;

$f_{V_{min}}, f_{V_{max}}$– мінімальна та максимальна частота ГС, що використовуються для пошуку f_{mt};

$A_{V_{min}+i}$ – модуль коефіцієнтів Фур'є, що відповідає частоті $f_{V_{min}+i}$.

Зазначимо, що вираз (3.112) відповідає спрощеному варіанту розрахунків частоти основного тону ГС. Однак, відповідно до результатів [269], вказане спрощення являється допустимим в задачі визначення рівня шуму, а його доцільність пояснюється необхідністю зменшення обсягу обчислень. Розрахунок виразу (3.114) передбачає попереднє визначення коефіцієнтів Фур'є для кожного із інтервалів τ_{em}. При цьому максимальна частота ГС дорівнює частоті Найквіста, тобто $f_{V_{max}} = 0{,}5 f_a$. Мінімальна частота ГС співвідноситься з нижнім порогом діапазону голосу людини. Базуючись на [23, 323], прийнято, що $f_{V_{min}} = 50$ Гц.

Перевірка допустимого рівня шуму шляхом співставлення частоти основного тону ГС з допустимими межами описується виразом:

$$if\ f_{mt} \in \left[f_{mt}^{min}, f_{mt}^{max}\right] \rightarrow RS_{\text{check}} = true \quad else\ RS_{\text{check}} = false, \tag{3.115}$$

де $f_{mt}^{min}, f_{mt}^{max}$ – нижній та верхній діапазон частот основного тону голосу людини.

На основі даних [266, 267] прийнято, що $f_{mt}^{min} = f_{V_{min}}$=50 Гц, $f_{mt}^{max} = 600$ Гц. Вказані величини прийняті з урахуванням можливостей сучасного апаратно-програмного забезпечення СДН.

Крок 3.2. Перевірка відсутності багатоголосся, що реалізується на основі визначення кількості основних тонів ГС. Для цього параметри основного тону (mt_1), визначеного на попередньому кроці, порівнюються із параметрами основного тону (mt_2), що розраховуються за допомогою виразів (3.113, 3.114) за умови вилучення із множини **Ф** елементу (f_{mt}, A_{mt}). Якщо в межах $[f_{mt}^{min}, f_{mt}^{max}]$ амплітуда для mt_1 перевищує амплітуду для mt_2 не менше ніж на 20%, то формується сигнал про наявність багатоголосся, що в свою чергу призводить до формування сигналу $RS_{\text{check}} = false$.

Крок 3.3. Перевірка часу реверберації, що реалізується на основі пошуку в ГС відлуння цього ж сигналу. На основі міжнародного стандарту ISO 3382 прийнято, що допустимий час реверберації $\tau_{rv} \in [0{,}4; 0{,}6]$ с. Пошук відлуння пропонується реалізовувати за допомогою методу, наведеного в [199], шляхом проведення спеціальних тестів до початку голосової взаємодії між слухачем та СДН. Тести полягають в трансляції аудіосистемою слухача наперед визначеного звукового сигналу з наступною реєстрацією параметрів озвучування. Якщо термін реверберації виходить за вказані межі, або тест реверберації не реалізовано, то приймається рішення, що $RS_{\text{check}} = false$.

По аналогії з першим етапом, виходом даного етапу являється сигнал RS_{check}, значення якого свідчить про можливість подальшої реалізації методу.

Етап 4. Формування інтервалів розпізнавання. Входом етапу являється $\langle \bar{V}_{reg} \rangle$ та $\boldsymbol{U}_{vp}$. Етап поділяється на три кроки.

Крок 4.1. Сегментація ГС. Для цього із зареєстрованого ГС видаляються

паузи. Зазначимо, що відомі процедури видалення пауз базуються на оцінці короткочасної енергії ГС та оцінці кількості переходів через 0. Однак розрахунок вказаних показників значно збільшує інертність системи розпізнавання, тому для видалення пауз використана спрощена процедура, котра базується на твердженні про те, що після виконання перших трьох етапів даного методу рівень приведеного ГС як мінімум на 10% вищий, ніж рівень фонового шуму. Таким чином, видалення пауз описується виразом:

$$if\ \bar{x}_{reg,i} \geq 0{,}1R_a \rightarrow \bar{x}_{reg,i} \in \boldsymbol{X}_{ss}\ , \tag{3.116}$$

де $\bar{x}_{reg,i}$ – амплітуда ГС в і-ий момент часу;

$\boldsymbol{X}_{ss}$ – множина, що містить значення амплітуд ГС з видаленими паузами.

Крок 4.2. Визначення інтервалів розпізнавання. На даному кроці із множини $\boldsymbol{X}_{ss}$ виділяються підмножини, які відповідають фрагментам розпізнавання з заданою тривалістю τ_{em}. Результат виконання даного кроку описується виразами:

$$\boldsymbol{X}_{ss} = \{\boldsymbol{x}_{ss}^1, \boldsymbol{x}_{ss}^2, \dots \boldsymbol{x}_{ss}^K\}\ , \tag{3.117}$$

$$\boldsymbol{x}_{ss}^k = \left\{x_{ss,1}^k, x_{ss,2}^k, \dots x_{ss,L}^k\right\}, \tag{3.118}$$

$$L = \tau_{em} \times f_a, \tag{3.119}$$

де $\boldsymbol{x}_{ss}^k$ – k-ий фрагмент ГС для розпізнавання емоцій та особи слухача;

L – кількість значень приведених амплітуд у фрагменті ГС.

Крок 4.3. Визначення квазістаціонарних фрагментів. Виконання кроку полягає у тому, що кожен фрагмент τ_{em} розділяється на ряд квазістаціонарних фрагментів тривалістю τ_{qsf}.

Зазначимо, що при реалізації кроків 4.2 та 4.3 враховується умова (3.85) та обмеження: $1{,}2\ \text{с} \leq \tau_{em} \leq 2\ \text{с}$; $0{,}01\ \text{с} \leq \tau_{qsf} \leq 0{,}06\ \text{с}$; , $\tau_{em}/\tau_{qsf} = z$; n, z – цілі числа. Вихід кроку та етапу в цілому є $\boldsymbol{X}_{ss}$, L і уточнені значення τ_{em} та τ_{qsf}.

Етап 5. Спектральний аналіз. Вхідними даними являються $\boldsymbol{X}_{ss}$ та f_a. Етап розділяється на два кроки.

Крок 5.1. Визначення віконної функції. В базовому варіанті передбачено

застосування вікна Хеммінга, заданого виразом (1.5).

Крок 5.2. Розрахунок амплітудно-частотної характеристики кожного із квазістаціонарних фрагментів, що входять до аналізованого інтервалу, здійснюється за допомогою віконного швидкого перетворення Фур'є. Використовується віконна функція визначена на кроці 5.1.

Результат виконання етапу для фіксованого фрагменту ГС, на якому проводиться розпізнавання, описується виразами:

$$\boldsymbol{A}_{ss} = \left\{\boldsymbol{a}_{qsf,1}, \boldsymbol{a}_{qsf,2} \dots \boldsymbol{a}_{qsf,Z}\right\}, \tag{3.120}$$

$$\boldsymbol{a}_{qsf,l} = \left\{a_{l,1}, a_{l,2}, \dots a_{l,Z}\right\}, \tag{3.121}$$

$$Z = 0{,}5 f_a, \tag{3.122}$$

де $\boldsymbol{A}_{ss}$ – множина значень модулів коефіцієнтів Фур'є для фрагменту аналізу ГС;

$\boldsymbol{a}_{qsf,l}$ – множина значень модулів коефіцієнтів Фур'є для l-го квазістаціонарного інтервалу;

$a_{l,z}$ – модуль значення коефіцієнту Фур'є для частоті z на l-му квазістаціонарному інтервалі.

Етап 6. Розрахунок мел-кепстральних коефіцієнтів. Вхідними даними цього являється множина $\boldsymbol{A}_{ss}$. Процес розрахунку реалізується відповідно до результатів [187, 323] та розділяється на 3 кроки.

Крок 6.1. Розрахунок параметрів набору 24 фільтрів, що реалізується за допомогою виразів (1.1-1.10).

Крок 6.2. Розрахунок енергії ГС, що реалізується межах кожного із фільтрів за допомогою виразу (1.11).

Крок 6.3. Реалізація дискретного косинус перетворення, що описується виразом (1.12) та застосовується до результатів, отриманих на кроці 6.2.

Вихід шостого етапу описується виразами:

$$\boldsymbol{K}_{mfcc} = \left\{\boldsymbol{k}_{mfcc,1}, \boldsymbol{k}_{mfcc,2}, \dots \boldsymbol{k}_{mfcc,Z}\right\}, \tag{3.123}$$

$$\boldsymbol{k}_{mfcc,z} = \left\{k_{mfcc,1}, k_{mfcc,2}, \dots k_{mfcc,24}\right\}, \quad (3.124)$$

де $\boldsymbol{K}_{mfcc}$ – множина значень мел-кепстральних коефіцієнтів для фрагменту аналізу ГС;

$\boldsymbol{k}_{mfcc,z}$ – множина значень мел-кепстральних коефіцієнтів для z-го квазістаціонарного інтервалу;

$k_{mfcc,i}$ – значення і-го мел-кепстрального коефіцієнту.

Етап 7. Формування вхідного поля. Входом етапу являються множини $\boldsymbol{K}_{mfcc}$ та $\boldsymbol{U}_{NN}$. Враховуючи апробовані підходи до визначення множини вхідних параметрів НММ та розроблену процедуру представлення множини мел-кепстральних коефіцієнтів у вигляді прямокутного напівтонового зображення, виконання етапу 7 розділено на 3 кроки.

Крок 7.1. Попередній розрахунок розміру вхідного поля. Для випадку $tp_{NN} = MLP$ цей розмір відповідає кількості нейронів у вхідному шарі ($N_{MLP,in}$) і розраховується за допомогою виразу (3.83). Для випадку $tp_{NN} = CNN$ висота вхідного поля (H_{CNN}) дорівнює кількості мел-кепстральних коефіцієнтів, що характеризують один квазістаціонарний фрагмент (K_{mel}). Ширина вхідного поля (H_{CNN}) дорівнює кількості квазістаціонарних фрагментів в межах піддослідного інтервалу ГС (K_{qt}). Розрахунок K_{qt} здійснюється за допомогою виразів (3.84, 3.85).

Крок 7.2. Визначення розмірності вхідного поля. Виконання даного кроку залежить від типу вхідного поля НММ, заданого виразом (3.80). Для випадку $tp_{NN} = MLP$ вхідні нейрони співвідносяться з мел-кепстральними коефіцієнтами ГС. Співвіднесення вхідного сигналу нейрону зі значенням мел-кепстрального коефіцієнту розраховується так:

$$j = 1 + Round\left(\frac{i}{K_{mel}}\right), \quad (3.125)$$

$$\psi_{n_{MLP}} = k_{mfcc,i,j},\ i = 1,2, \dots N_{MLP,in}, \quad (3.126)$$

де $\psi_{n_{MLP}}$ – вхідний сигнал нейрону з номером n_{MLP};

$k_{mfcc,i,j}$ – значення і-го мел-кепстрального коефіцієнту на j-му квазістаціонарному фрагменті приведеного ГС.

Для $tp_{NN} = CNN$ вісь ординат співвідноситься з мел-кепстральними коефіцієнтами ГС, а вісь абцис – із квазістаціонарними фрагментами ГС. Співвіднесення вхідних нейронів та мел-кепстральних коефіцієнтів описується виразом:

$$n_{cnn}(i,j) = n_{mel}(i,j), i = 1,2, \dots K_{mel},\ j = 1,2, \dots K_{mel}, \qquad (3.127)$$

де $n_{cnn}(i,j)$ – (i,j)-ий вхідний нейрон;

$n_{mel}(i,j)$ – і-ий мел-кепстральний коефіцієнт на j-му квазістаціонарному фрагменті приведеного ГС.

Крок 7.3. Приведення вхідного поля. Даний крок виконується у випадку наявності обмежень на розмір вхідного поля ЗНМ. Якщо заданий в множині $\boldsymbol{U}_{NN}$ розмір вхідного поля менший ніж $(K_{mel} \times K_{mel})$, то вхідне поле ЗНМ обмежується вказаними величинами, а значення мел-кепстральних коефіцієнтів, що виходять за вказані обмеження, не враховуються. Якщо обмеження перевищують розмір $(K_{mel} \times K_{mel})$, то вхідне поле доповнюється нулями. Кількість доповнюючих рядками та стовбців розраховується за допомогою виразів (3.62-3.64).

Виходом даного кроку, що одночасно є виходом сьомого етапу і методу в цілому є визначені виразами (3.81, 3.82) множини $\boldsymbol{\Psi}_{MLP}^{V}$ та $\boldsymbol{\Psi}_{CNN}^{V}$.

Деталізована схема 1-3 етапів методу показана на рис. 3.21, а деталізована схема етапів 4-7 – на рис. 3.22. Даний метод передбачає можливість розпаралелювання розрахунків, пов’язаних з його виконанням. Так, на етапі 2 можливо паралельно виконувати: кроки 2.3, 2.4 та крок 2.5, кроки 2.6 та 2.7, а на етапі 3 можливо розпаралелити виконання кроків 3.2 та 3.3. Зазначимо, що наукова новизна розробленого методу формування вхідного поля НММ аналізу ГС полягає у застосуванні процедур перевірки характеристик клієнтського АПЗ, валідації величин зареєстрованих параметрів, перевірки відсутності багатоголосся і часу

реверберації, приведення параметрів ГС до заданого діапазону та адаптації параметрів ГС до аналізу ЗНМ.

Рис. 3.21. Деталізована схема 1-3 етапів методу формування вхідного поля НММ аналізу ГС

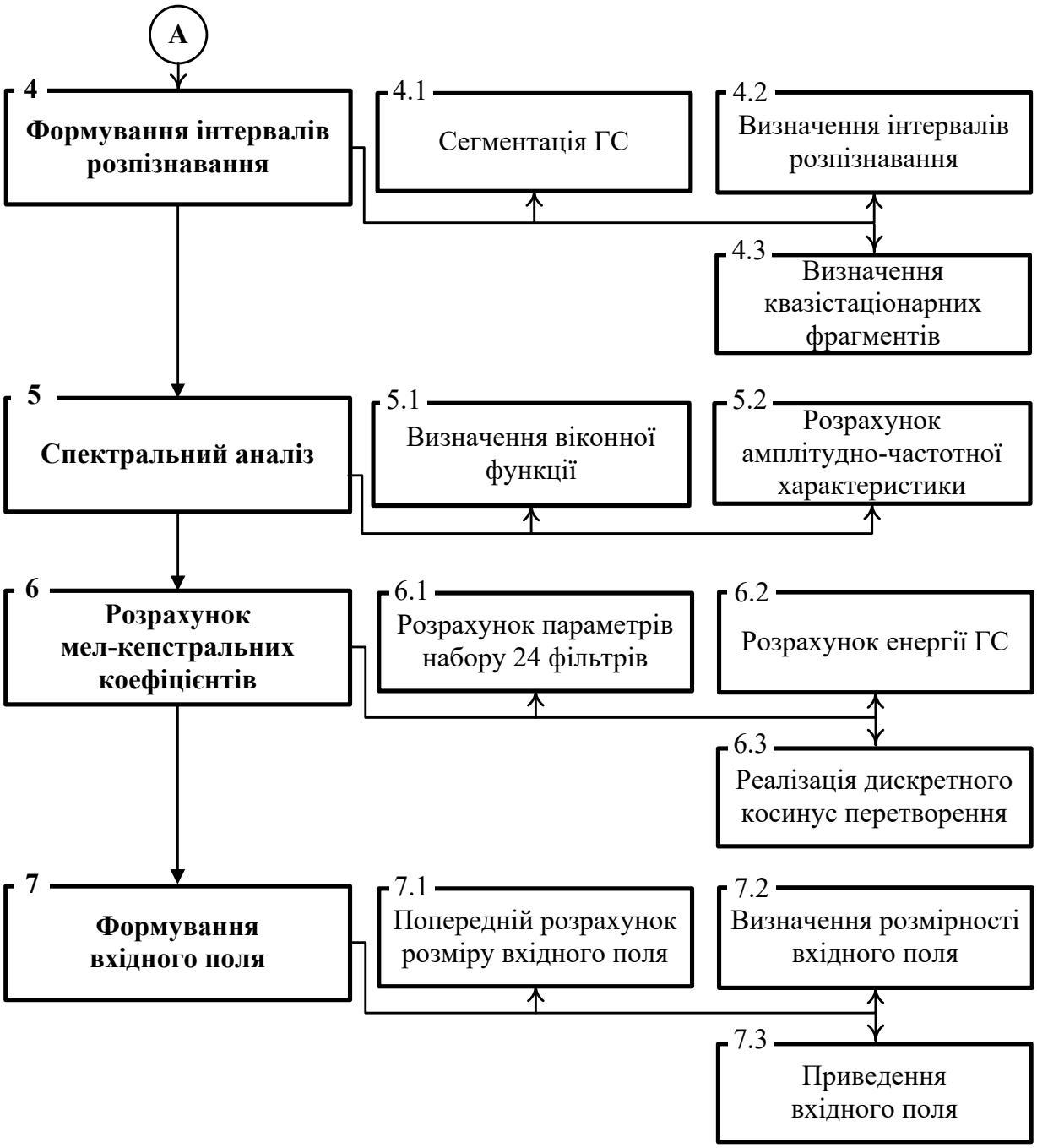

Рис. 3.22. Деталізована схема 4-7 етапів методу формування вхідного поля НММ аналізу ГС

Крім того, новизна методу полягає у можливості представлення параметрів обробленого ГС у вигляді напівтонового прямокутного рисунку, пристосованого до

аналізу за допомогою ЗНМ. Вказані особливості забезпечують можливість створення ефективної методології розпізнавання емоцій та особи слухача СДН.

Перспективи розвитку розробленого методу формування вхідного поля НММ аналізу ГС пов’язані з забезпеченням ефективного розпізнавання емоцій та особи слухача СДН в умовах багатоголосся та нерозбірливості мовлення, що характерно при проведенні дистанційних занять з групою слухачів.

3.5. Висновки до третього розділу

В даному розділі вирішувалось наукове завдання розвитку методів формування вхідного поля нейромережевих моделей розпізнавання емоцій слухачів системи дистанційного навчання. Основні результати розділу наступні:

1. Отримав подальший розвиток метод застосування вейвлет-перетворень для фільтрації вхідного поля нейромережевої моделі аналізу біометричних параметрів, асоційованих з зображеннями, в якому на відміну від відомих, за рахунок реалізації процедури визначення найбільш ефективного типу базисного вейвлету, геометрична форма якого найбільш подібна до геометрії піддослідного процесу та використання результатів вказаної процедури для розрахунку та інтеграції вейвлет-коефіцієнтів послідовно зареєстрованих зображень, забезпечується можливість фільтрації завад типових для системи дистанційного навчання.

2. Удосконалено метод формування вхідного поля нейромережевої моделі аналізу біометричних параметрів, асоційованих з зображеннями, що за рахунок адаптації процедур, пов’язаних із обробкою зображень, нівелюванням впливу типових завад та процедур визначення параметрів ключових та контрольних точок, до вимог нейромережевого аналізу зображень обличчя, вушної раковини, райдужної оболонки ока, сітківки ока та відбитку пальця, забезпечує можливість ефективного розпізнавання емоцій та особи слухача системи дистанційного навчання.

3. Вперше розроблено метод формування вхідного поля нейромережевої моделі аналізу клавіатурного почерку слухача системи дистанційного навчання, що за рахунок забезпечення можливості подання параметрів клавіатурного почерку у вигляді багатоканального кольорового рисунку придатного для аналізу за допомогою згорткової нейронної мережі забезпечує можливість ефективного розпізнавання особи та емоцій слухача системи дистанційного навчання.

4. Вперше розроблено метод формування вхідного поля нейромережевої моделі аналізу голосового сигналу, що на відміну від відомих за рахунок адаптації процедур, пов'язаних з перевіркою вимог до параметрів апаратно-програмного забезпечення, валідації величин зареєстрованих параметрів, перевірки відсутності багатоголосся та часу реверберації, приведення параметрів голосового сигналу до заданого типу запису, глибини і частоти дискретизації до умов системи дистанційного навчання та за рахунок впровадження можливості представлення параметрів обробленого голосового сигналу у вигляді напівтонового прямокутного зображення, пристосованого до аналізу за допомогою згорткової нейронної мережі забезпечує можливість ефективного розпізнавання особи та емоцій слухача системи дистанційного навчання.

5. Перспективи подальших досліджень пов'язані з підвищенням ефективності запропонованих методів за рахунок застосування в них процедур:

- Вибору принципу інтеграції вейвлет-коефіцієнтів в залежності від параметрів відеореєстрації та властивостей завад на зображенні, що асоційоване з біометричним параметром.

- Кодування параметрів кожного із каналів зображення, асоційованого з біометричним параметром за допомогою вейвлет-перетворень.

- Виділення та підвищення якості голосового сигналу слухача системи дистанційного навчання, зареєстрованого в умовах багатоголосся та нерозбірливості мовлення.

РОЗДІЛ 4
МЕТОДИ НЕЙРОМЕРЕЖЕВОГО РОЗПІЗНАВАННЯ ЕМОЦІЙ ТА ОСОБИ СЛУХАЧА СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ

4.1. Метод формування навчальних прикладів для нейромережевої моделі аналізу біометричних параметрів

Використовуючи результати [46, 237, 244, 289], прийнято, що у загальному випадку r-ий навчальний приклад НММ аналізу БП може бути записаний за допомогою виразів виду:

$$\boldsymbol{x_r} \rightarrow \boldsymbol{y_r}, \tag{4.1}$$

$$\boldsymbol{x_r} = |x_1, x_2, \dots x_K|, \tag{4.2}$$

$$\boldsymbol{x_r} = \begin{vmatrix} x_{1,1} & \dots & x_{1,L} \\ \dots & \dots & \dots \\ x_{1,H} & \dots & x_{H,L} \end{vmatrix}, \tag{4.3}$$

$$\boldsymbol{y_r} = |y_1, y_{,2}, \dots y_{N,}|, \tag{4.4}$$

де $\boldsymbol{x_r}$ – множина значень вхідних параметрів НММ для r-го навчального прикладу;

$\boldsymbol{y_r}$ – множина значень вихідних параметрів НММ для r-го навчального прикладу;

K, N – кількість вхідних та вихідних параметрів НММ;

H, L – розміри вхідного поля ЗНМ.

Вираз (4.2) використовується у випадку, коли вхідне поле НММ описується вектором, що характерно для НММ на основі БШП. Вираз (4.3) використовується для опису вхідного поля ЗНМ. При використанні НММ, які здатні самонавчатись, $\boldsymbol{y_r} = \emptyset$.

З урахуванням наведеного опису навчальних прикладів та результатів

досліджень третього розділу, перетворення інформації, що реалізується в методі формування навчальних прикладів для НММ аналізу БП, можливо записати за допомогою виразів виду:

$$\langle \boldsymbol{Q}, b, R, \boldsymbol{G}, \boldsymbol{U}, \boldsymbol{E}, \boldsymbol{S}, \boldsymbol{D} \rangle \rightarrow \langle \boldsymbol{X}, \boldsymbol{Y} \rangle, \tag{4.5}$$

$$\boldsymbol{Q} = \{\boldsymbol{q}\}_R, \tag{4.6}$$

$$\boldsymbol{X} = \{\boldsymbol{x}\}_R, \tag{4.7}$$

$$\boldsymbol{Y} = \{\boldsymbol{y}\}_R, \tag{4.8}$$

де $\boldsymbol{Q}$ – множина значень показників, що характеризують БП, який підлягає реєстрації;

$\boldsymbol{q_r}$ **-** множина значень показників, що характеризують БП для r-го навчального прикладу;

b – тип БП;

R – кількість навчальних прикладів;

$\boldsymbol{G}$ – множина показників, що характеризують БП;

$\boldsymbol{U}$ – множина вимог до вхідних параметрів НММ;

$\boldsymbol{E}$ – множина емоцій, що мають бути розпізнані;

$\boldsymbol{S}$ – множина слухачів, особи яких мають бути розпізнані;

$\boldsymbol{D}$ – множина експертних даних, що використовуються для обробки БП.

При цьому

$$b \in \{\boldsymbol{B}, \Upsilon, \Psi\}, \tag{4.9}$$

де $\boldsymbol{B}$ – множина БП, асоційованих з зображеннями;

Υ – КП слухача;

Ψ – голос слухача.

Зазначимо, що розв’язок завдання, заданого виразами (4.5-4.9), розділяється на дві частини. Перша частина стосується формування множини значень вхідних параметрів НММ, характерних для даного навчального прикладу, друга частина – формування множини значень вихідних параметрів. При цьому для кожного із БП,

заданих виразом (4.9), визначення множини значень вхідних параметрів НММ передбачено реалізовувати за допомогою методів, розроблених в третьому розділі даної дисертаційної роботи. Вирішення другої частини завдання (4.5) передбачено реалізовувати із врахуванням тієї частини запропонованого підходу до навчання НММ, що стосується доцільності врахування в навчальних прикладах близькості еталонів класів, які мають бути розпізнані.

Деталізуємо завдання відображення у вихідному сигналі навчальних прикладів близькості еталонів на прикладі нейромережевого розпізнавання друкованих літер українського алфавіту, які відображаються чорним кольором на білому фоні. Класична процедура кодування параметрів еталонів літер, проілюстрована рис. 4.1, складається в реалізації п'яти етапів.

1. Кожна літера розміщується в окремому прямокутнику фіксованого розміру, який розділений координатною сіткою на окремі комірки. Кількість комірок визначає кількість вхідних параметрів НММ. На рис. 4.1 ця кількість дорівнює:

$$K = l \times h = 11 \times 10 = 110, \tag{4.10}$$

де l, h – висота та ширина прямокутника, що описує літеру.

У межах окремого прямокутника кожній з комірок призначається номер, який відповідає номеру вхідного параметра НММ.

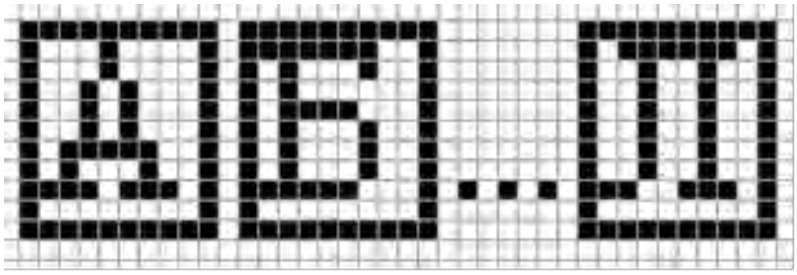

Рис. 4.1. Ілюстрація процедури кодування друкованих літер

2. Для окремої літери значення і-го вхідного параметра дорівнює 1, якщо відповідна і-та комірка заповнена чорним кольором, та дорівнює 0 у випадку, коли

комірка заповнена білим кольором.

3. Літери записуються і нумеруються в алфавітному порядку.

4. Номер літери в алфавіті визначає необроблене значення очікуваного вихідного сигналу НММ. Таким чином, для літери А необроблене значення очікуваного вихідного сигналу дорівнює 1, для літери Б – 2, а для літери Л – 16.

5. Для кожної з літер необроблене значення перетвориться до вигляду, придатного до використання в НММ. Залежно від структури НММ, можливі два варіанти перетворення.

Варіант 1. В НММ є тільки один вихідний нейрон. В цьому випадку для еталону n-ої літери очікуваний вихідний сигнал визначається так:

$$y(n) = {y_{reg}(n)}/{N} = {n}/{N}, \quad (4.11)$$

де $y_{reg}(n)$ – необроблене значення вихідного сигналу для еталону n-ої літери;

n – номер літери в алфавіті;

N – кількість літер в алфавіті.

Варіант 2. Вихідний сигнал НММ визначається набором вихідних нейронів, кількість яких дорівнює кількості літер в алфавіті. Кожен вихідний нейрон співвідноситься з однією окремою літерою. Для навчального прикладу, що відповідає даній літері, вихід цього нейрону дорівнює 1, а виходи всіх інших нейронів дорівнюють 0. Це так зване «one-hot» кодування.

В цьому випадку для навчального прикладу еталону n-ої букви очікуваний вихідний сигнал визначається так:

$$\begin{cases} y_n(n) = \dfrac{y_{reg}(n)}{n} = 1, n = 1,..N \\ y_k(n) = y_{reg}(n) - n = 0, k = 1,..N, k \neq n \end{cases}, \quad (4.12)$$

де k – номер вихідного нейрону;

$y_n(n)$ – очікуваний вихідний сигнал n-го вихідного нейрону для n-го еталону;

$y_k(n)$ – очікуваний вихідний сигнал k-го вихідного нейрону для n-го еталону.

Принциповий недолік описаної процедури кодування полягає в тому, що визначена на четвертому етапі величина необробленого очікуваного вихідного сигналу погано співвідноситься з геометричною близькістю зображень, що мають бути розпізнані.

Так, наприклад, при використанні першого варіанту кодування літер українського алфавіту очікуваний вихідний сигнал для літери А $y_{\text{А}} = {}^{1}/_{33}$, для літери Б $y_{\text{Б}} = {}^{2}/_{33}$, а для літери Л $y_{\text{Л}} = {}^{16}/_{33}$. Таким чином, значення очікуваного вихідного сигналу для літери А значно ближче до очікуваного вихідного сигналу для літери Б, ніж до сигналу для літери Л. При цьому геометрична подібність літер А та Л більша, ніж для літер А та Б.

Для усунення цього недоліку запропоновано для оцінки близькості еталонів використовувати мережу PNN, в навчальних прикладах якої очікуваний вихідний сигнал представляє собою не число, а назву класу [103]:

$$|x|_k \to N\ , \tag{4.13}$$

де $|x|_k$ – вектор вхідних параметрів для k-го навчального прикладу;

N – назва класу, до якого належить k-ий навчальний приклад.

Для прикладу, на рис. 4.2 показана структура мережі PNN, яка призначена для віднесення невідомих образів до одного з трьох відомих класів – A, B або C. Передбачається, що кожен з графічних образів розміщено в окремий прямокутник розміром $l \times h$ пікселів. Кількість входів мережі PNN дорівнює числу ознак класу і визначається за допомогою (4.7).

Показана на рис. 4.2 мережа складається з чотирьох шарів нейронів:

- Вхідний шар (Ln_{in}) – призначений виключно для отримання вхідної інформації.

Нейронам даного шару притаманна лінійна функція активації виду:

$$x = y, \tag{4.14}$$

де x – вхідний сигнал;

y – вихідний сигнал.

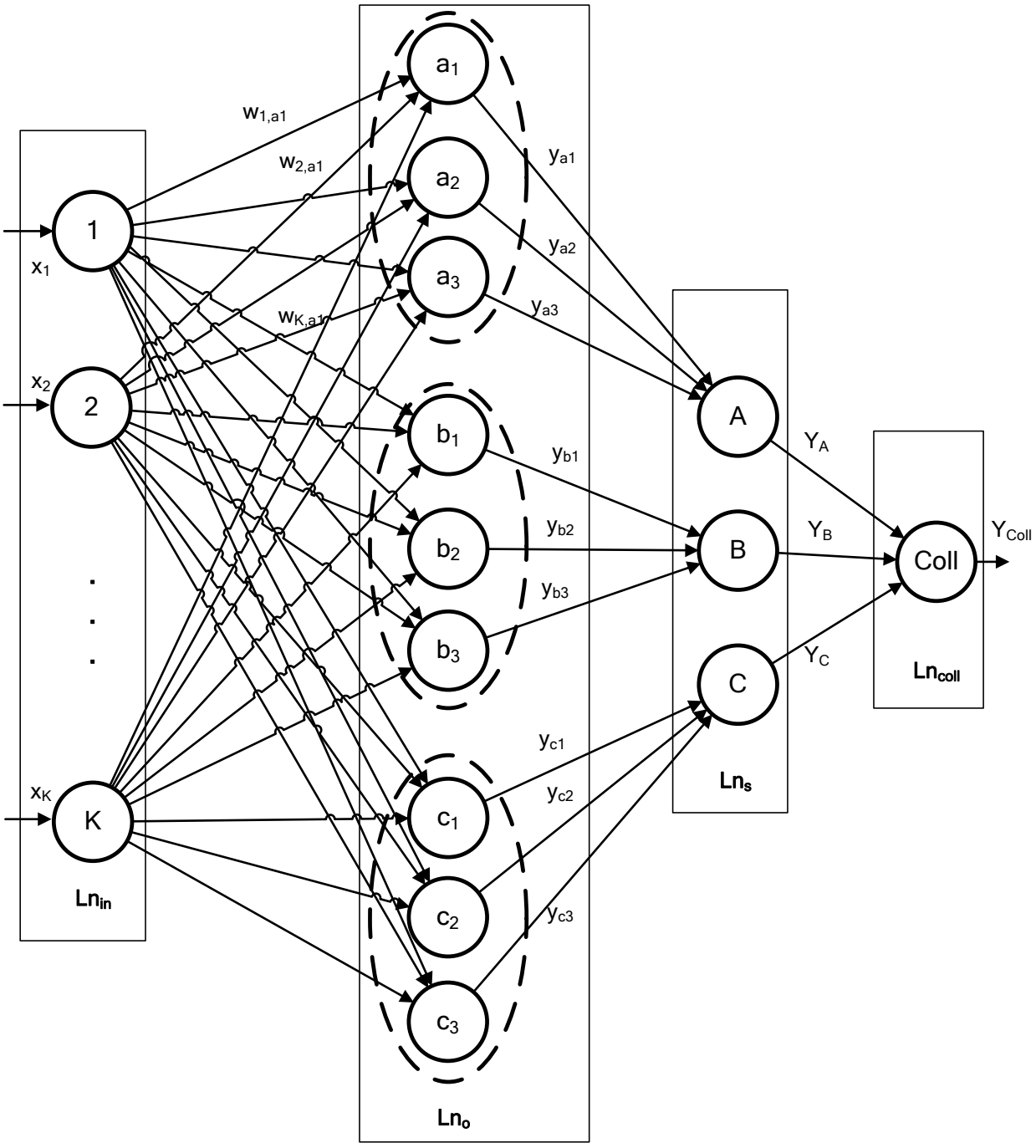

Рис. 4.2. Приклад структури PNN, призначеної для віднесення невідомого образу до одного з класів A, B або C

- Шар образів (Ln_o). Нейрони даного шару в режимі навчання призначені для зберігання навчальних прикладів. В режимі розпізнавання нейрони даного шару генерують вихідний сигнал, величина якого свідчить про схожість невідомого образу з навчальним прикладом, з яким асоційовано даний нейрон образів.

- Шар додавання (Ln_s). Нейрони, котрі входять до складу даного шару, в режимі розпізнавання інтегрують вихідні сигнали пов’язаних з ними нейронів шару образів. Нейрони шару додавання асоціюються з класами, що мають бути розпізнані.

- Вихідний шар (Ln_{coll}) – призначений для генерації результатів розпізнавання невідомого образу. В класичному вигляді результатом розпізнавання являється назва класу, до якого мережа відносить невідомий вхідний образ.

Кількість нейронів шару образів мережі PNN дорівнює кількості навчальних прикладів, а кількість нейронів шару додавання дорівнює кількості класів, що мають бути розпізнані. Нейрон шару образів пов'язаний тільки з тим нейроном шару додавання, якому відповідає клас образу.

Тому вагові коефіцієнти зв’язків між нейронами шару додавання та пов’язаним з ними нейроном шару образів дорівнюють 1, а вагові коефіцієнти зв’язків між нейронами шару додавання та непов’язаним з ними нейроном шару образів дорівнюють 0. На рис. 4.2 зв’язки між нейронами шару образів та нейронами шару додавання, для яких вагові коефіцієнти дорівнюють 0, не показані.

Вагові коефіцієнти зв'язків, що входять у вихідний елемент, дорівнюють 1, тобто

$$w_l = 1 \quad l = 1,2, \dots L\ , \tag{4.15}$$

де w_l – ваговий коефіцієнт зв’язку між l-им нейроном шару додавання та вихідним нейроном;

L – кількість нейронів шару додавання.

Навчання мережі полягає у тому, що для синаптичних зв'язків, які входять в

нейрон шару образів, вагові коефіцієнти встановлюються такими ж, як складові частини відповідного навчального прикладу, тобто

$$w_{i,j} = x_{i,j}, \tag{4.16}$$

де $w_{i,j}$ – ваговий коефіцієнт між і-им вхідним нейроном та j-им нейроном шару образів;

x_i – значення і-го вхідного параметру навчального прикладу, що співвідноситься з j-им нейроном шару образів.

Таким чином, всі параметри мережі PNN безпосередньо визначаються навчальними даними.

Вихідний сигнал довільного j-го нейрона шару образів розраховується так:

$$y_j = \sum_{k=1}^{K} exp\left(\frac{-\left(w_{k,j} - x_k\right)^2}{\sigma^2}\right), \tag{4.17}$$

де x_k – k-та компонента невідомого образу, що дорівнює вихідному сигналу k-го вхідного нейрону;

$w_{k,j}$ – ваговий коефіцієнт зв'язку між k-им вхідним та j-им нейроном шару образів;

K – кількість компонент вхідного образу, що дорівнює кількості нейронів вхідного шару;

σ – радіус функції Гауса.

Прийнято, що для більшості випадків можливо прийняти $\sigma = 0{,}5$. Разом з тим, в теоретичних роботах [42, 59] відзначено, що радіус функції Гауса не має значного впливу на результат розпізнавання.

В нейронах шару додавання використовується лінійна функція активації виду:

$$y_{c,i} = x_{c,i}^{\Sigma}, \tag{4.18}$$

де $y_{c,i}$ – вихідний сигнал і-го нейрону шару додавання;

$x_{c,i}^{\Sigma}$ – сумарний вхідний сигнал і-го нейрону шару додавання.

Вихідний сигнал n-го нейрону даного шару (Y_n) розраховується так:

$$Y_n = \frac{1}{I}\sum_{i=1}^{I} y_i, \tag{4.19}$$

де I – кількість нейронів шару образів, пов'язаних з n-им нейроном шару додавання;

y_i – вихідний сигнал і-го нейрона шару образів, пов'язаного з n-им нейроном шару додавання.

Величина Y_n представляє собою результат віднесення вхідного образу до класу, який відповідає даному нейрону. Завданням вихідного елемента є тільки визначення нейрона шару додавання з максимальною активністю. Хоча результатом розпізнавання мережі PNN є тільки назва класу, до якого віднесено вхідний приклад, але величини Y_n вказують на приналежність невідомого вхідного прикладу до одного з класів.

Запропонована процедура використання мережі PNN для визначення очікуваного вихідного сигналу полягає в наступному:

- За допомогою виразу (4.13) формується множина навчальних прикладів, які відповідають множині еталонів класів, що мають бути розпізнані.

- Реалізується навчання мережі.

- На вхід навченої мережі послідовно подаються приклади класів. Для кожного прикладу за допомогою виразів (4.14-4.19) розраховуються величини вихідних сигналів нейронів шару додавання. Масштабовані величини цих сигналів і будуть очікуваним вихідним сигналом для класичних ГНМ, у яких вихідний сигнал генерується у числовому вигляді.

Розглянемо застосування розробленої процедури використання мережі PNN для визначення очікуваного вихідного сигналу на прикладі розпізнавання 8 чорно-білих фігур, показаних на рис. 4.3. В спрощеному вигляді вказані фігури можливо

асоціювати з БП, що співвідносяться з монохромними зображеннями.

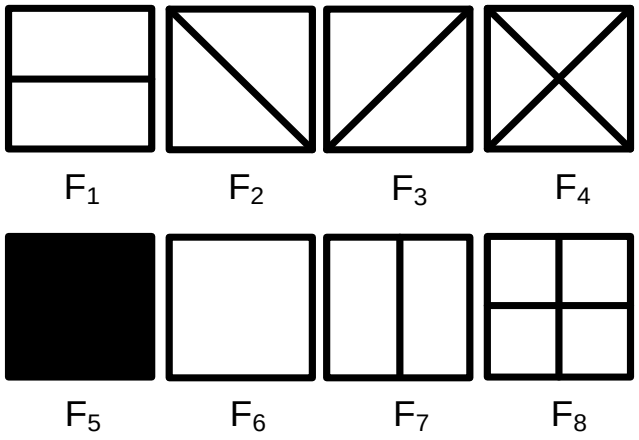

Рис. 4.3. Фігури, що співвідносяться з класами, які мають бути розпізнані

Показані на рис. 4.3 фігури записані в квадрат розміром 3х3. Таким чином, кількість вхідних параметрів НММ дорівнює 9. Кожна з фігур відповідає еталону класу, що має бути розпізнаний.

Відповідно запропонованого підходу до застосування мережі PNN, параметри навчальних прикладів визначені за допомогою виразу (4.13). Наприклад, еталони фігур F_1 та F_5 записано у вигляді:

$$|0,0,0,1,1,1,0,0,0| \rightarrow F_1, \tag{4.20}$$

$$|1,1,1,1,1,1,1,1,1| \rightarrow F_5. \tag{4.21}$$

Параметри навчальних прикладів, що відповідають фігурам наведеним на рис. 4.3 зведені в табл. 4.1.

Таблиця 4.1

Параметри навчальних прикладів

Номер прикладу	**Вхідні параметри**									**Вихідний параметр**
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
1	0	0	0	1	1	1	0	0	0	F_1
2	1	0	0	0	1	0	0	0	1	F_2

Таблиця 4.1 (продовження)

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
3	0	0	1	0	1	0	1	0	0	F_3
4	1	0	1	0	1	0	1	0	1	F_4
5	1	1	1	1	1	1	1	1	1	F_5
6	0	0	0	0	0	0	0	0	0	F_6
7	0	1	0	0	1	0	0	1	0	F_7
8	0	1	0	1	1	1	0	1	0	F_8

Вагові коефіцієнти зв'язків мережі PNN, що відповідають навчальним прикладам, котрі представлені в табл. 4.1, наведені в табл. 4.2.

Таблиця 4.2

Величини вагових коефіцієнтів синаптичних зв'язків

Ln_o	Ln_{in}									Ln_s							
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8
$f1_1$	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0
$f2_1$	1	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0
$f3_1$	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	0
$f4_1$	1	0	1	0	1	0	1	0	1	0	0	0	1	0	0	0	0
$f5_1$	1	1	1	1	1	1	1	1	1	0	0	0	0	1	0	0	0
$f6_1$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
$f7_1$	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0
$f8_1$	0	1	0	1	1	1	0	1	0	0	0	0	0	0	0	0	1

Після цього з використанням даних табл. 4.1 та табл. 4.2 була побудована структура мережі PNN, що показана на рис. 4.4.

Зазначимо, що в побудованій мережі PNN кожному класу відповідає лише

один навчальний приклад, тому на рис. 4.4 всі нейрони шару образів позначені індексом 1.

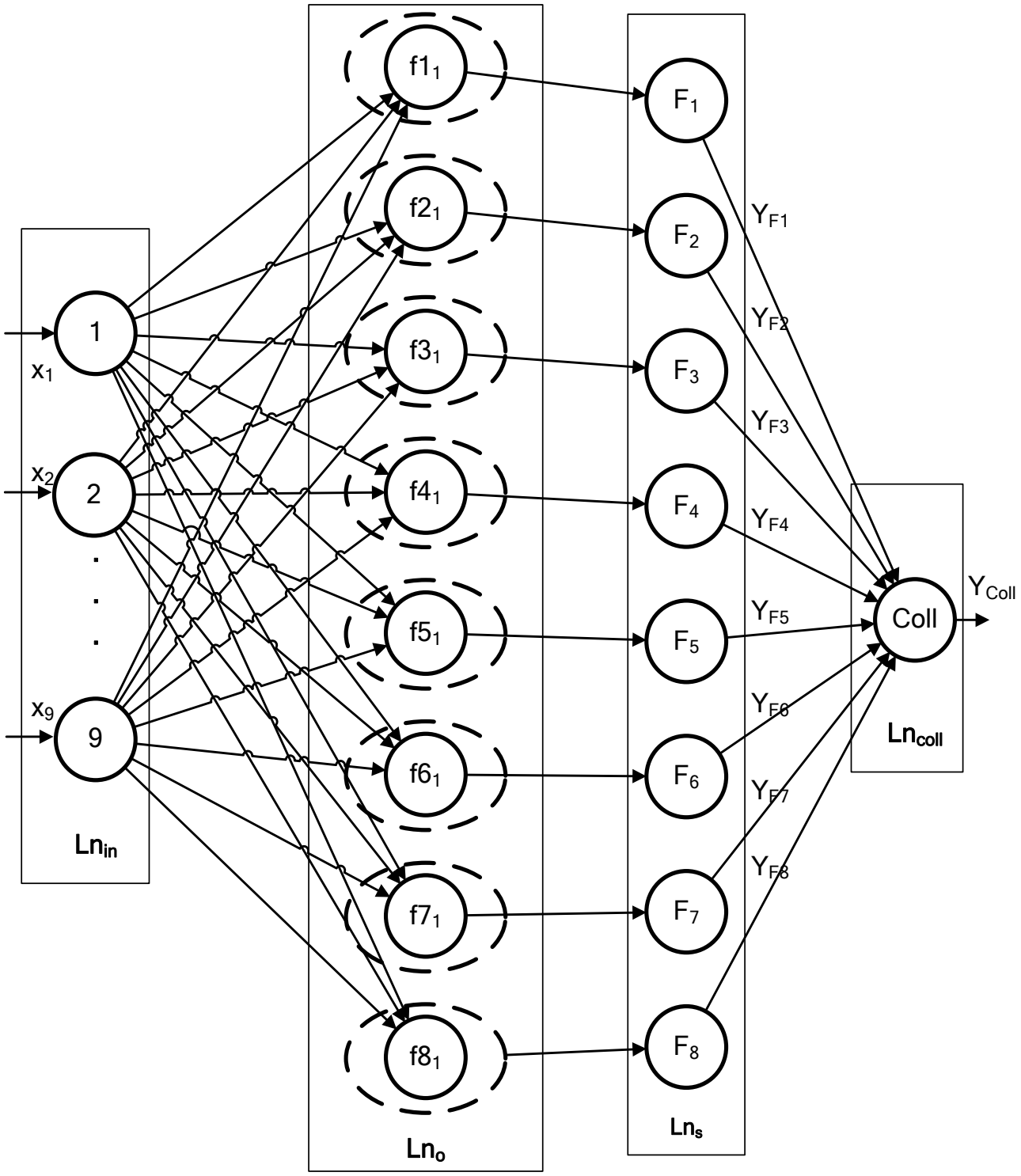

Рис. 4.4. Структура PNN для розпізнавання геометричних образів

Застосування побудованої мережі дозволило провести кодування вихідного сигналу для кожного з еталонів. Значення цих сигналів представлені в табл. 4.3.

Таблиця 4.3

Значення очікуваного вихідного сигналу

Вихідний сигнал	**Назва еталону**							
	F_1	**F_2**	**F_3**	**F_4**	**F_5**	**F_6**	**F_7**	**F_8**
Y_{F1}	9,00	5,54	5,54	3,81	3,81	6,41	5,54	7,27
Y_{F2}	5,54	9,00	5,54	7,27	3,68	6,41	5,54	3,81
Y_{F3}	5,54	5,54	9,00	7,27	3,81	6,41	5,54	3,81
Y_{F4}	3,81	7,27	7,27	9,00	5,54	4,68	3,81	2,08
Y_{F5}	3,81	3,81	3,81	5,54	9,00	1,22	3,81	5,54
Y_{F6}	6,41	6,41	6,41	4,68	1,22	9,00	6,41	4,68
Y_{F7}	5,54	5,54	5,54	3,81	3,81	6,41	9,00	7,27
Y_{F8}	7,27	3,81	3,81	2,08	5,54	4,68	7,27	9,00

Для застосування отриманих результатів в навчальних прикладах, орієнтованих на НММ з сигмоїдальною функцією активації, необхідно провести їх масштабування в діапазоні від 0 до 1. Для цього використано вираз:

$$y_m = \frac{y_{pr} - y_{min}}{y_{max} - y_{min}}, \tag{4.22}$$

де y_m – масштабоване значення очікуваного вихідного сигналу;

y_{pr} – початкове значення очікуваного вихідного сигналу, визначене за допомогою мережі PNN;

y_{min}, y_{max} – мінімальне та максимальне значення очікуваного вихідного сигналу, визначене за допомогою мережі PNN.

Застосувавши до даних табл. 4.3 вираз (4.22), розраховані масштабовані

значення вихідного сигналу, котрі наведено в табл. 4.4.

У випадку використання ГНМ з функцією активації типу ReLU реалізовувати масштабування необов'язково, хоча і рекомендовано в [139, 294, 316, 325].

Таблиця 4.4

Масштабовані значення очікуваного вихідного сигналу

Вихідний сигнал	**Назва еталону**							
	F_1	**F_2**	**F_3**	**F_4**	**F_5**	**F_6**	**F_7**	**F_8**
Ym_{F1}	1,00	0,56	0,56	0,33	0,33	0,67	0,56	0,78
Ym_{F2}	0,56	1,00	0,56	0,78	0,32	0,67	0,56	0,33
Ym_{F3}	0,56	0,56	1,00	0,78	0,33	0,67	0,56	0,33
Ym_{F4}	0,33	0,78	0,78	1,00	0,56	0,44	0,33	0,11
Ym_{F5}	0,33	0,33	0,33	0,56	1,00	0,00	0,33	0,56
Ym_{F6}	0,67	0,67	0,67	0,44	0,00	1,00	0,67	0,44
Ym_{F7}	0,56	0,56	0,56	0,33	0,33	0,67	1,00	0,78
Ym_{F8}	0,78	0,33	0,33	0,11	0,56	0,44	0,78	1,00

Отримані масштабовані значення очікуваних вихідних сигналів були використані при формуванні навчальних прикладів для ДШП.

Проведені порівняльні числові експерименти показали, що використання таких навчальних прикладів дозволяє в 1,3-1,5 рази зменшити кількість навчальних ітерацій для досягнення допустимої помилки навчання в межах 1%, відносно випадку застосування прикладів, в яких використана загальновідома процедура кодування очікуваного вихідного сигналу. Таким чином, ефективність застосування розробленої процедури визначення очікуваного вихідного сигналу НМЗ аналізу БП можна вважати доведеною.

З урахуванням виразів (4.1-4.9), розроблених методів формування вхідного поля НММ та запропонованої процедури використання мережі PNN для

визначення очікуваного вихідного сигналу, метод формування навчальних прикладів для НММ аналізу БП доцільно розділити на 5 етапів.

Етап 1. Формування вхідного поля. В залежності від виду БП виконання етапу полягає у реалізації одного із трьох розроблених методів формування вхідного поля НММ аналізу БП. Етап виконується для кожної реєстрації БП. Вхідними даними етапу 1 являються $\boldsymbol{Q}$, b, R, $\boldsymbol{U}$, $\boldsymbol{E}$, $\boldsymbol{S}$ та $\boldsymbol{D}$. Виходом етапу є множина $\boldsymbol{X}$.

Етап 2. Формування навчальних прикладів з символьним визначенням очікуваного вихідного сигналу. На вхід етапу подаються множини $\boldsymbol{X}$ та $\boldsymbol{D}$**.** Етап розділяється на три кроки.

Крок 2.1. Визначення кількості класів, що мають бути розпізнані. У випадку розпізнавання емоцій кількість класів дорівнює кількості емоцій, що мають бути розпізнані:

$$N = len(\boldsymbol{E}), \tag{4.23}$$

де N – кількість класів, що мають бути розпізнані;

len – функція визначення кількості елементів множини.

У випадку розпізнавання особи слухача СДН кількість класів дорівнює кількості слухачів:

$$N = len(\boldsymbol{S}). \tag{4.24}$$

Крок 2.3. Визначення символьних позначень класів, що мають бути розпізнані. У випадку розпізнавання емоцій очікуваний вихід співвідноситься з назвою емоції:

$$y_{sb,n} = e_n,\ \ e_n \in \boldsymbol{E},\ \ n \in [1,..N]\,, \tag{4.25}$$

де $y_{sb,n}$ – значення вихідного сигналу у символьному вигляді для n-го класу;

e_n – назва n-ої емоції.

У випадку розпізнавання особи слухача СДН назви класів співвідносяться з іменами слухачів:

$$y_{sb,n} = s_n,\ \ \ s_n \in \boldsymbol{S},\ \ n \in [1,..N]\,, \tag{4.26}$$

де s_n – ім’я n-го користувача.

Крок 2.3. Формування навчальних прикладів мережі PNN. Для цього використовується вираз (4.13).

Виходом етапу 2 являється кортеж виду:

$$\langle\langle \boldsymbol{x_1}, y_{sb,1}\rangle, \langle \boldsymbol{x_2}, y_{sb,2}\rangle, \dots, \langle \boldsymbol{x_N}, y_{sb,N}\rangle\rangle. \quad (4.27)$$

Етап 3. Побудова мережі PNN, що виконується за 5 кроків.

Крок 3.1. Визначення вхідних нейронів, котре полягає у співвіднесенні кожного із вхідних нейронів з певним вхідним параметром навчального прикладу:

$$k_{in} \to x_k, \;\; k = 1, \dots, len(\boldsymbol{x}), \quad (4.28)$$

де k_{in} – k-ий вхідний нейрон;

x_k – k-ий вхідний параметр.

Крок 3.2. Визначення нейронів шару образів, що описується виразом виду

$$n_o \to \langle \boldsymbol{x}_n, y_{sb,n}\rangle, \; n = 1, \dots N, \quad (4.29)$$

де n_o – n-ий нейрон шару образів;

$\langle \boldsymbol{x}_n, y_{sb,n}\rangle$ – n-ий навчальний приклад.

Крок 3.3. Визначення вагових коефіцієнтів нейронів шару образів. Реалізується за допомогою виразу (4.16).

Крок 3.4. Визначення нейронів шару додавання. Процес визначення описується так:

$$n_s \to y_{sb,n}, \quad (4.30)$$

де n_s – n-ий нейрон шару додавання.

Крок 3.5. Визначення вагових коефіцієнтів нейронів шару додавання.

Входом етапу є кортеж, заданий виразом (4.30). Виходом етапу є значення архітектурних параметрів мережі PNN.

Етап 4. Розрахунок очікуваного вихідного сигналу. На вхід етапу подаються вхідні параметри навчальних прикладів, задані виразом (4.27), та значення архітектурних параметрів мережі PNN, розраховані на кроці 3.5. Етап виконується

для кожного із навчальних прикладів та поділяється на два кроки.

Крок 4.1. Розрахунок величин вихідних сигналів нейронів шару додавання. Математичний апарат розрахунку складають вирази (4.14-4.19)

Крок 4.2. Масштабування значень вихідного сигналу. Для цього застосовуються вирази виду (4.22).

Виходом етапу є множина виду:

$$\{\boldsymbol{y}_{m.1}, \boldsymbol{y}_{m.2}, \dots, \boldsymbol{y}_{m.N}\}, \tag{4.31}$$

де $\boldsymbol{y}_{m.n}$ – множина масштабованих значень вихідних сигналів n-го навчального прикладу.

Етап 5. Визначення параметрів навчальних прикладів. Виконання етапу полягає в співставленні для кожного навчального прикладу множини значень вхідних та вихідних параметрів. Входом етапу являються множина вхідних параметрів $\boldsymbol{X}$ та множина вихідних параметрів, задана виразом (4.31). Виходом етапу, що співпадає із виходом методу, є кортеж виду:

$$\langle\langle \boldsymbol{x_1}, \boldsymbol{y}_{m,1}\rangle, \langle \boldsymbol{x_2}, \boldsymbol{y}_{m,2}\rangle, \dots, \langle \boldsymbol{x_N}, \boldsymbol{y}_{m,N}\rangle \rangle, \tag{4.32}$$

де $\langle \boldsymbol{x}_i, \boldsymbol{y}_{m,i}\rangle$ – кортеж, що складається із множини вхідних параметрів та множини масштабованих вихідних параметрів для і-го навчального прикладу.

Схема розробленого методу формування навчальних прикладів для НММ аналізу БП показана на рис. 4.5, а його математичне забезпечення складають вирази (4.1-4.32). Розпаралелювання розрахунків в даному методі забезпечується за рахунок того, що етап 1 може виконуватись паралельно з виконанням етапу 2.

Таким чином, в результаті проведених досліджень отримав подальший розвиток метод формування навчальних прикладів для НММ аналізу БП, що за рахунок обробки навчальних прикладів мережею PNN забезпечує можливість приблизно в 1,3-1,5 рази зменшити кількість навчальних ітерацій для досягнення помилки навчання в межах 1%, відносно випадку застосування прикладів, в яких

використана загальновідома процедура кодування очікуваного вихідного сигналу.

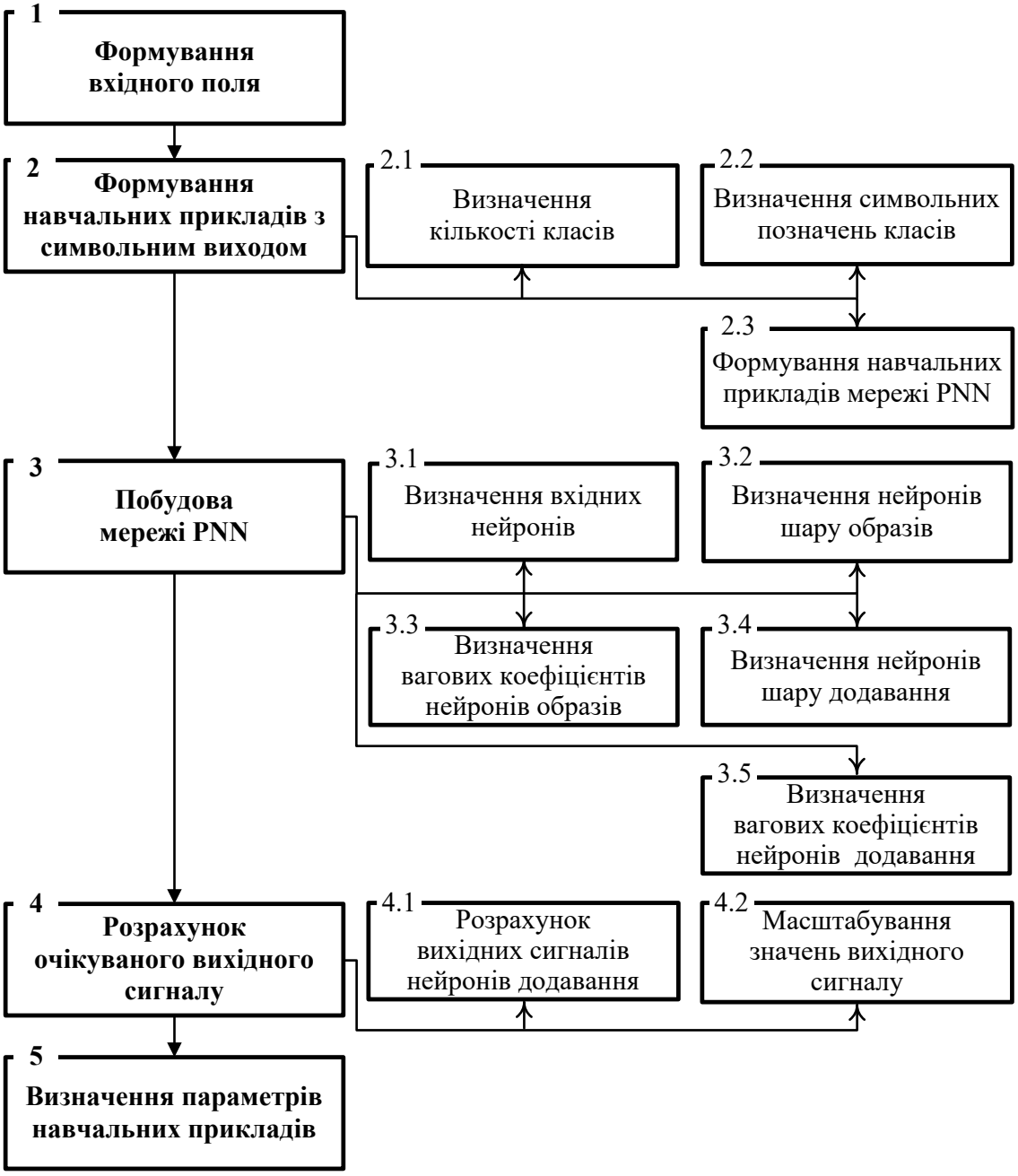

Рис. 4.5. Схема методу формування навчальних прикладів для нейромережевої моделі аналізу біометричних параметрів

Перспективи подальших досліджень полягають у вдосконаленні математичного забезпечення методу формування навчальних прикладів в напрямку збільшення схожості між очікуваними вихідними сигналами еталонів класів, що

мають бути розпізнані. Для цього слід розробити нову функцію активації нейронів шару образів, оскільки використання функції виду (4.17) призводить до того, що при подачі на вхід PNN навчального прикладу, який відповідає k-му еталону, вихідний сигнал k-го нейрону шару додавання є близьким до числа, котре дорівнює кількості вхідних параметрів НММ, що як мінімум в 5 разів перевищує вихідні сигнали інших нейронів шару додавання. Ще одним шляхом вдосконалення може бути заміна в методі мережі PNN на іншу НММ з більш високою обчислювальною ефективністю.

4.2. Метод розробки архітектури згорткової нейронної мережі, призначеної для аналізу біометричних параметрів

Базуючись на дослідженнях, наведених в першому розділі даної дисертаційної роботи, використовуючи запропонований підхід до визначення конструктивних параметрів та аналогію з відомими методами подібного призначення [96, 127, 273, 295], визначено, що перетворення інформації, яке здійснюється методом розробки архітектури ЗНМ, призначеної для аналізу БП, може бути описане за допомогою виразів:

$$\langle \boldsymbol{U}^{CNN}, \boldsymbol{V}_I^{CNN}, \boldsymbol{R}^{CNN}, \boldsymbol{\Psi}, \boldsymbol{Y} \rangle \rightarrow \langle \boldsymbol{v}_{max}^{CNN}, \boldsymbol{m}_{max}^{CNN} \rangle \tag{4.33}$$

$$h_{D1}^{CNN}\left(v_i^{CNN}\right) \rightarrow max, \qquad v_i^{CNN} \in \boldsymbol{V}_I^{CNN}, \tag{4.34}$$

$$h_{D2}^{CNN}\left(v_{max}^{CNN}, m_k^{CNN}\right) \rightarrow max, \qquad m_k^{CNN} \in \{M^{CNN}\}_I, v_{max}^{CNN} \in \{V^{CNN}\}_I, \tag{4.35}$$

де $\boldsymbol{U}^{CNN}$ – множина вимог до ЗНМ;

$\boldsymbol{R}^{CNN}$ – множина критеріїв ефективності видів ЗНМ;

$\boldsymbol{\Psi}$ – множина, що визначає вхідне поле ЗНМ;

$\boldsymbol{Y}$ – множина, що визначає вихідне поле ЗНМ;

$\boldsymbol{v}_{max}^{CNN}$ – множина найбільш ефективних видів ЗНМ;

$\boldsymbol{m}_{max}^{CNN}$ – множина параметрів найбільш ефективних видів ЗНМ;

v_i^{CNN} – і-ий вид ЗНМ;

$h_{D1}^{CNN}\left(v_i^{CNN}\right)$ – ефективність і-го виду ЗНМ;

$h_{D2}^{CNN}\left(v_i^{CNN}, m_k^{CNN}\right)$ – ефективність і-го виду ЗНМ при k-ій комбінації конструктивних параметрів;

$\boldsymbol{V}_I^{CNN}$ – множина прийнятних видів ЗНМ;

$\{M^{CNN}\}_I$ – множина допустимих комбінацій параметрів ЗНМ.

Зазначимо, що вирази (4.34, 4.35) являються версією (2.5, 2.6), адаптованою до задачі визначення конструктивних параметрів ЗНМ. Використавши (2.9), отримано:

$$h_{D1}^{CNN}\left(v_i^{CNN}\right) = \sum_{n=1}^{N}\left(\alpha_n^{CNN} \times 2^{R_n^{CNN}\left(v_i^{CNN}\right)}\right), \tag{4.36}$$

де α_n^{CNN} – ваговий коефіцієнт n-го критерію ефективності ЗНМ.

Значення n-го вагового коефіцієнту відображає важливість n-го критерію ефективності в умовах поставленої задачі розпізнавання емоцій та особи слухача СДН. Тобто значення критеріїв ефективності відображають можливості ЗНМ, а значення вагових коефіцієнтів – вимоги до ЗНМ.

Використавши [254, 263, 293, 302], в базовому випадку множину прийнятних видів ЗНМ записано у вигляді:

$$\boldsymbol{V}_I^{CNN} = \{v_{AN}^{CNN}, v_{Of}^{CNN}, v_{VG16}^{CNN}, v_{VG19}^{CNN}, v_{Ic}^{CNN}, v_{GN}^{CNN}, v_{MN}^{CNN}, v_{RN}^{CNN}, v_{Nn}^{CNN}, v_{EN}^{CNN}, v_{SN}^{CNN}, v_{LN}^{CNN}, v_{CNE}^{CNN}\}, \tag{4.37}$$

де v_{AN}^{CNN} – ЗНМ типу AlexNet;

v_{Of}^{CNN} – ЗНМ типу Overfeat;

v_{VG16}^{CNN} – ЗНМ типу VGG-16;

v_{VG19}^{CNN} – ЗНМ типу VGG-19;

v_{Ic}^{CNN} – ЗНМ типу Inception;

v_{GN}^{CNN} – ЗНМ типу GoogleNet;

v_{MN}^{CNN} – ЗНМ типу MobileNet;

v_{RN}^{CNN} – ЗНМ типу ResNet;

v_{Nn}^{CNN} – ЗНМ типу Network-in-network;

v_{EN}^{CNN} – ЗНМ типу ENet;

v_{SN}^{CNN} – ЗНМ типу SqueezeNet;

v_{LN}^{CNN} – ЗНМ типу LeNet;

v_{CNE}^{CNN} – ЗНМ типу CapsNetEm.

В подальшому множину прийнятних видів ЗНМ слід адаптувати до нових досягнень в області НМ. Визначений на основі вимог до НММ, призначених для розпізнавання емоцій та особи слухача СДН, перелік критеріїв ефективності ЗНМ наведено в табл. 4.5.

Таблиця 4.5

Перелік критеріїв ефективності ЗНМ

Позначення критерію	**Вимога, що відповідає критерію**
R_1^{CNN}	Наявність доступних навчальних БД БП
R_2^{CNN}	Наявність доступного та апробованого інструментарію для реалізації ЗНМ
R_3^{CNN}	Низька ресурсоємність
R_4^{CNN}	Низький термін навчання
R_5^{CNN}	Достатня точність розпізнавання
R_6^{CNN}	Можливість розпаралелювання обчислень
R_7^{CNN}	Низький термін розпізнавання
R_8^{CNN}	Пристосованість до розпізнавання повернутих зображень

Передбачається, що кожен із критеріїв ефективності може прийняти значення

в діапазоні від -1 до 1. Значення критерію дорівнює 1, коли відповідна йому вимога забезпечується в даному виді ЗНМ, 0 – коли забезпечується частково і -1 – коли вимога не забезпечується. Зазначимо, що однією із відмінностей видів ЗНМ між собою є різний розмір вхідного поля. При цьому однією із основних характеристик навчальних БД БП також являється фіксований розмір вхідного поля. Тому можливість застосування навчальної БД залежить від співвідношення вказаних розмірів вхідного поля. Розраховані за допомогою методу експертного оцінювання значення наведених критеріїв для прийнятних видів ЗНМ наведено в табл. 4.6.

Таблиця 4.6

Значення критеріїв ефективності для прийнятних видів ЗНМ

R	**Вид ЗНМ**												
	v_{AN}^{CNN}	v_{Of}^{CNN}	v_{VG16}^{CNN}	v_{VG19}^{CNN}	v_{Ic}^{CNN}	v_{GN}^{CNN}	v_{MN}^{CNN}	v_{RN}^{CNN}	v_{Nn}^{CNN}	v_{EN}^{CNN}	v_{SN}^{CNN}	v_{LN}^{CNN}	v_{CNE}^{CNN}
R_1^{CNN}	0	0	0	0	0	0	0	0	0	0,5	0,5	-0,1	-0,9
R_2^{CNN}	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,9	0,9	1	-0,1
R_3^{CNN}	0	0	0	0	0	0	0,9	0,8	0,8	0,9	0,9	0	0,5
R_4^{CNN}	0,5	0,5	0,5	0,5	0,5	0,5	0,8	0,5	0,5	0,8	0,9	0,5	0
R_5^{CNN}	0,9	0,9	0,9	0,9	0,9	0,9	0,8	0,9	0,9	0,5	0,8	0,3	0,7
R_6^{CNN}	0,8	0,9	0,8	0,8	0,9	0,9	0,9	0,9	0,9	0,9	0,9	-0,5	-0,5
R_7^{CNN}	-0,3	-0,2	-0,3	-0,2	-0,2	-0,2	-0,2	-0,2	-0,2	0	0	0,9	0
R_8^{CNN}	0	0	0	0	0	0,1	0	0	0	0	0	-0,5	0,7

Базуючись на аналізі структур прийнятних видів ЗНМ, з урахуванням результатів теоретичних робіт [88, 100, 171, 314], визначено, що основними конструктивними параметрами даного типу НММ є:

- Кількість вхідних нейронів - L_{in}.

- Кількість вихідних нейронів - L_{out}.

- Кількість нейронів в повнозв'язному шарі - L_f.

- Кількість шарів згортки - K_{ls}.

- Кількість карт ознак в кожному шарі згортки - $K_{h,k}, k \in [1; K_{ls}]$.

- Кількість шарів підвибірки (субдіскретізації) - K_{ld}.

- Масштабний коефіцієнт для кожного шару підвибірки - $m_l, l \in [1; K_{ld}]$.

- Розмір ядра згортки для кожного k-го шару згортки - $(b \times b)_k, k \in [1; K_{ls}]$.

- Зміщення рецептивного поля для k-ої процедури згортки - $d_k \in [1; K_{ls}]$.

- Структура зв'язків між сусідніми шарами згортки/підвибірки. Цю структуру можна представити у вигляді матриці:

$$\boldsymbol{Q}_{i,i+1} = \begin{Vmatrix} q_{(i,1),(i+1,1)} & \cdots & q_{(i,1),(i+1,J)} \\ \cdots & \cdots & \cdots \\ q_{(i,G),(i+1,1)} & \cdots & q_{(i,G),(i+1,J)} \end{Vmatrix}, \tag{4.38}$$

де G - кількість карт в і-му шарі;

J - кількість карт в (і + 1) -му шарі;

$\boldsymbol{Q}_{i,i+1}$- матриця, компоненти якої визначають наявність зв'язків між і-им та (і+1)-им прихованими шарами;

$q_{(i,g),(i+1,j)}$- компонент, який вказує на наявність/відсутність зв'язку між g-ою картою і-го шару та j-ою картою (і + 1)-го шару.

При цьому $q_{(i,g),(i+1,j)} = 1$, якщо зв'язок між g-ою картою і-го шару та j-ою картою (і + 1)-го шару є, а $q_{(i,g),(i+1,j)} = 0$, якщо зв'язку між g-ою картою і-го шару і j-ою картою (і + 1)-го шару немає. Використовуючи визначені конструктивні параметри ЗНМ та з урахуванням того, що ефективність ЗНМ може бути співставлена з помилкою розпізнавання, вираз (4.35) можливо переписати так:

$$\begin{cases} \Delta\left(L_{in}, L_{out}, L_f, K_{ls}, K_{h,k}, K_{ld}, b_k, d_k, \left|\boldsymbol{Q}_{i,i+1}\right|_{K_{ls}}\right) \to \min \\ R \le R_{max} \end{cases}, \tag{4.39}$$

де Δ – помилка розпізнавання;

R – ресурсоємність ЗНМ;

R_{max} – максимально допустима ресурсоємність ЗНМ;

$\left|\boldsymbol{Q}_{i,i+1}\right|_{K_{ls}}$ – вектор, що складається з матриць, які визначають зв'язки між сусідніми схованими шарами нейронів.

Відсутність у виразі (4.39) інших конструктивних параметрів ЗНМ пояснюється тим, що вони є або похідними від компонентів, які використовуються у цьому виразі, або визначення величин цих параметрів є слабко формалізованим. Після визначення наведених конструктивних параметрів розрахунок інших структурних параметрів ЗНМ можливо реалізувати так:

$$H_{CNN,k} = \left(H_{CNN,k-1} - b_k + 2r_k\right)/d_k + 1, \tag{4.41}$$

$$H_{S,j} = H_{CNN,k}/m_j, \tag{4.42}$$

$$L_{f,s} \leq \mathrm{P}, \tag{4.43}$$

$$2\left(L_{h,end} + 1\right) \leq L_{f,s}, \tag{4.44}$$

$$L_{f,k} = {L_{f,s}}\Big/{L_{FC}}, \tag{4.45}$$

де $H_{CNN,k}$ – розмір карти ознак для k-го шару згортки;

r_i – кількість доповнюючих нулів для і-го шару згортки;

$H_{S,j}$ – розмір j-го шару субдискретизації;

$L_{h,end}$ – кількість карт ознак в останньому шарі згортки;

$L_{f,s}$ – загальна кількість нейронів в повнозв'язних шарах;

$L_{f,k}$ – кількість нейронів в k-му повнозв'язному шарі;

L_{FC} – кількість повнозв’язних шарів нейронів;

P – кількість навчальних прикладів.

Інтеграція загальноприйнятої методології побудови нейромережевих систем розпізнавання [183, 189, 191, 246], сформованих у другому розділі принципів адаптації структури ЗНМ до умов застосування, та типовими обмеженнями на процес попередньої обробки БП слухачів СДН, дозволила запропонувати метод

розробки архітектури ЗНМ, призначеної для аналізу БП. Реалізація методу передбачає виконання 7 етапів, а його узагальнена схема показана на рис. 4.6.

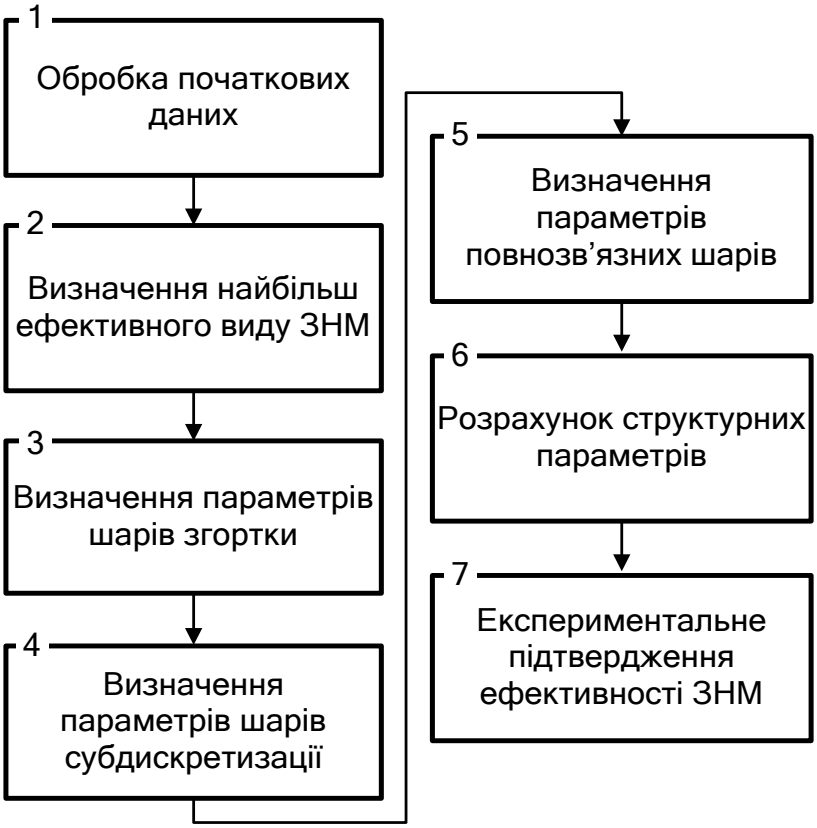

Рис. 4.6. Узагальнена схема методу розробки архітектури ЗНМ, призначеної для аналізу БП

Етап 1. Обробка початкових даних. Даний етап призначений для підготовки даних, що використовуються для розрахунку значень функції ефективності кожного із прийнятних видів ЗНМ. На вхід етапу подаються $\boldsymbol{U}^{CNN}$, $\boldsymbol{R}^{CNN}$ та $\{V^{CNN}\}_I$. Виконання етапу розділяється на три кроки.

Крок 1.1. Модернізація множини прийнятних видів ЗНМ. Виконання кроку полягає в аналізі розробок нових видів ЗНМ та, за необхідності, в модернізації множини $\{V^{CNN}\}_I$.

Крок 1.2. Визначення множини апробованих видів ЗНМ, що полягає в обмеженні множини прийнятних видів ЗНМ відповідно до можливості їх навчання

та описується за допомогою виразу (2.10).

Крок 1.3. Визначення вагових коефіцієнтів. На даному кроці шляхом експертного оцінювання вимог до ЗНМ здійснюється визначення критеріїв ефективності ЗНМ.

Виходом етапу 1 являється множина апробованих видів ЗНМ - $\{V^{CNN}\}_I$ та множина вагових коефіцієнтів критеріїв ефективності ЗНМ - $\boldsymbol{\alpha}^{CNN}$.

Етап 2. Визначення найбільш ефективних видів ЗНМ. На вхід етапу подаються $\boldsymbol{\alpha}^{CNN}$, $\boldsymbol{R}^{CNN}$ та $\{V^{CNN}\}_I$. Виконання етапу розділяється на три кроки.

Крок 2.1. Розрахунок значення функції ефективності, що реалізується для кожного прийнятного виду ЗНМ. Для цього застосовується вираз (4.36).

Крок 2.2. Визначення найбільш ефективного виду ЗНМ, тобто виду, для якого значення функції ефективності максимальне.

Крок 2.3. Формування множини ефективних видів ЗНМ, що здійснюється за допомогою виразу (4.34). До вказаної множини входять види ЗНМ, ефективність яких відрізняється від максимального значення ефективності не більше, ніж на 10%.

Виходом другого етапу являється множина $\boldsymbol{v}_{max}^{CNN}$.

Етап 3. Визначення параметрів шарів згортки, що реалізується за допомогою експертного оцінювання на основі запропонованої у другому розділі підходу до визначення конструктивних параметрів ЗНМ та запропонованих принципів адаптації ЗНМ. Етап розділяється на чотири кроки.

Крок 3.1. Визначення кількості шарів згортки, що здійснюється на основі першого запропонованого принципу.

Крок 3.2. Визначення кількості карт ознак. Визначення здійснюється для кожного шару згортки на основі другого запропонованого принципу. Для останнього шару згортки додатково використовується вираз (4.44).

Крок 3.3. Формування структури зв’язків, що реалізується на основі третього запропонованого принципу.

Крок 3.4. Визначення розміру ядра згортки, що здійснюється для кожного шару згортки на основі четвертого та п’ятого запропонованих принципів.

Виходом етапу являється K_{ls}, $K_{h,k}$, b_k, d_k, r_k, $\left|\boldsymbol{Q}_{i,i+1}\right|_{K_{ls}}$.

Етап 4. Визначення параметрів шарів субдискретизації. Вхідними даними етапу являються $\boldsymbol{\Psi}$, K_{ls}. Етап розділяється на два кроки.

Крок 4.1. Розрахунок кількості шарів субдискретизації. В базовому варіанті їх кількість дорівнює кількості шарів згортки.

Крок 4.2. Розрахунок масштабних коефіцієнтів, що здійснюється для кожного із шарів субдискретизації. Розрахунок може здійснюватися із позицій зведення зображення, що подається на вхід ЗНМ, до однієї точки.

Виходом етапу являється K_{ld} та множина m_l.

Етап 5. Визначення параметрів повнозв’язних шарів. На вхід етапу подаються K_{ls}, $K_{h,k}$, $\boldsymbol{Y}$. Етап розділяється на два кроки.

Крок 5.1. Розрахунок загальної кількості нейронів. Для розрахунку використовується вираз (4.45).

Крок 5.2. Розрахунок параметрів кожного шару. В першому наближенні кількість повнозв’язних шарів дорівнює L_{FC}=2. Кількість нейронів в кожному шарі розраховується за допомогою виразу (4.44).

Виходом етапу являються L_{FC}, $L_{f,s}$ та $L_{f,L_{FC}}$.

Етап 6. Розрахунок структурних параметрів. Вхідними даними етапу являються $\boldsymbol{\Psi}$, K_{ls}, $K_{h,k}$, K_{ld}, m_l. Етап розділяється на два кроки.

Крок 6.1. Розрахунок розмірів шарів згортки, що здійснюється за допомогою виразу (4.41).

Крок 6.2. Розрахунок розмірів шарів субдискретизації, що здійснюється за допомогою вирізу (4.42).

Виходом етапу є $H_{CNN,k}$ та $H_{S,j}$ для всіх шарів згортки та субдискретизації.

Етап 7. Експериментальне підтвердження ефективності ЗНМ. На вхід етапу

подаються всі види ефективних ЗНМ та їх конструктивні параметри - $\boldsymbol{\Psi}$, $\boldsymbol{Y}$, K_{ld}, $\boldsymbol{v}_{max}^{CNN}$, m_l, b_k, d_k, r_k, $\left|\boldsymbol{Q}_{i,i+1}\right|_{K_{ls}}$, K_{ls}, $K_{h,k}$, K_{ld}, m_l, L_{FC}, $L_{f,s}$ та $L_{f,L_{FC}}$, $H_{CNN,k}$ та $H_{S,j}$. Етап розділяється на два кроки.

Крок 7.1. Проведення комп'ютерних досліджень, що спрямовані на визначення ефективності ЗНМ.

Крок 7.2. Уточнення архітектурних параметрів ЗНМ.

Виходом етапу є уточненні значення вказаних параметрів. Деталізована схема 1-3 етапів методу розробки архітектури ЗНМ, призначеної для аналізу БП, показана на рис. 4.7, а схема 4-7 етапів на рис. 4.8.

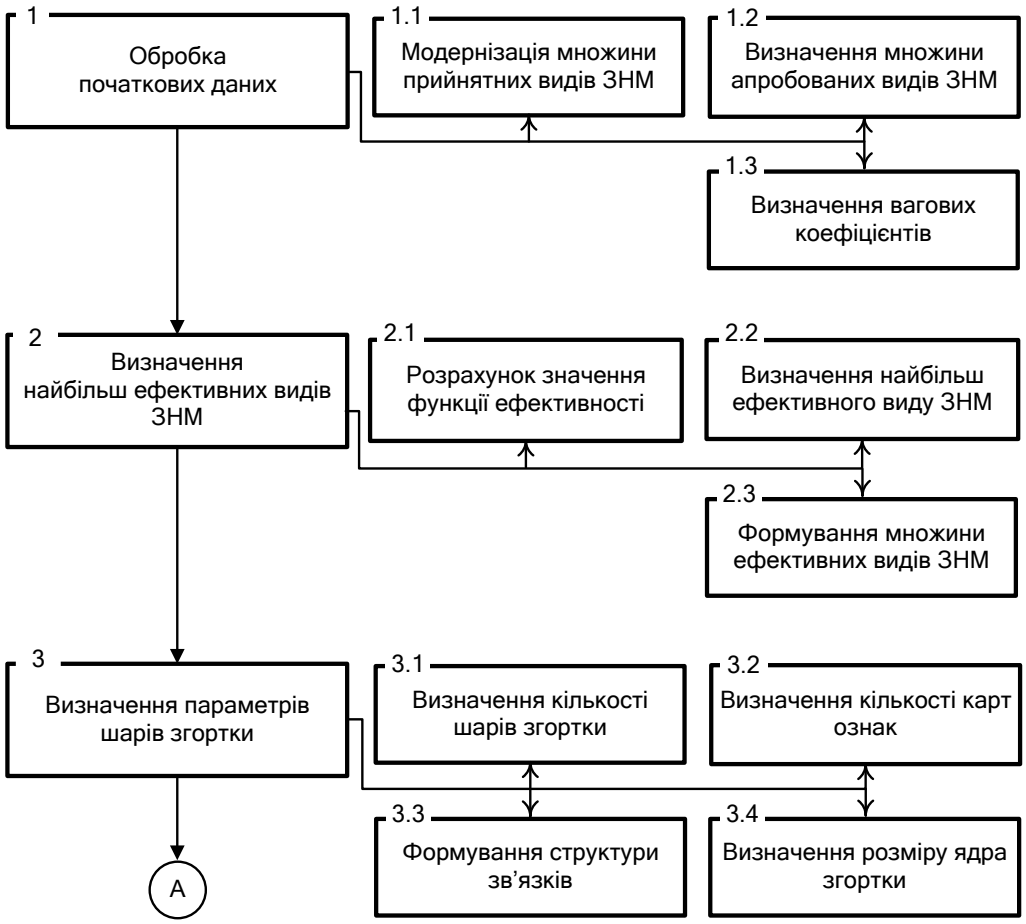

Рис. 4.7. Деталізована схема 1-3 етапів методу розробки архітектури ЗНМ, призначеної для аналізу БП

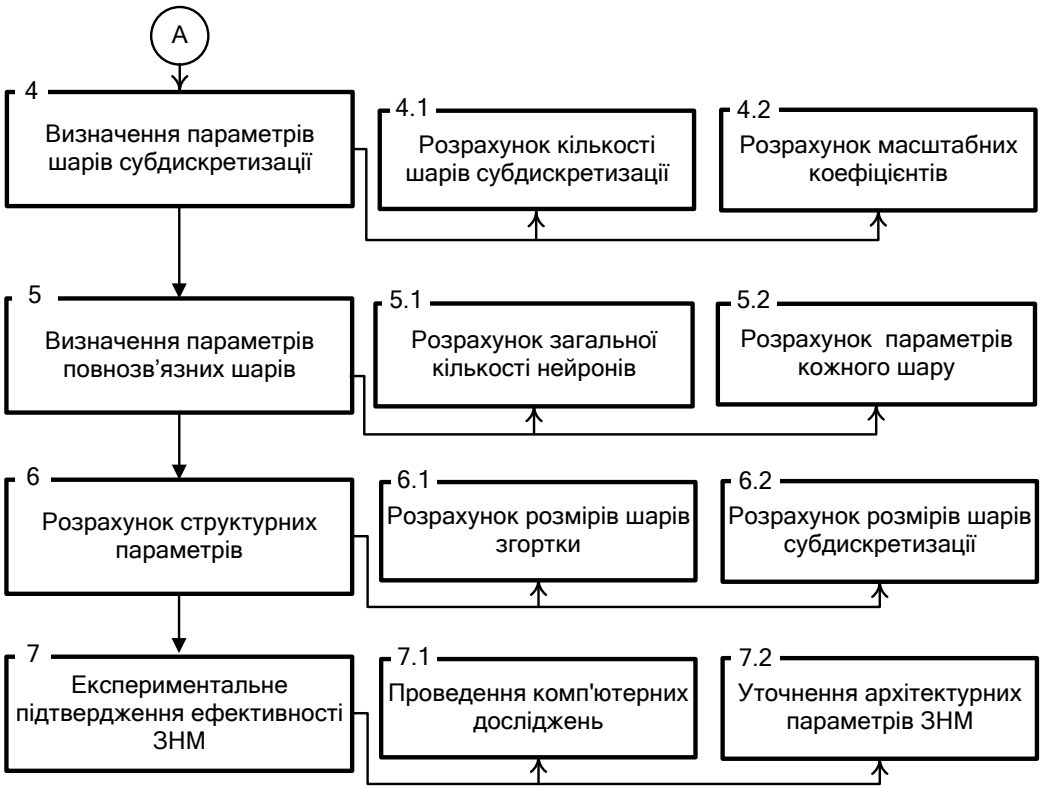

Рис. 4.8. Деталізована схема 4-7 етапів методу розробки архітектури ЗНМ, призначеної для аналізу БП

Таким чином, удосконалено метод розробки архітектури ЗНМ, призначеної для аналізу БП, що за рахунок запропонованих критеріїв ефективності виду ЗНМ та принципів адаптації конструктивних параметрів забезпечує можливість зменшення обсягу експериментальних досліджень, пов'язаних з розробкою архітектури ЗНМ. Це пояснюється зменшенням діапазону параметрів ЗНМ, що підлягають дослідженню. Перспективи вдосконалення методу полягають у модифікації множини прийнятних видів ЗНМ за рахунок долучення новітніх архітектур, що можуть бути реалізовані за допомогою доступного інструментарію. Крім того, доцільно імплементувати в метод процедури розрахунку параметрів блоків Stem Network, Inception та Residul, що використовуються в ЗНМ типу GoogleNet та ResNet. Також представляє науковий інтерес застосування отриманих рішень в області медичної діагностики [161].

4.3. Метод нейромережевого розпізнавання емоцій

Оскільки результат виконання запропонованого методу співвідноситься з розпізнаними емоціями слухача СДН, то передбачається, що на вхід даного методу може подаватись тільки один із основних БП – або зображення обличчя, або ГС, або КП. Використавши дані [264, 272, 304] і результати другого та третього розділу, визначено, що в загальному вигляді перетворення інформації, яке здійснюється в цьому методі, може бути описане виразами (4.46-4.50).

$$\langle \boldsymbol{BP^b}, \boldsymbol{E}, \boldsymbol{V^a}, \langle D^a \rangle, \boldsymbol{R^a}, \boldsymbol{M^{Va}}, \boldsymbol{\Theta^b}, \boldsymbol{\Theta^{hv,b}}, \boldsymbol{\rho} \rangle \rightarrow \boldsymbol{O^E} , \quad (4.46)$$

де $\boldsymbol{BP^b}$ – множина, що містить зареєстровані значення одного із основних БП;

$\boldsymbol{E}$ – множина базових емоцій, що мають бути розпізнані;

$\boldsymbol{V^a}$ – множина прийнятних видів НММ;

$\boldsymbol{R^a}$ – множина критеріїв ефективності прийнятних видів НММ;

$\langle D^a \rangle$ – кортеж, що складається із множин, котрі містять експертні дані стосовно НММ, які застосовуються для аналізу основних БП, та експертні дані, що стосуються обробки БП;

$\boldsymbol{M^{Va}}$ – множина параметрів прийнятних видів НММ;

$\boldsymbol{\Theta^b}$ – множина параметрів, що стосуються умов реєстрації та обробки основних БП;

$\boldsymbol{\Theta^{hv,b}}$ – множина параметрів, що характеризують АПЗ, яке використовується для реєстрації основних БП;

$\boldsymbol{\rho}$ – множина мінімально допустимих показників ефективності НММ;

$\boldsymbol{O^E}$ – множина, що містить результати розпізнавання базових емоцій.

$$\boldsymbol{BP^b} \in \langle \boldsymbol{BP^f}, \boldsymbol{BP^v}, \boldsymbol{BP^{ks}} \rangle, \quad (4.47)$$

де $\boldsymbol{BP^f}$ – множина, що містить зареєстровані параметри зображення обличчя;

$\boldsymbol{BP^v}$ – множина, що містить зареєстровані параметри ГС;

$\boldsymbol{BP^{ks}}$ – множина, що містить зареєстровані параметри КП.

$$\langle D^a \rangle = \langle \boldsymbol{D_{Va}^{f,E}}, \boldsymbol{D_{\alpha}^{f,E}}, \boldsymbol{D_{Va}^{v,E}}, \boldsymbol{D_{\alpha}^{v,E}}, \boldsymbol{D_{Va}^{ks,E}}, \boldsymbol{D_{\alpha}^{ks,E}}, \boldsymbol{D_{pr}^{f}}, \boldsymbol{D_{pr}^{v}}, \boldsymbol{D_{pr}^{ks}} \rangle, \quad (4.48)$$

де $\boldsymbol{D_{\alpha}^{f,E}}$, $\boldsymbol{D_{\alpha}^{v,E}}$, $\boldsymbol{D_{\alpha}^{ks,E}}$ – множини експертних даних щодо вагових коефіцієнтів критеріїв ефективності апробованих видів НММ для розпізнаванні емоцій за зображенням обличчя, ГС та КП;

$\boldsymbol{D_{V\alpha}^{f,E}}$, $\boldsymbol{D_{V\alpha}^{v,E}}$, $\boldsymbol{D_{Va}^{ks,E}}$ – множини експертних даних щодо можливості застосування апробованих видів НММ для розпізнавання емоцій за зображенням обличчя, ГС та КП.

$\boldsymbol{D_{pr}^{f}}, \boldsymbol{D_{pr}^{v}}, \boldsymbol{D_{pr}^{ks}}$ – множини експертних даних, що стосуються обробки параметрів зображення обличчя, ГС та КП.

$$\boldsymbol{\Theta^{b}} = \{\boldsymbol{\Theta^{f}}, \boldsymbol{\Theta^{v}}, \boldsymbol{\Theta^{ks}}\}, \quad (4.49)$$

де $\boldsymbol{\Theta^{f}}, \boldsymbol{\Theta^{v}}, \boldsymbol{\Theta^{ks}}$ – множини параметрів, що стосуються умов реєстрації та обробки зображення обличчя, ГС та КП.

$$\boldsymbol{\Theta^{hv,b}} = \{\boldsymbol{\Theta^{hb,f}}, \boldsymbol{\Theta^{hb,v}}, \boldsymbol{\Theta^{hb,ks}}\}, \quad (4.50)$$

де $\boldsymbol{\Theta^{hb,f}}, \boldsymbol{\Theta^{hb,v}}, \boldsymbol{\Theta^{hb,ks}}$ – множини параметрів, що характеризують АПЗ, яке використовується для реєстрації зображення обличчя, ГС та КП.

Передбачено, що формування $\boldsymbol{\Theta^{f}}$ та $\boldsymbol{\Theta^{hb,f}}$ здійснюється за допомогою виразів (3.23, 3.25, 3.26), формування $\boldsymbol{\Theta^{v}}$ та $\boldsymbol{\Theta^{hb,v}}$ – за допомогою виразів (3.73-3.80), а формування множин $\boldsymbol{\Theta^{ks}}$ та $\boldsymbol{\Theta^{hb,ks}}$ – за допомогою виразів (3.54-3.61). Базуючись на розробленій аналітичній моделі, що описується виразами (4.46-4.50, 3.23, 3.25, 3.26, 3.54-3.61, 3.73-3.80), використавши наведені в [94, 99, 245] апробовані підходи до створення НМЗ аналізу складних процесів та враховуючи розроблені елементи методологічного забезпечення технології нейромережевого розпізнавання, методи формування вхідного поля НММ та метод формування навчальних прикладів визначено, що метод нейромережевого розпізнавання емоцій доцільно розділити на 7 етапів, узагальнена схема виконання яких показана на рис. 4.9.

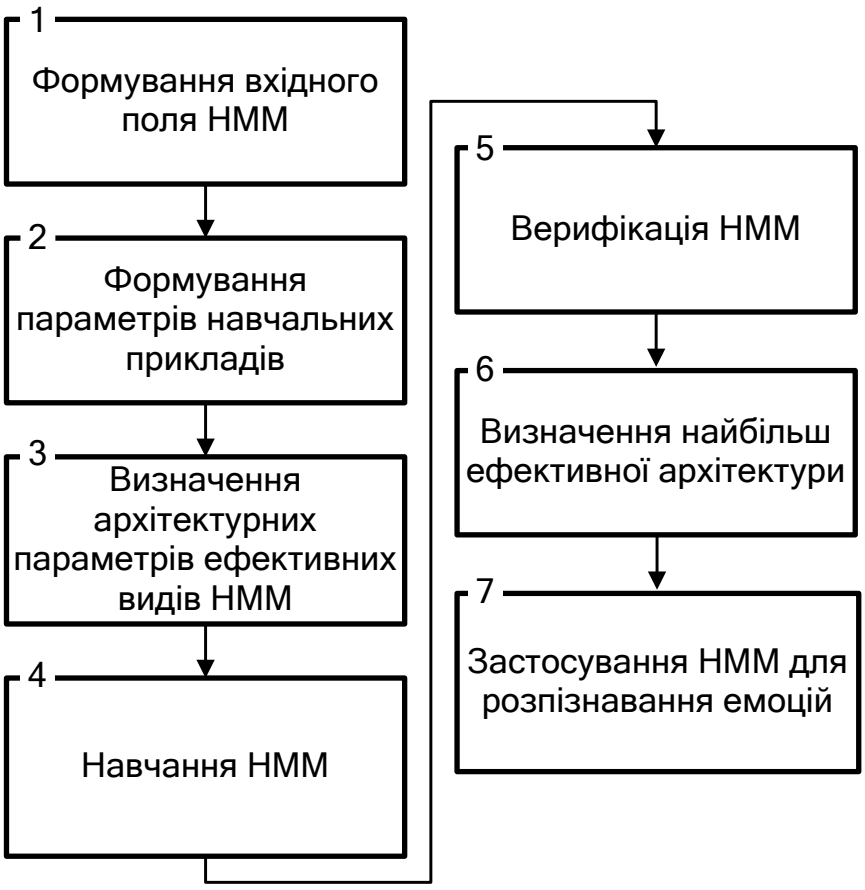

Рис. 4.9. Узагальнена схема методу нейромережевого розпізнавання емоцій

Етап 1. Формування вхідного поля НММ. На даному етапі проводиться аналіз можливості перетворення зареєстрованих показників БП у вхідні параметри НММ, призначеної для розпізнавання емоцій. На основі результатів аналізу визначається множина вхідних параметрів НММ, множина показників БП, які підлягають реєстрації та процедури обробки зареєстрованих значень вказаних показників. Також визначається процедура співставлення кожного із оброблених значень показників з конкретним вхідним параметром НММ.

Етап 2. Формування параметрів навчальних прикладів. Виконання даного етапу полягає у формуванні вихідного поля НММ та його співставленні з вхідним полем, що було отримане в результаті виконання попереднього етапу. При цьому у вихідному полі НММ у числовому вигляді повинні відображатись емоції слухача

СДН, які мають бути розпізнані.

Етап 3. Визначення архітектурних параметрів ефективних видів НММ. На даному етапі на основі прийнятних видів формується множина апробованих видів НММ. Після цього послідовно визначається множина ефективних видів НММ, визначаються конструктивні параметри та параметри НММ, що розраховуються на їх основі.

Етап 4. Навчання НММ. На цьому етапі спочатку з використанням результатів другого етапу визначається необхідний обсяг навчальної вибірки та здійснюється її наповнення. Базуючись на практичному досвіді, враховуючи високу складність створення достатньої кількості навчальних прикладів БП, призначених для НМЗ розпізнавання емоцій, та використовуючи рекомендації [89, 92], пропонується в якості джерела наповнення навчальної вибірки використовувати вільнодоступні апробовані БД. Після наповнення навчальна вибірка використовується для навчання НММ, що входять до множини ефективних видів. Результатом навчання являється множина вагових коефіцієнтів для кожної НММ, яка входить до множини ефективних видів.

Етап 5. Верифікація НММ. На даному етапі спочатку розраховуються значення показників, які характеризують ефективність НММ, що пройшла навчання. Після цього на основі порівняння означених показників з мінімально допустимими показниками визначається множина верифікованих НММ.

Етап 6. Визначення найбільш ефективної архітектури. На цьому етапі шляхом порівняння значень ефективності визначається НММ, найбільш ефективна серед верифікованих. Після цього за допомогою комп'ютерних експериментів проводиться уточнення її архітектурних параметрів.

Етап 7. Застосування НММ для розпізнавання емоцій. Виконання даного етапу полягає в отриманні нових зареєстрованих показників БП слухача СДН. Використовується БП того ж типу, який був використаний для формування

навчальних прикладів та побудови НММ. Надалі на основі вказаних показників формується множина параметрів, що подаються на вхід навченої НММ, архітектура якої була модифікована на попередньому етапі методу. Результатом подачі являється вихідний сигнал НММ, що свідчить про емоції слухача СДН.

Деталізація описаних етапів методу нейромережевого розпізнавання проведена з позицій необхідності врахування особливості нейромережевого аналізу кожного із основних БП.

Етап 1. Формування вхідного поля НММ. Виконання даного етапу співвідноситься з реалізацією розроблених в третьому розділі методів формування вхідного поля НММ. Реалізація проводиться з урахуванням необхідності розпізнавання виключно емоцій слухачів СДН. Входом даного етапу являються $\boldsymbol{BP^b}$, $\boldsymbol{\Theta^b}$, $\boldsymbol{\Theta^{hv,b}}$. Поділ етапу на кроки реалізується в залежності від виду зареєстрованого БП. Оскільки формування вхідного поля можливо реалізувати на основі трьох різних БП, то передбачено виконання даного етапу в трьох різних режимах. Режим А – на вхід етапу подається зображення обличчя, режим В – на вхід етапу подається ГС, режим С – на вхід етапу подається КП. Крім того, передбачено, що виконання даного етапу в одному із двох варіантів – спрощеному та розширеному. Спрощений варіант виконання етапу Simple відповідає випадку, коли вхідне поле НММ формується на основі типових БД, що містять зареєстровані та оброблені значення БП. Розширений варіант виконання Extended відповідає випадку, коли вхідне поле НММ формується на основі зареєстрованих значень, які потребують обробки. Кроки етапу залежать від режиму та варіанту виконання.

Режим А. Варіант Extended.

Крок А1.1. Центрування зареєстрованого зображення.

Крок А1.2. Обмеження розміру зображення.

Крок А1.3. Масштабування обмеженого зображення.

Крок А1.4. Первинна обробка масштабованого зображення.

Крок А1.5. Нівелювання впливу типових завад.

Крок А1.6. Визначення параметрів ключових точок для розпізнавання емоцій.

Крок А1.7. Нормалізація параметрів зображення та параметрів ключових точок.

Крок А1.8. Кодування параметрів зображення та параметрів ключових точок.

У випадку варіанту Simple виконуються кроки А1.7 та А1.8.

Режим В. Варіант Extended.

Крок В1.1. Базова оцінка значень зареєстрованих параметрів ГС.

Крок В1.2. Приведення параметрів ГС.

Крок В1.3. Додаткова оцінка значень зареєстрованих параметрів ГС.

Крок В1.4. Формування інтервалів розпізнавання.

Крок В1.5. Спектральний аналіз.

Крок В1.6. Розрахунок мел-кепстральних коефіцієнтів.

Крок В1.7. Формування вхідного поля.

У випадку варіанту Simple виконуються кроки В1.4-В1.7.

Режим С. Варіант Extended.

Крок С1.1. Визначення параметрів символьних даних.

Крок С1.2. Визначення параметрів вхідного поля.

Крок С1.3. Розрахунок параметрів растрів.

Крок С1.4. Розрахунок параметрів КП.

Крок С1.5. Формування вхідного поля ЗНМ.

У випадку варіанту Simple виконуються кроки С1.3-С1.5.

Виходом першого етапу являється множина $\boldsymbol{X}$, до складу якої входять:

- для режиму А - $\boldsymbol{\Psi}^{f}$ – множина параметрів, що описують вхідне поле ЗНМ, призначеної для розпізнавання емоцій слухача СДН на основі аналізу зображення обличчя, що представлене за допомогою виразів (3.23-3.25);

- для режиму В - $\boldsymbol{\Psi}^{v}$ - сформоване на основі параметрів ГС вхідне поле БШП

або ЗНМ, що представлене за допомогою виразів (3.81, 3.82);

- для режиму С - $\boldsymbol{\Psi}^{ks}$ - сформоване на основі параметрів КП вхідне поле ЗНМ, що представлене за допомогою виразу (3.72).

Етап 2. Формування параметрів навчальних прикладів. Передбачено два варіанти виконання даного етапу – розширений (Extended) та спрощений (Simple). У варіанті Extended виконання етапу співвідноситься з реалізацією 2-5 етапів наведеного в п. 4.1 методу формування навчальних прикладів для НММ аналізу БП. При цьому враховується, що вихідний сигнал в навчальному прикладі співвідноситься лише з результатом розпізнаванням емоцій. У варіанті Simple застосовується спрощена процедура кодування вихідного сигналу, що не передбачає врахування близькості еталонів образів, які мають бути розпізнані. Входом даного етапу являється $\boldsymbol{X}$, $\boldsymbol{E}$, $\boldsymbol{D}^{\boldsymbol{f}}_{\boldsymbol{pr}}, \boldsymbol{D}^{\boldsymbol{v}}_{\boldsymbol{pr}}, \boldsymbol{D}^{\boldsymbol{ks}}_{\boldsymbol{pr}}$. Етап поділяється на 4 кроки.

Варіант виконання Extended.

Крок 2.1. Формування навчальних прикладів з символьним виходом.

Крок 2.2. Побудова мережі PNN.

Крок 2.3. Розрахунок очікуваного вихідного сигналу.

Крок 2.4. Визначення параметрів навчальних прикладів.

У випадку варіанту Simple виконуються лише кроки 2.3 та 2.4.

Виходом етапу являється заданий виразом (4.32) кортеж $\langle Z \rangle$, що складається із множини вхідних параметрів та множини масштабованих вихідних параметрів для кожного навчального прикладу.

Етап 3. Визначення архітектурних параметрів ефективних видів НММ. На вхід етапу подаються $\boldsymbol{V^a}$, $\boldsymbol{R^a}$, $\langle D^a \rangle$, $\boldsymbol{M^{Va}}$, Z. Етап розділяється на 4 кроки.

Крок 3.1. Формування множини апробованих видів НММ, що реалізується шляхом експертного оцінювання видів НММ, які входять до складу множини $\boldsymbol{V^a}$.

Крок 3.2. Формування множини ефективних видів НММ. На даному кроці шляхом експертного оцінювання визначаються значення вагових коефіцієнтів

критеріїв ефективності НММ. Після цього послідовно за допомогою виразу (2.9) розраховуються значення функції ефективності прийнятних видів НММ, за допомогою виразу (2.5) визначається найбільш ефективний вид НММ та визначається множина ефективних видів НММ, для яких значення функції ефективності відрізняється від максимального значення не більше, ніж на 10%.

Крок 3.3. Визначення конструктивних параметрів ефективних видів НММ, що полягає у визначені переліку цих параметрів. Реалізація кроку залежить від виду НММ та передбачає використання результатів для БШП [104, 107], для ЗНМ [216], для рекурентних НМ [157, 177, 190, 286], а при використанні блоків кодеру-декодеру [156].

Крок 3.4. Розрахунок параметрів НММ. На даному кроці розраховуються величини параметрів НММ, перелік яких визначено на кроці 3.3. Визначення значень конструктивних параметрів та способів їх розрахунку залежать від виду НММ та реалізується з позицій досягнення максимальної точності розпізнавання емоцій при використанні допустимого обсягу обчислювальних ресурсів.

Виходом етапу являється множина ефективних видів НММ - $\boldsymbol{v_{max}}$ та множина параметрів ефективних видів НММ - $\boldsymbol{M^{v_{max}}}$.

Етап 4. Навчання НММ. На вхід етапу подаються $\boldsymbol{v_{max}}$, $\boldsymbol{M^{v_{max}}}$, $\langle Z \rangle$. Виконання етапу розділяється на 4 кроки.

Крок 4.1. Розрахунок мінімального обсягу навчальної вибірки. В базовому випадку, при використанні маркованих навчальних прикладів для НММ з вхідним полем у вигляді вектору, розрахунок реалізується так:

$$P = 10 \times N_x \times |\boldsymbol{E}|, \tag{4.51}$$

де N_x – кількість вхідних параметрів;

$|\boldsymbol{E}|$ – потужність множини $\boldsymbol{E}$.

Для ЗНМ мінімальна кількість навчальних прикладів розраховується так:

$$P = 6 \times L \times H \times |\boldsymbol{E}|, \tag{4.52}$$

де L, H - ширина та висота вхідного поля ЗНМ.

Крок 4.2. Наповнення навчальної вибірки, що доцільно реалізовувати на основі даних [46, 91, 176, 235].

Крок 4.3. Розрахунок вагових коефіцієнтів синаптичних зв'язків, що при використанні маркованих навчальних прикладів тренувальної вибірки реалізується на основі сучасних модифікацій алгоритму зворотнього поширення помилки, опис яких наведено в [230, 250].

Виходом етапу являється множина вагових коефіцієнтів синаптичних зв'язків $\boldsymbol{w}^{v_{max}}$ для кожної НММ, що входить до складу $\boldsymbol{v_{max}}$.

Етап 5. Верифікація НММ. На вхід даного етапу подаються $\boldsymbol{v_{max}}$, $\boldsymbol{w}^{v_{max}}$, $\boldsymbol{\rho}$. В базовому варіанті прийнято, що ефективність НММ співвідноситься з похибкою розпізнавання валідаційних даних δ. Виконання п'ятого етапу розділяється на два кроки.

Крок 5.1. Розрахунок точності розпізнавання, що полягає у розрахунку середньої похибки розпізнавання прикладів з валідаційної вибірки - δ_{av}.

Крок 5.2. Визначення множини верифікованих НММ, що полягає у вилученні із множини $\boldsymbol{v_{max}}$ НММ, для яких:

$$\delta_{av} \leq \delta_{min}, \tag{4.53}$$

де δ_{min} – мінімально допустима похибка розпізнавання.

Також на цьому кроці для верифікованих НММ визначається множина вагових коефіцієнтів.

Виходом етапу являється множина верифікованих НММ - $\boldsymbol{v_{ver}}$, множина вагових коефіцієнтів для НММ, що входять до складу $\boldsymbol{v_{ver}}$ - $\boldsymbol{w}^{v_{ver}}$ та множина $\boldsymbol{\delta_{ver}}$, що містить величини похибок розпізнавання для $\boldsymbol{v_{ver}}$.

Якщо $\boldsymbol{v_{ver}} = \emptyset$, то реалізується перехід на третій етап, в протилежному випадку наступним виконується шостий етап.

Етап 6. Визначення найбільш ефективної архітектури. На вхід етапу

подаються множини $\boldsymbol{v_{ver}}$, $\boldsymbol{w^{v_{ver}}}$, $\boldsymbol{\delta_{ver}}$. Кроки етапу:

Крок 6.1. Розрахунок НММ з мінімальним значенням δ_{av}. Виконання кроку описується виразом:

$$v_{eff} = \min_{\delta_{av}}(\boldsymbol{v_{ver}}) , \tag{4.54}$$

де v_{eff} - найбільш ефективна НММ.

Крок 6.2. Уточнення архітектурних параметрів, що реалізується шляхом комп'ютерних експериментів, проведення яких регламентується [265, 307].

Виходом етапу являється v_{eff}, множина параметрів цієї НММ - $\boldsymbol{M^{v_{eff}}}$ та множина вагових коефіцієнтів її синаптичних зв'язків - $\boldsymbol{w^{v_{eff}}}$.

Етап 7. Застосування НММ для розпізнавання емоцій. На вхід етапу надходять v_{eff}, $\boldsymbol{M^{v_{eff}}}$, $\boldsymbol{w^{v_{eff}}}$, $\boldsymbol{BP^b}$. Етап розділяється на два кроки.

Крок 7.1. Формування вхідного поля НММ. Виконання даного кроку співвідноситься з виконання першого етапу даного методу з урахуванням необхідності формування вхідного поля на основі зареєстрованих показників БП.

Крок 7.2. Розпізнавання невідомого образу, що полягає у розрахунку вихідного сигналу НММ при подачі на її вхід параметрів, визначених на попередньому кроці даного етапу. Результатом виконання даного кроку являється $\boldsymbol{O^E}$ – множина розпізнаних емоцій слухача СДН.

Вказана множина $\boldsymbol{O^E}$ являється виходом етапу 7 та виходом представленого методу в цілому.

Таким чином, вдосконалено метод нейромережевого розпізнавання емоцій слухача СДН за одним із основних БП, що за рахунок застосування запропонованих рішень щодо формування вхідного поля та архітектури НММ для аналізу зображення обличчя, ГС та КП та рішень щодо відображення у вихідному полі НММ близькості еталонів класів дозволяє підвищити точність розпізнавання. Перспективи подальших досліджень полягають у розробці та застосуванні способів навчання НММ, адаптованих до паралельної реалізації.

4.4. Етапи реалізації методології автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання

Розроблена в даному розділі методологія являється інтеграцією запропонованих підходів, моделей та методів з відомими моделями та методами створення НМЗ для розпізнавання емоцій та особи людини на основі аналізу її БП. Вхідними даними методології є кортеж виду:

$$\boldsymbol{T^{in}} = \langle \boldsymbol{N^{BP}}, \langle \boldsymbol{BP^{s}} \rangle, \boldsymbol{E}, \boldsymbol{I}, \boldsymbol{V^{t}}, \boldsymbol{M^{V^{t}}}, \langle D^{r} \rangle, \langle D^{a} \rangle, \boldsymbol{\Theta^{p}}, \boldsymbol{\Theta^{p,a}}, \boldsymbol{\Theta^{hv}}, \boldsymbol{\rho}, \langle \boldsymbol{BP^{r}} \rangle \rangle, \quad (4.55)$$

де $\boldsymbol{N^{BP}}$ – перелік ідентифікаторів БП, що можуть бути зареєстровані;

$\langle \boldsymbol{BP^{s}} \rangle$ – кортеж, що містить збережені зареєстровані значення показників БП, які використовуються для побудови НММ;

$\boldsymbol{E}$ – множина емоцій, що мають бути розпізнані;

$\boldsymbol{I}$ – множина слухачів, особи яких мають бути розпізнані;

$\boldsymbol{V^{t}}$ – множина апробованих видів НММ;

$\boldsymbol{M^{V^{t}}}$ – множина параметрів апробованих видів НММ;

$\langle D^{r} \rangle$ – експертні дані, що стосуються реєстрації та обробки БП;

$\langle D^{a} \rangle$ – експертні дані, що стосуються НММ;

$\boldsymbol{\Theta^{p,m}}$ – множина показників основних БП;

$\boldsymbol{\Theta^{p,a}}$ – множина показників додаткових БП;

$\boldsymbol{\Theta^{hv}}$ – множина показників АПЗ, що використовується для реєстрації БП;

$\boldsymbol{\rho}$ – множина мінімально допустимих показників ефективності розпізнавання;

$\langle \boldsymbol{BP^{r}} \rangle$ – кортеж, що містить зареєстровані значення БП, що призначені для розпізнавання емоційного стану та особи слухача СДН.

Результат реалізації методології описується так:

$$\boldsymbol{M^{out}} = \langle \boldsymbol{O^{E}}, \boldsymbol{O^{ES}}, \boldsymbol{O^{I}} \rangle, \quad (4.56)$$

де $\boldsymbol{O^{E}}$ – результати розпізнавання емоцій слухача;

$\boldsymbol{O}^{ES}$ – результати розпізнавання поточного емоційного стану слухача;

$\boldsymbol{O}^{I}$ – результати розпізнавання особи слухача.

В узагальненому вигляді методологія включає наступні основні етапи:

1. Визначення множин БП, що підлягають реєстрації.
2. Формування множин параметрів, що підлягають нейромережевому аналізу.
3. Формування вимог до обробки зареєстрованих параметрів.
4. Формування вхідного поля НММ.
5. Формування навчальних прикладів.
6. Розпізнавання емоцій слухача.
7. Розпізнавання особи слухача.
8. Інтеграція результатів розпізнавання.

При розробці методології враховані дані [42, 163, 174, 177] та авторські дослідження [290, 292, 302] , що стосуються визначення множини сучасних видів НММ, які довели свою ефективність в задачах аналізу БП для розпізнавання емоцій та особи людини. В результаті сформована множина апробованих видів НММ:

$$\boldsymbol{V^t} = \{v^{TLP}, v^{MLP}, v^{CNN}, v^{LSTM}, v^{GRU}\}, \quad (4.57)$$

де v^{TLP} – ДШП;

v^{MLP} – БШП;

v^{CNN} – ЗНМ;

v^{LSTM} – мережа довгої короткострокової пам'яті;

v^{GRU} – мережа GRU.

Також на основі даних [243, 276, 293] та практичного досвіду щодо аналізу БП визначені критерії ефективності видів НММ, що входять до складу множини (4.55). До вказаних критеріїв відносяться:

- R_1^{NN}– можливість навчання на зашумлених даних;
- R_2^{NN} – обсяг пам'яті, необхідний для навчання НМ;
- R_3^{NN} – термін навчання мережі;

- R_4^{NN} – точність розпізнавання;

- R_5^{NN} – термін прийняття рішення;

- R_6^{NN} – пристосованість до аналізу динамічних рядів даних;

- R_7^{NN} – можливість інтерпретації виходу у вигляді ймовірності;

- R_8^{NN} – апробованість в задачах аналізу БП, асоційованих з зображеннями;

- R_9^{NN} – апробованість в задачах аналізу ГС;

- R_{10}^{NN} – апробованість в задачах аналізу КП.

Як і для випадку ЗНМ, передбачається, що кожен із критеріїв ефективності може прийняти значення в діапазоні від -1 до 1. Розраховані значення вказаних критеріїв для прийнятних видів НММ наведено в табл. 4.7. Розрахунок реалізовано за допомогою методу експертного оцінювання.

Таблиця 4.7

Значення критеріїв ефективності для прийнятних видів НММ

Критерій	**Вид НММ**				
	v^{TLP}	v^{MLP}	v^{CNN}	v^{LSTM}	v^{GRU}
R_1^{NN}	0,9	1	1	0.8	0,8
R_2^{NN}	0,5	0,7	1	0,2	0,1
R_3^{NN}	0,9	0,8	0	0,1	0,2
R_4^{NN}	1	1	1	0,8	0,8
R_5^{NN}	1	0,9	-0,5	-0,6	-0,5
R_6^{NN}	-1	-1	-1	0,9	0,7
R_7^{NN}	1	1	1	0,5	0,5
R_8^{NN}	0,7	0,8	1	-0,9	-0,9
R_9^{NN}	0,7	0,7	0,5	0,8	0,7
R_{10}^{NN}	0,5	0,8	0,5	0,5	0,5

Деталізація етапів методу здійснена з позицій відображення в ньому рішень,

що базуються на запропонованих підходах до нейромережевого аналізу основних та додаткових БП.

Етап 1. Визначення множин БП, що підлягають реєстрації. На вхід етапу подаються $\boldsymbol{N}^{BP}$, $\langle D^r \rangle$, $\langle D^a \rangle$, $\boldsymbol{\Theta}^{p,m}$, $\boldsymbol{\Theta}^{p,a}$, $\boldsymbol{\Theta}^{hv}$, $\langle \boldsymbol{BP}^{\boldsymbol{s}} \rangle$. Реалізація етапу 1 розділяється на три кроки.

Крок 1.1. Визначення переліку ідентифікаторів БП, що підлягають реєстрації. Визначення відбувається з використанням експертних даних, що стосуються задач управління учбовим процесом, які залежать від поточного емоційного стану та результатів розпізнавання особи слухача, можливостей АПЗ СДН, доступних БД навчальних прикладів БП та можливостей НМЗ розпізнавання емоцій та особи слухача. Виконання кроку 1.1 описується виразами виду:

$$\boldsymbol{N}_r^{BP,m} = f(\boldsymbol{N}^{BP,m}, \langle D^r \rangle, \boldsymbol{\Theta}^{p,m}, \boldsymbol{\Theta}^{hv}), \tag{4.58}$$

$$\boldsymbol{N}_r^{BP,a} = f(\boldsymbol{N}^{BP,a}, \langle D^r \rangle, \boldsymbol{\Theta}^{p,a}, \boldsymbol{\Theta}^{hv}), \tag{4.59}$$

$$\boldsymbol{N}_r^{BP} = \boldsymbol{N}_r^{BP,m} \cup \boldsymbol{N}_r^{BP,a}, \tag{4.60}$$

де $\boldsymbol{N}_r^{BP,m}, \boldsymbol{N}_r^{BP,a}$ – перелік основних та додаткових БП, що підлягають реєстрації.

Крок 1.2. Сепарація основних БП. Крок здійснюється для того, щоб на основі $\boldsymbol{N}_r^{BP,m}$ визначити $\boldsymbol{N}_r^{BP,E}$ - множину основних БП, що використовуються для розпізнавання емоцій та $\boldsymbol{N}_r^{BP,I}$ - множину основних БП, що використовуються для розпізнавання особи. Селекція основних БП реалізується за допомогою експертного оцінювання та описується виразом виду:

$$\left\{\boldsymbol{N}_r^{BP,E}, \boldsymbol{N}_r^{BP,I}\right\} = f\left(\boldsymbol{N}_r^{BP,m}, \langle D^r \rangle, \langle D^a \rangle, \boldsymbol{\Theta}^{p,m}, \boldsymbol{\Theta}^{hv}\right). \tag{4.61}$$

Крок 1.3. Визначення показників БП, що підлягають реєстрації. Визначення здійснюється за рахунок експертного оцінювання доцільності реєстрації показників БП, що входять до складу множин, визначених виразами (4.60, 4.61). В результаті виконання даного кроку формується $\boldsymbol{\Theta}_r^{p,m}$ - множина показників основних БП та $\boldsymbol{\Theta}_r^{p,a}$ - множина показників додаткових БП, що мають бути зареєстровані.

Виходом першого етапу являються множини $\boldsymbol{N}_r^{BP}$, $\boldsymbol{N}_r^{BP,E}$, $\boldsymbol{N}_r^{BP,I}$, компоненти яких визначаються виразами (4.60, 4.61), та множини $\boldsymbol{\Theta}_r^{p,m}$, $\boldsymbol{\Theta}_r^{p,a}$.

Етап 2. Формування множин параметрів для нейромережевого аналізу. На вхід етапу подаються $\boldsymbol{N}_r^{BP}$, $\boldsymbol{N}_r^{BP,E}$, $\boldsymbol{N}_r^{BP,I}$, $\langle D^r \rangle$, $\boldsymbol{\Theta}_r^{p,m}$, $\boldsymbol{\Theta}_r^{p,a}$, $\boldsymbol{\Theta}^{hv}$, $\langle \boldsymbol{BP}^s \rangle$. Передбачається, що множини значень параметрів, на основі яких реалізується розпізнавання емоцій слухача СДН, можуть бути отримані із доступних БД БП, а множини параметрів, на основі яких реалізується розпізнавання особи слухача, формуються в процесі введення вказаних слухачів в СДН. Реалізація етапу 2 розділяється на два кроки.

Крок 2.1. Формування кортежу значень показників основних БП для розпізнавання емоцій. Виконання кроку полягає в обробці доступних БД БП для отримання значень показників, що входять до складу множини $\boldsymbol{\Theta}_r^{p,m}$ та призначені для розпізнавання емоцій.

Крок 2.2. Формування кортежу значень показників БП для розпізнавання особи. Реалізація кроку співвідноситься з реєстрацією показників основних та додаткових БП, що входять до складу $\boldsymbol{\Theta}_r^{p,m}$ і $\boldsymbol{\Theta}_r^{p,a}$ та призначені для розпізнавання особи.

В результаті реалізації другого етапу формується кортеж виду:

$$\langle \boldsymbol{BP}_r \rangle = \langle \boldsymbol{BP}_r^{m,E}, \boldsymbol{BP}_r^{m,I}, \boldsymbol{BP}_r^{a,I} \rangle , \qquad (4.62)$$

де $\boldsymbol{BP}_r^{m,E}$ – множина показників основних БП, що використовуються для розпізнавання емоцій;

$\boldsymbol{BP}_r^{m,I}$ – множина показників основних БП, що використовуються для розпізнавання особи;

$\boldsymbol{BP}_r^{a,I}$ – множина показників додаткових БП.

Виходом другого етапу являється кортеж $\langle \boldsymbol{BP}_r \rangle$, компоненти якого визначаються виразом (4.62).

Етап 3. Формування вимог до обробки зареєстрованих параметрів. На

другому етапі на основі поданих на вхід етапу множин $\boldsymbol{N}_r^{BP}$, $\boldsymbol{N}_r^{BP,E}$, $\boldsymbol{N}_r^{BP,I}$,$\boldsymbol{\Theta}_r^{p,m}$, $\boldsymbol{\Theta}_r^{p,a}$, $\boldsymbol{V}^t$, $\boldsymbol{M}^{\boldsymbol{V}^t}$, $\boldsymbol{\Theta}^{hv}$ визначаються вимоги до результатів попередньої обробки БП. Визначення реалізується за допомогою методів експертного оцінювання. Етап реалізується за три кроки, перший виконується для БП, асоційованих з зображенням, другий – для ГС, третій – для КП.

Крок 3.1. Визначення множини вимог $\boldsymbol{U}_{img}$ до обробки зображень. Вказана множина описується так:

$$\boldsymbol{U}_{img} = \left\{\bigcup_{i=1}^{15} u_{img,i}\right\} = \{u_{img,1}, u_{img,2}, \dots, u_{img,15}\}, \tag{4.63}$$

де $u_{img,1}$ – наявність зовнішніх кордонів зображення;

$u_{img,2}$ – розміщення зображення в центрі вхідного поля ЗНМ;

$u_{img,3}$ – співпадіння вертикальної осі зображення з вертикаллю вхідного поля ЗНМ;

$u_{img,4}, u_{img,5}$ – максимально та мінімально допустима ширина зображення;

$u_{img,6}$, $u_{img,7}$ – максимально та мінімально допустима висота зображення;

$u_{img,8}$ – коефіцієнт допустимого викривлення при масштабуванні зображення;

$u_{img,9}$ – тип корекції яскравості пікселів;

$u_{img,10}$ – тип обробленого зображення (монохромне, напівтонове, кольорове);

$u_{img,11}$ – необхідність нівелювання завад, що закривають частину зображення;

$u_{img,12}$ – необхідність відновлення повернутого зображення;

$u_{img,13}$ – необхідність нормалізації параметрів зображення;

$u_{img,14}$ – шаблон визначення ключових точок;

$u_{img,15}$ – шаблон визначення контрольних точок.

Крок 3.2. Визначення множини вимог $\boldsymbol{U}_{ks}$ до обробки КП. Вказана множина описується так:

$$\boldsymbol{U}_{ks} = \left\{\bigcup_{i=1}^{6} u_{ks,i}\right\} = \{u_{ks,1}, u_{ks,2}, \dots, u_{ks,6}\}, \tag{4.64}$$

де $u_{ks,1}$– обмеження на символи, що можуть використовуватись при введені тексту;

$u_{ks,2}$– обмеження на тип клавіатури;

$u_{ks,3}$– обмеження на параметри ДКП, що підлягають реєстрації;

$u_{ks,4}$– мінімально допустимий обсяг тексту;

$u_{ks,5}$– ширина та висота вхідного поля ЗНМ;

$u_{ks,6}$– кількість растрів у вхідному полі ЗНМ.

Крок 3.3. Визначення множини вимог $\boldsymbol{U}_{v}$ до обробки ГС. Вказана множина описується так:

$$\boldsymbol{U}_{v} = \left\{\bigcup_{i=1}^{14} u_{vr,i}\right\} \cup \left\{\bigcup_{i=1}^{6} u_{vp,i}\right\} \cup \left\{\bigcup_{i=1}^{4} u_{NN,i}\right\} = \{u_{v,1}, u_{v,2}, \dots, u_{v,24}\}, \tag{4.65}$$

де $u_{v,1}$–$u_{v,14}$ – елементи множини $\boldsymbol{U}_{vr}$, задані виразом (3.74);

$u_{v,15}$–$u_{v,20}$ – елементи множини $\boldsymbol{U}_{vp,}$, задані виразом (3.77);

$u_{v,21}$–$u_{v,24}$ – елементи множини $\boldsymbol{U}_{NN}$, задані виразом (3.79).

Виходом етапу є множина вимог $\boldsymbol{U}_{bp}$ до реєстрації та обробки кожного із БП, що підлягають реєстрації.

Етап 4. Формування вхідного поля НММ. Вхідними даними являються множини $\boldsymbol{N}_r^{BP}$, $\boldsymbol{N}_r^{BP,E}$, $\boldsymbol{N}_r^{BP,I}$, $\boldsymbol{\Theta}_r^{p,m}$, $\boldsymbol{\Theta}_r^{p,a}$, $\boldsymbol{U}_{bp}$ та кортеж $\langle \boldsymbol{BP}_r \rangle$. Як і для попереднього, виконання даного етапу розділяється на три кроки, реалізація яких залежить від виду БП.

Крок 4.1. Реалізація методу формування вхідного поля НММ аналізу БП, асоційованих з зображеннями. Враховуються вимоги до обробки БП, задані

виразом (4.63). Узагальнена схема даного методу показана на рис. 3.6, а деталізована схема показана на рис. 3.7 та на рис.3.8, виконання методу описується виразами (3.23-3.53). У випадку наявності типових завад, викликаних перекриттям частини зображення та поворотом зображення, виконання п'ятого етапу методу формування вхідного поля полягає у реалізації запропонованого методу застосування вейвлет-перетворень для фільтрації вхідного поля (рис. 3.4), що описується виразами (3.1-3.22). Результатом виконання кроку 4.1 являється формування множини вхідних параметрів $\mathbf{\Psi}^{Im}$ для ЗНМ, що призначена для аналізу зображень.

Крок 4.2. Реалізація методу формування вхідного поля НММ аналізу КП. Математичне забезпечення цього методу описується виразами (3.54-3.72), структурна схема показана на рис. 3.16, а вимоги до обробки параметрів КП описуються виразом (4.64). Виходом кроку є множина $\mathbf{\Psi}^{KS}$, що визначає сформоване на основі параметрів КП вхідне поле ЗНМ, задане виразом (3.72).

Крок 4.3. Реалізація методу формування вхідного поля НММ аналізу ГС. Узагальнена схема даного методу показана на рис. 3.17, деталізована - на рис. 3.22 та рис.3.23. Математичне забезпечення методу складають вирази (3.73-3.127), а вимоги до обробки ГС задані виразом (4.65). Особливістю методу являється можливість формування вхідного поля у вигляді вектору, придатного для подання в НММ з лінійним типом вхідного поля та можливість формування вхідного поля у вигляді прямокутника, придатного для подання в ЗНМ. Результатом реалізації кроку 3.3 являється множина $\mathbf{\Psi}^{V}$, що містить значення параметрів вхідного поля НММ, яка призначена для аналізу ГС та описується виразами (3.81, 3.82).

Виходом четвертого етапу являється множина $\mathbf{\Psi}^{BP} = \{\mathbf{\Psi}^{Im}, \mathbf{\Psi}^{KS}, \mathbf{\Psi}^{V}\}$, елементи якої містять значення параметрів вхідного поля НММ для кожного виду БП, що підлягають нейромережевому аналізу.

Етап 5. Формування навчальних прикладів. Виконання даного етапу

співвідноситься з реалізацією запропонованого методу формування навчальних прикладів для НММ аналізу БП, схема якого показана на рис. 4.5, а математичне забезпечення складають вирази (4.1-4.9, 4.14-4.32). Особливістю даного методу являється врахування у вихідному сигналі навчальних прикладів близькості еталонів класів, що мають бути розпізнані. Також при реалізації етапу врахована можливість створення навчальних прикладів з використанням експертних знань та можливість спрощеного визначення вихідного сигналу. На вхід етапу подаються множини $\boldsymbol{\Psi}^{BP}$, $\boldsymbol{E}$, $\boldsymbol{I}$, $\boldsymbol{M}^{\boldsymbol{V^t}}$, $\boldsymbol{V}^t$ та кортеж $\langle D^r \rangle$. Виконання етапу розділяється на три кроки.

Крок 5.1. Визначення параметрів навчальних прикладів з урахуванням близькості еталонів, що полягає у реалізації 2-5 етапів методу формування навчальних прикладів.

Крок 5.2. Визначення параметрів навчальних прикладів з one-hot кодуванням вихідного сигналу. У випадку виконання даного кроку реалізується лише 5 етап означеного методу формування навчальних прикладів. При цьому очікуваний вихідний сигнал визначається за допомогою загальновідомого методу one-hot кодування.

Крок 5.3. Визначення параметрів навчальних прикладів за допомогою експертних правил, що визначаються виразом виду:

$$If\ (x_1 = const_1 \wedge x_2 = const_2 \wedge \ldots \wedge x_K = const_K) \rightarrow Y_m\ , \quad (4.66)$$

де $x_1, x_2, \ldots, x_K$ – параметри вхідного поля НММ;

$const_1, const_2, \ldots, const_K$ – константи, визначені шляхом експертного оцінювання БП;

Y_m – очікуваний вихідний сигнал НММ, що асоціюється з m-ою емоцією або m-ою особою слухача СДН.

Виконання кроку 5.3 базується на результатах [102, 103, 244, 304].

Виходом п’ятого етапу являється кортеж параметрів навчальних прикладів

$\langle \boldsymbol{X}, \boldsymbol{Y} \rangle$, що описується виразом (4.32).

Етап 6. Розпізнавання емоцій слухача. Реалізація етапу співвідноситься з виконанням методу нейромережевого розпізнавання емоцій та методу розробки архітектури ЗНМ, призначеної для аналізу БП, що описуються виразами (4.46-4.54) та (4.33-4.45). Схеми виконання вказаних методів показано на рис. 4.8-4.8 та на рис. 4.5-4.7. Вхідними даними етапу являються $\boldsymbol{N}_r^{BP,E}$, $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle$, $\boldsymbol{E}$, $\boldsymbol{V}^t$, $\boldsymbol{M}^{\boldsymbol{V}^t}$, $\langle D^a \rangle$, $\boldsymbol{\rho}$, а його виконання розділяється на чотири кроки, що виконуються для кожного із основних БП, які входять до складу $\boldsymbol{N}_r^{BP,E}$.

Крок 6.1. Визначення ефективних нейромережевих архітектур для розпізнавання емоцій. Виконання кроку співвідноситься з реалізацією 2-4 кроків третього етапу методу нейромережевого розпізнавання емоцій.

Крок 6.2. Реалізація методу розробки архітектури ЗНМ розпізнавання емоцій. Даний крок виконується лише у випадку, якщо на попередньому кроці ЗНМ була включена до складу ефективних видів НММ.

Крок 6.3. Визначення архітектурних параметрів НММ розпізнавання емоцій. Виконання цього кроку полягає у реалізації 4-6 етапів методу нейромережевого розпізнавання емоцій.

Крок 6.4. Розпізнавання емоцій. Виконання кроку полягає у реалізації сьомого етапу методу нейромережевого розпізнавання емоцій.

Виходом етапу являється $\boldsymbol{v}_{eff}^E$ – множина найбільш ефективних нейромережевих архітектур для розпізнавання емоцій на основі аналізу кожного елементу $\boldsymbol{N}_r^{BP,E}$, $\boldsymbol{M}^{\boldsymbol{v}_{eff}^E}$ – множина параметрів вказаних архітектур, $\boldsymbol{w}^{\boldsymbol{v}_{eff}^E}$ – множина вагових коефіцієнтів синаптичних зв'язків для $\boldsymbol{v}_{eff}^E$ та $\boldsymbol{Y}^{\boldsymbol{E}}$ – множина, що містить результат нейромережевого розпізнавання емоцій на основі аналізу кожного елементу $\boldsymbol{N}_r^{BP,E}$.

Етап 7. Розпізнавання особи слухача. Даний етап являється аналогією попереднього етапу з урахуванням відмінностей нейромережевого розпізнавання

особи слухача від нейромережевого розпізнавання емоцій. Так, на вхід цього етапу замість множин $\boldsymbol{N}_r^{BP,E}$ та $\boldsymbol{E}$ подаються множини $\boldsymbol{N}_r^{BP,I}$, $\boldsymbol{N}_r^{BP,a}$ та $\boldsymbol{I}$. Крім того, вхідними даними етапу 7 являються $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle$, $\boldsymbol{V}^t$, $\boldsymbol{M}^{\boldsymbol{V}^t}$, $\langle D^a \rangle$, $\boldsymbol{\rho}$. Як і попередній, етап 7 розділяється на чотири кроки.

Крок 7.1. Визначення ефективних нейромережевих архітектур для розпізнавання особи. Виконання кроку співвідноситься з реалізацією 2-4 кроків третього етапу методу розпізнавання емоцій з урахуванням необхідності нейромережевого аналізу $\boldsymbol{N}_r^{BP,I}$ та $\boldsymbol{N}_r^{BP,a}$.

Крок 7.2. Реалізація методу розробки архітектури ЗНМ розпізнавання особи. Даний крок виконується у випадку, якщо ЗНМ входить до множини ефективних видів НММ, призначених для розпізнавання особи.

Крок 7.3. Визначення архітектурних параметрів НММ розпізнавання особи. Виконання цього кроку полягає у реалізації 4-6 етапів методу нейромережевого розпізнавання емоцій з урахуванням того, що навчання НММ реалізується з використанням навчальних прикладів, в яких вихідний сигнал відображає результат розпізнавання особи, а вхідне поле може формуватися на основі контрольних точок зображень, асоційованих з БП.

Крок 7.4. Розпізнавання особи. Виконання кроку полягає у реалізації сьомого етапу методу нейромережевого розпізнавання емоцій з урахуванням формування вхідного поля НММ на основі $\boldsymbol{N}_r^{BP,I}$ та $\boldsymbol{N}_r^{BP,a}$.

Виходом етапу являється $\boldsymbol{v}_{eff}^I$ – множина найбільш ефективних нейромережевих архітектур для розпізнавання особи на основі аналізу кожного елементу $\boldsymbol{N}_r^{BP,I}$ та $\boldsymbol{N}_r^{BP,a}$, $\boldsymbol{M}^{\boldsymbol{v}_{eff}^I}$ - множина параметрів вказаних архітектур, $\boldsymbol{w}^{\boldsymbol{v}_{eff}^I}$ - множина вагових коефіцієнтів синаптичних зв'язків для $\boldsymbol{v}_{eff}^I$ та $\boldsymbol{Y}^I$ - множина, що містить результат нейромережевого розпізнавання особи на основі аналізу кожного елементу $\boldsymbol{N}_r^{BP,I}$ та $\boldsymbol{N}_r^{BP,a}$.

Етап 8. Інтеграція результатів розпізнавання. Метою виконання даного етапу

являється інтеграція результатів розпізнавання особи і емоцій, що реалізовані на основі аналізу кожного із зареєстрованих БП, та розпізнавання емоційного стану на основі інтеграції результатів розпізнавання емоцій по кожному із $\boldsymbol{N}_r^{BP,E}$. На вхід етапу подаються $\boldsymbol{N}_r^{BP,E}$, $\boldsymbol{N}_r^{BP,I}$, $\boldsymbol{N}_r^{BP,a}$ та $\boldsymbol{Y^E}$, $\boldsymbol{Y}^I$ - результати розпізнавання емоцій та особи на протязі певного періоду часу $T = \{t_1, t_2, \dots, t_K\}$. У базовому варіанті передбачається, що правило прийняття остаточного (інтегрованого) рішення про результати нейромережевого аналізу може бути сформульовано так:

$$Y_R = max(Y_1, Y_2, \dots Y_N), \tag{4.67}$$

$$Y_n = \frac{1}{J}\sum_{j=1}^{J}(\alpha_j \times y_j), \tag{4.68}$$

де Y_R – вихідний сигнал, що свідчить про результати інтегрованого нейромережевого аналізу всіх зареєстрованих БП;

Y_n – оцінка n-го альтернативного рішення;

N – кількість можливих альтернативних рішень;

α_j – ваговий коефіцієнт j-го параметру;

J – кількість параметрів, що підлягають обліку;

y_j – результат нейромережевого аналізу j-го параметру.

Значення вагових коефіцієнтів, що використовуються в (4.68), можуть визначатися шляхом експертного оцінювання.

Надалі правило прийняття рішення виду (4.67, 4.68) можливо вдосконалити за рахунок використання однієї або декількох НММ, призначених для узагальнення результатів нейромережевого аналізу перерахованих вище БП.

На вхід такої мережі можуть подаватись вихідні сигнали про результати розпізнавання особи, емоцій та емоційного стану. Вихід мережі асоціюється з кортежем виду:

$$Y_R = \langle \{K_I\}, \{K_E\}, \{K_{ES}\} \rangle, \tag{4.69}$$

де $\{K_I\}, \{K_E\}, \{K_{ES}\}$ – множина вихідних сигналів НММ, що свідчать про

приналежність особи, емоцій та емоційного стану слухача до одного із заздалегідь визначених класів.

Етап розділяється на три кроки.

Крок 8.1. Інтегроване розпізнавання особи. Крок призначено для розпізнавання особи слухача СДН на основі інтеграції результатів розпізнавання слухача, що були отримані на основі аналізу $\boldsymbol{N}_r^{BP,I}$ та $\boldsymbol{N}_r^{BP,a}$. Оскільки для аналізу різних БП передбачено використання НММ з однаковою кількістю вихідних сигналів, то видозмінений вираз (4.68) для інтеграції вихідних сигналів вказаних НММ записується так:

$$O_{k,\Sigma}^{I} = \frac{1}{N}\sum_{n=1}^{N}\left(\alpha_n^I \times y_{k,n}^I\right), \tag{4.70}$$

де $O_{k,\Sigma}^{I}$ – інтегроване значення k-го вихідного сигналу НММ для розпізнавання особи слухача СДН;

$y_{k,n}^I$ – значення k-го вихідного сигналу НММ при розпізнаванні особи слухача СДН на основі n-го БП;

N – кількість БП, що використовуються для розпізнавання особи;

α_n^I – ваговий коефіцієнт n-го БП при розпізнаванні особи слухача, що визначається методом експертних оцінок.

Висновок про остаточне розпізнавання особи слухача СДН реалізується за допомогою (4.67). Якщо особа слухача не може бути розпізнана, то передбачено формування відповідного повідомлення та припинення процесу розпізнавання. Неможливість розпізнавання формується у випадку правдивості виразу (4.71).

$$Y_R \le 1{,}1K^{-1}, \tag{4.71}$$

де K – кількість слухачів СДН.

Крок 8.2. Інтегроване розпізнавання емоцій. Крок призначено для розпізнавання емоцій слухача СДН на основі інтеграції результатів розпізнавання емоцій, що були отримані на основі аналізу $\boldsymbol{N}_r^{BP,E}$. Розрахунок інтегрованого

значення k-го вихідного сигналу НММ при розпізнаванні емоцій описується так:

$$O_{k,\Sigma}^{E} = \frac{1}{N}\sum_{n=1}^{N}\left(\alpha_n^E \times y_{k,n}^E\right), \tag{4.72}$$

де $O_{k,\Sigma}^{E}$ – інтегроване значення k-го вихідного сигналу НММ при розпізнаванні емоцій слухача СДН;

$y_{k,n}^E$ – значення k-го вихідного сигналу НММ при розпізнаванні емоцій слухача СДН на основі n-го БП;

N – кількість БП, що використовуються для розпізнавання емоцій;

α_n^E – ваговий коефіцієнт n-го БП при розпізнаванні емоцій.

За необхідності, домінуюча емоція визначається за допомогою (4.67). Результат розпізнавання емоцій слухача СДН описується так:

$$\boldsymbol{O}^{E} = \left\{O_{k,\Sigma}^{E}\right\}_{K}. \tag{4.73}$$

Крок 8.3. Розпізнавання емоційного стану. Використавши наведений в [48, 166, 308] постулат про те, що емоційний стан являється узагальненням емоцій на протязі певного інтервалу часу, реалізація даного кроку описана виразом виду:

$$\bar{O}_k^{Es} = \frac{1}{N}\sum_{n=1}^{N} O_{k,\Sigma}^{E}(t_n), \tag{4.74}$$

$$\boldsymbol{O}^{Es} = \{\bar{O}_k^{Es}\}_K, \tag{4.75}$$

де $\bar{O}_{k,\Sigma}^{E}(t_n)$ – результат інтегрованого розпізнавання k-ої емоції на t_n-му часовому інтервалі;

N – кількість піддослідних часових інтервалів аналізу БП.

В базовому випадку тривалість одного інтервалу аналізу БП становить $t_{rg} = 1{,}5$ с, а $N = 20$. Зазначимо, що для зображення обличчя вираз (4.74) модифікується з урахуванням кількості окремих реєстрацій такого зображення на t_{rg}. Виходом даного етапу являються множини $\boldsymbol{O}^{E}, \boldsymbol{O}^{ES}, \boldsymbol{O}^{I}$, що відповідно до (4.56), визначають кортеж $\boldsymbol{M}^{\boldsymbol{out}}$, що являється результатом реалізації запропонованої методології,

структурно-аналітичне відображення якої показано на рис. 4.10 та рис. 4.11.

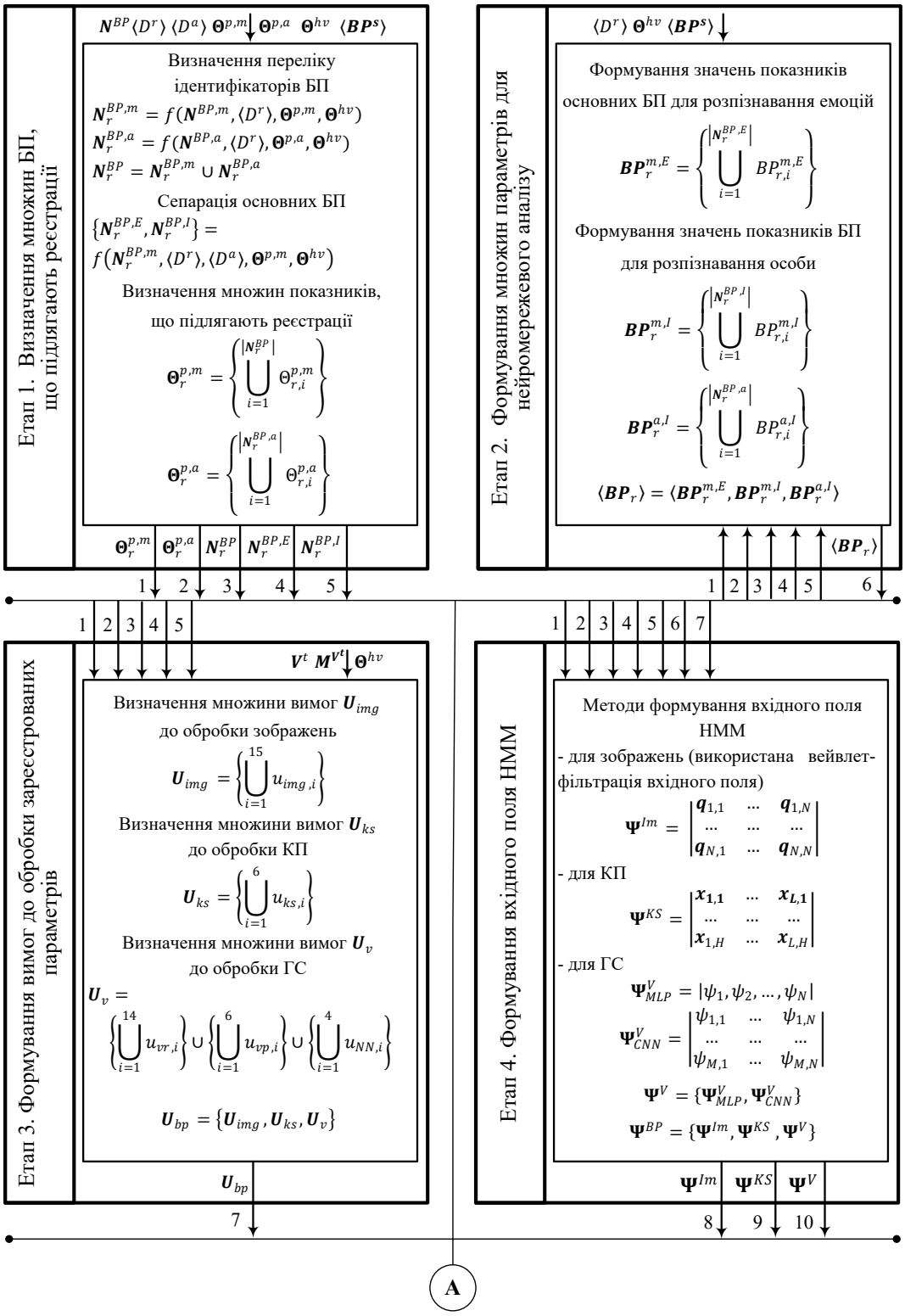

Рис. 4.10. Структурно-аналітичне відображення етапів 1-4 методології автоматизованого розпізнавання емоційного стану слухача СДН

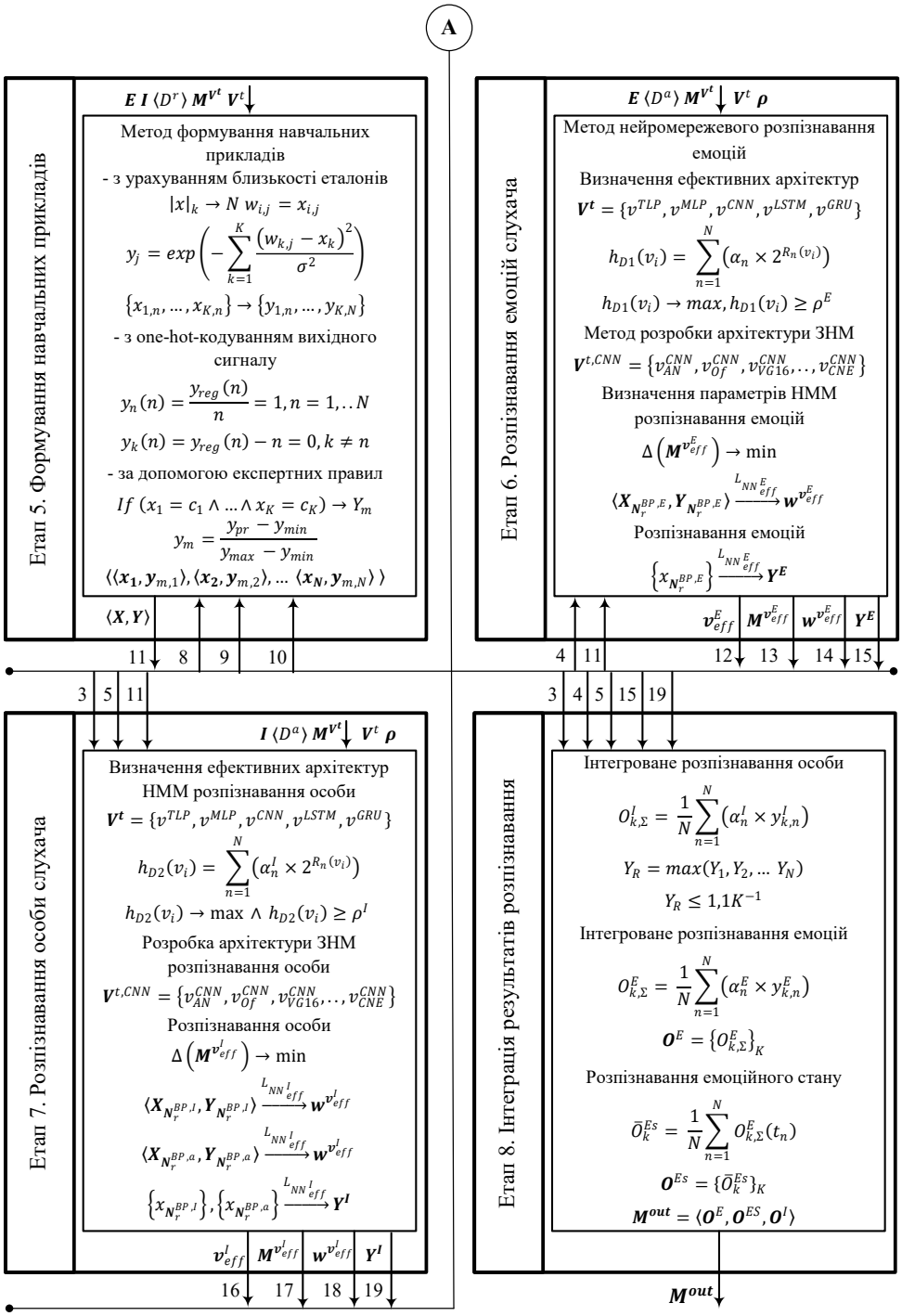

Рис. 4.11. Структурно-аналітичне відображення етапів 5-8 методології автоматизованого розпізнавання емоційного стану слухача СДН

Схема запропонованої методології, що відображає зв’язки між базовими та розробленими підходами, моделями та методами, показана на рис. 4.12.

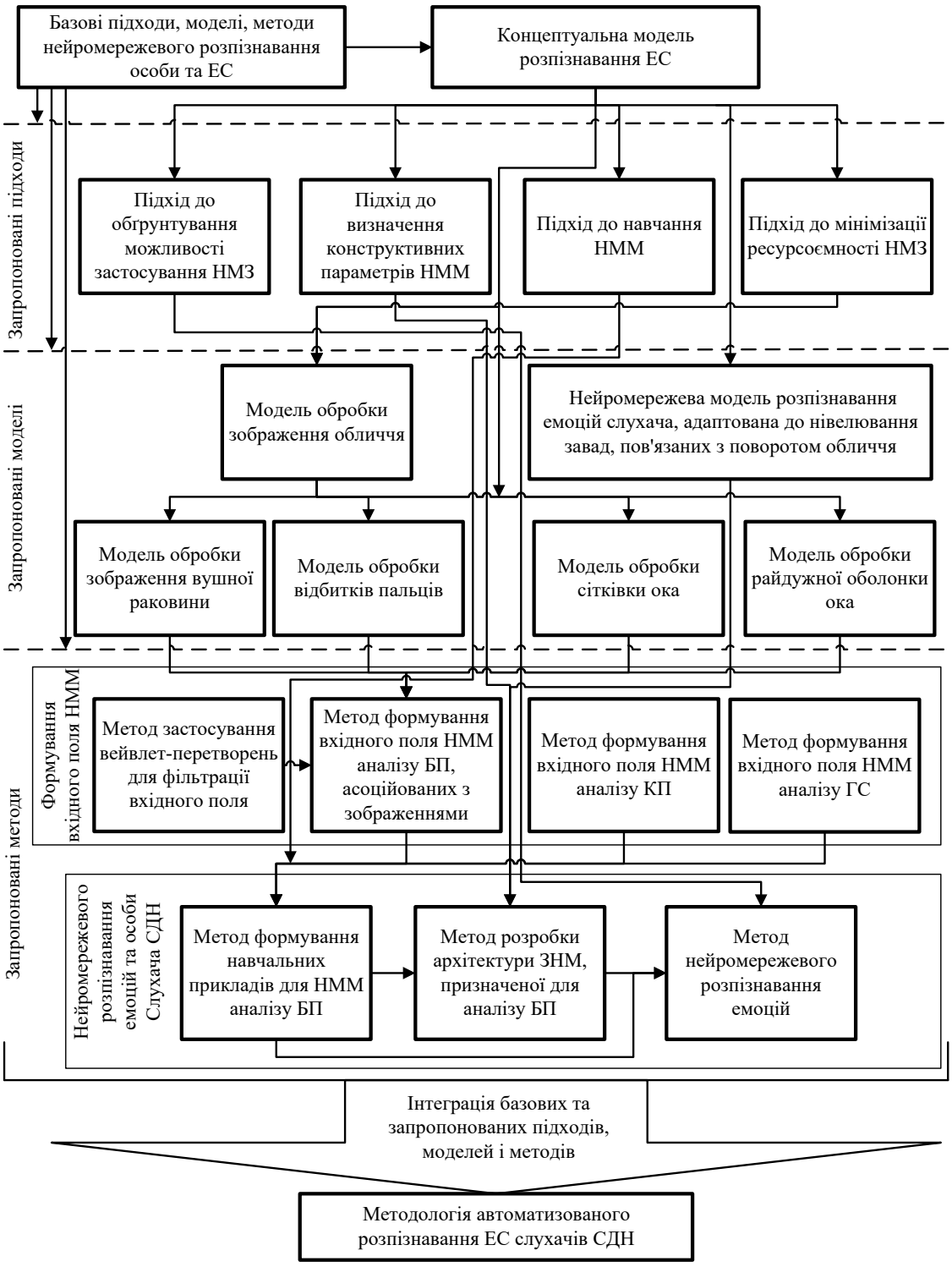

Рис. 4.12. Схема методології автоматизованого розпізнавання ЕС слухачів СДН

В результаті проведених досліджень вперше розроблено методологію автоматизованого розпізнавання емоційного стану слухача СДН, що за рахунок

формалізованого механізму використання розроблених методів формування вхідного поля НММ аналізу ГС, КП, БП, асоційованих з зображеннями, методів застосування вейвлет-перетворень для фільтрації вхідного поля, формування навчальних прикладів, розробки архітектури ЗНМ та методу нейромережевого розпізнавання емоцій дозволяє розширити функціональні можливості СДН, обрати найбільш ефективний нейромережевий засіб та забезпечити підтримку процесів створення систем нейромережевого розпізнавання ЕС та особи слухача. Перспективи розвитку даної методології пов'язані із вдосконаленням процедур інтеграції результатів розпізнавання ЕС та особи слухача СДН за сукупністю БП.

4.5. Висновки до четвертого розділу

В даному розділі вирішувались наукові завдання розвитку методів та побудови методології нейромережевого розпізнавання емоційного стану та особи слухача системи дистанційного навчання. Основні результати розділу наступні:

1. Отримав подальший розвиток метод формування навчальних прикладів для НММ аналізу БП, який за рахунок обробки навчальних прикладів мережею PNN дозволяє врахувати в очікуваному вихідному сигналі близькість еталонів класів, котрі мають бути розпізнані, що забезпечує можливість приблизно в 1,3-1,5 рази зменшити кількість навчальних ітерацій для досягнення помилки навчання в межах 1% відносно випадку застосування прикладів, в яких використана загальновідома процедура кодування очікуваного вихідного сигналу.

2. Удосконалено метод розробки архітектури ЗНМ, призначеної для аналізу БП слухача СДН, що за рахунок запропонованих критеріїв ефективності виду ЗНМ та запропонованих принципів адаптації конструктивних параметрів ЗНМ до умов функціонування, забезпечує можливість зменшення обсягу експериментальних досліджень, пов'язаних з розробкою архітектури ЗНМ.

3. Удосконалено метод нейромережевого розпізнавання емоцій слухача СДН за одним із основних БП, що на відміну від відомих за рахунок застосування запропонованих рішень щодо формування вхідного поля НММ для аналізу зображення обличчя, ГС та КП та рішень щодо відображення у вихідному полі НММ близькості еталонів класів, які мають бути розпізнані, дозволяє підвищити точність розпізнавання.

4. В результаті проведених досліджень вперше розроблено методологію автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання, що за рахунок формалізованого механізму використання розроблених методів формування вхідного поля НММ аналізу ГС, КП, БП, асоційованих з зображеннями, методів застосування вейвлет-перетворень для фільтрації вхідного поля, формування навчальних прикладів, розробки архітектури ЗНМ та методу нейромережевого розпізнавання емоцій дозволяє розширити функціональні можливості СДН, обрати найбільш ефективний нейромережевий засіб та забезпечити підтримку процесів створення систем нейромережевого розпізнавання емоційного стану та особи слухача.

5. Перспективи подальших досліджень пов’язані з підвищенням ефективності запропонованих методів та методології за рахунок застосування в них процедур:

- Модифікації архітектурних параметрів ЗНМ за рахунок використання сучасних досягнень в області НМ.

- Навчання НММ, адаптованих до паралельної реалізації.

- Інтеграції результатів розпізнавання емоційного стану та особи слухача СДН за сукупністю БП.

РОЗДІЛ 5
СИСТЕМА АВТОМАТИЗОВАНОГО РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО СТАНУ ТА ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

5.1. Архітектура системи автоматизованого розпізнавання емоційного стану

Відповідно до загальноприйнятої технології розробки ІС, для опису архітектури CAPEC слухачів СДН використано мову моделювання Unified Modeling Language. Вибір вказаної мови моделювання пояснюється її універсальністю, апробованістю та доступністю засобів автоматизації процесу проєктування. Зазначимо, що для опису архітектури був використаний сервіс Rational Rose виробництва компанії IBM.

Відповідно [97, 117], в першу чергу було вирішено завдання побудови діаграм варіантів використання, які необхідні для виявлення загальних особливостей поведінки системи, що проєктується. Основні елементи діаграми - актори і варіанти використання. Актором може бути як людина, що взаємодіє з системою, так і зовнішня система, пристрій, програма, яка здійснює обмін інформацією з системою. Кожен варіант використання відповідає окремому сервісу, який надається за запитом актора. Коли система закінчує обробку запиту від актора, вона повертається в початковий стан і готова до наступних запитів. Актор завжди знаходиться поза розроблюваної системи і взаємодіє з нею за допомогою варіантів використання. Один актор може виконувати кілька варіантів використання. У свою чергу у варіанті використання може бути кілька акторів.

Завданнями CAPEC є:

- розпізнавання особи слухача при вході в СДН;

- розпізнавання емоційного стану слухача в процесі його взаємодії з СДН при навчанні;

- розпізнавання особи слухача в процесі його взаємодії з СДН при навчанні.

Орієнтуючись на зазначені завдання, побудована діаграма варіантів використання CAPEC при розпізнаванні особи слухача при вході в СДН, а також діаграма, що відповідає розпізнаванню особи і емоцій слухача при навчанні. У побудованих діаграмах відображені зв'язки (відношення) між варіантами використання і акторами, а також зв'язки між декількома варіантами використання. Використано відношення асоціації, включення і узагальнення.

Відношення асоціації, що виконує комунікативну функцію, сигналізує про взаємодію суб'єкта з системою в рамках певного варіанту використання. Напрямок зв'язку вказує, хто є ініціатором взаємодії, суб'єкт або система. Відношення включення ("include") застосовувалося на діаграмах, коли поведінка одного варіанту використання включається як складовий компонент в послідовність поведінки іншого. Відношення узагальнення використано для вказівки факту, що деякий варіант використання A може бути узагальнено до варіанту використання B. Таким чином, цей показник вказує, що дочірні варіанти використання мають атрибути і особливості поведінки батьківських варіантів використання. При цьому дочірні варіанти використання можуть мати нові властивості поведінки.

Перша з розроблених діаграм показана на рис. 5.1. До складу діаграми входять 3 актора (Student, Defender, Admin) і 16 варіантів використання.

Актор Student – слухач СДН, чиї основні та допоміжні БП, зображення обличчя, голос, клавіатурний почерк, райдужна оболонка ока, зображення вушної раковини, відбитки пальців, сітківка ока, підлягають нейромережевому аналізу з метою визначення особи (ідентифікації) слухача при вході в систему.

Актор Admin представляє собою адміністратора представленої системи автоматизованого розпізнавання емоційного стану та особи слухача СДН. Даний актор відповідає за налаштування параметрів системи при нейромережевому розпізнаванні особи слухача при вході в СДН.

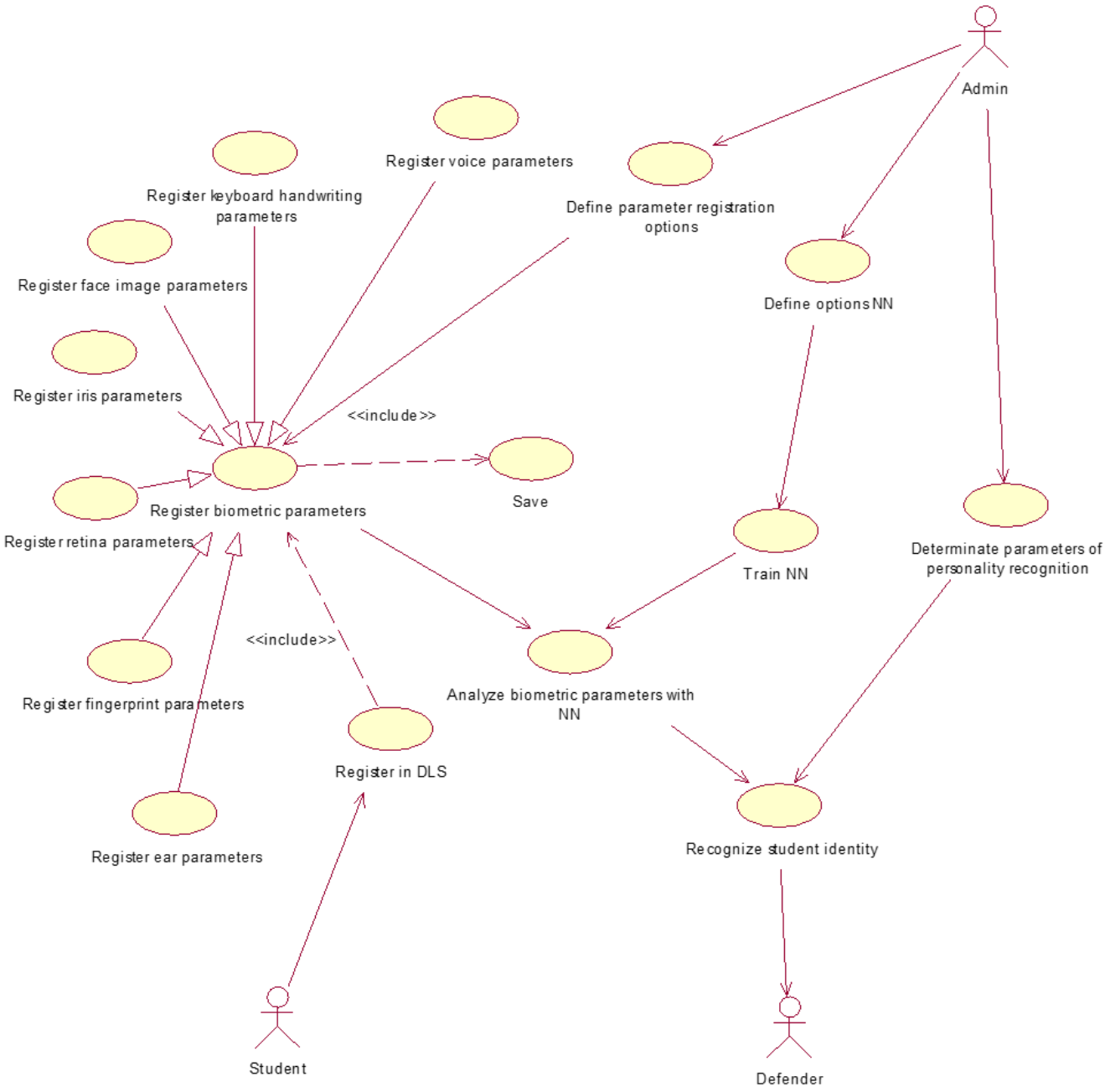

Рис. 5.1. Діаграма варіантів використання системи автоматизованого розпізнавання емоційного стану при аутентифікації слухача при вході в СДН

Актор Defender співвідноситься з блоком захисту СДН. За його запитом представлена система надає сервіс – нейромережевий аналіз особи слухача за його БП. За результатами даного запиту актор Defender приймає рішення щодо необхідних захисних заходів та реалізує ці заходи.

До вказаних заходів відносяться: дозвіл на вхід в СДН, запит на повторне введення БП, повідомлення користувачеві про заборону доступу до системи, заборона доступу, блокування системи вводу БП. Вказані заходи можуть бути реалізовані ізольовано або в комплексі.

Номенклатуру можливих захисних заходів та методи їх реалізації можливо визначити на основі даних [3, 42].

Відповідність варіантів використання функціям CAPEC, що використовується при вході слухача в СДН, наведені в табл. 5.1.

Слід зазначити, що перераховані в табл. 5.1 варіанти використання відповідають розширеним можливостям CAPEC, які стосуються розпізнавання особи слухача при вході в СДН. На практиці частина функцій може не використовуватися.

Також зауважимо, що варіант використання *Register biometric parameters* є батьківським відносно варіантів використання, що описують функції CAPEC, які стосуються реєстрації таких БП, як зображення обличчя, КП, зображення вушної раковини, голос, сітківка ока, відбитки пальців та райдужна оболонка ока.

Таблиця 5.1

Опис варіантів використання нейромережевого аналізатора

Варіант використання	**Функції нейромережевого аналізатора**
1	2
Register in DLS	Реєстрація слухача при вході в СДН
Register biometric parameters	Реєстрація біометричних параметрів слухача
Register ear parameters	Реєстрація зображення вушної раковини
Register fingerprint parameters	Реєстрація відбитків пальців
Register retina parameters	Реєстрація сітківки ока
Register face image parameters	Реєстрація зображення обличчя

Таблиця 5.1 (продовження)

1	2
Register iris parameters	Реєстрація райдужної оболонки ока
Register keyboard handwriting parameters	Реєстрація параметрів клавіатурного почерку
Register voice parameters	Реєстрація параметрів голосового сигналу
Save	Збереження зареєстрованих біометричних параметрів
Define parameter registration options	Визначення опцій реєстрації біометричних параметрів
Define options NN	Визначення значень конструктивних параметрів архітектури нейронної мережі
Train NN	Навчання нейронної мережі
Analyze biometric parameters with NN	Нейромережевий аналіз біометричних параметрів
Recognize student identity	Розпізнавання особи слухача СДН
Determinate parameters of personality recognition	Визначення параметрів розпізнавання особи слухача СДН

Діаграма варіантів використання запропонованої CAPEC слухача СДН в процесі навчання показана на рис. 5.2. Відносно діаграми використання, що показана на рис. 5.1, особливості даної діаграми визначаються відмінностями завдання розпізнавання особи і емоційного стану слухача в процесі навчання від завдання розпізнавання особи слухача при вході в СДН.

Базуючись на запропонованій методології автоматизованого розпізнавання, передбачено, що в процесі навчання як розпізнавання особи, так і розпізнавання емоційного стану слухача СДН здійснюється на підставі нейромережевого аналізу

всіх трьох основних БП (КП, зображення обличчя та ГС), що можуть бути зареєстровані за допомогою розповсюдженого АПЗ.

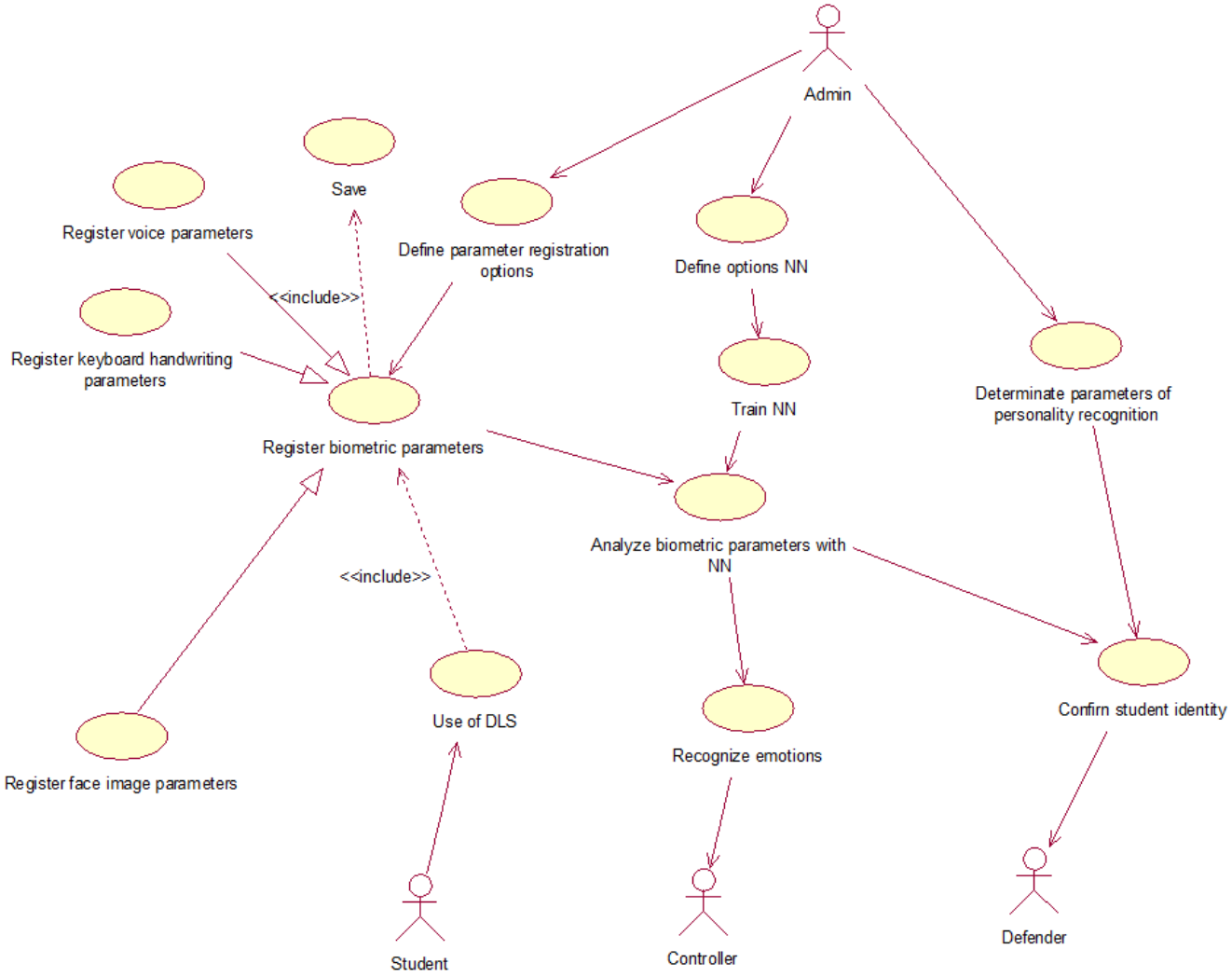

Рис. 5.2. Діаграма варіантів використання системи автоматизованого розпізнавання емоційного стану слухача в процесі навчання

Це продиктовано прагненням зменшити ресурсоємність САРЕС шляхом зменшення обсягу параметрів, що підлягають аналізу.

Прийнято до уваги, що згідно з [3, 42, 131], найбільш відповідальний етап розпізнавання особи слухача, реалізується при його вході в СДН. Тому в процесі навчання в сенсі захисту інформації завдання нейромережевого аналізу зводиться

виключно до підтвердження особи слухача. Як наслідок, для розпізнавання особи слухача в процесі навчання використовуються тільки ті БП, які одночасно використовуються і для розпізнавання емоцій.

Разом з тим, необхідність розпізнавання емоційного стану слухача визначила включення в діаграму актора Controller. Передбачається, що актор Controller представляє собою модуль СДН, що призначений для корекції виду і обсягу поданих навчальних матеріалів в залежності від розпізнаного поточного емоційного стану слухача.

Таким чином, представлена на рис. 5.2 діаграма включає в себе 13 варіантів використання. Особливості варіантів використання CAPEC що використовується для розпізнавання особи і емоцій слухача в процесі навчання полягають в наступному:

- Use of DLS – використання слухача СДН в процесі навчання.

- Recognize emotions – розпізнавання емоційного стану слухача.

- Confirm student identity – підтвердження особи слухача.

Базуючись на розроблених діаграмах варіантів використання CAPEC при розпізнаванні особи слухача при вході в СДН та в процесі навчання (рис. 5.1 і рис. 5.2), побудована показана на рис. 5.3 структурна схема CAPEC та особи слухача. Відповідно до опису акторів діаграм варіантів використання CAPEC, передбачено, що вхідна інформація включає в себе:

- ***УЕ*** – множину параметрів, які характеризують умови експлуатації.

- ***БП*** – множину БП, що підлягають реєстрації.

- ***УПР*** – множину параметрів, що характеризують умови перемикання режимів функціонування CAPEC.

Вихідна інформація CAPEC складається з:

- ***I*** – множини параметрів, що описують результати розпізнавання особи слухача.

- ***E*** – множини параметрів, що описують результати розпізнавання емоційного стану слухача.

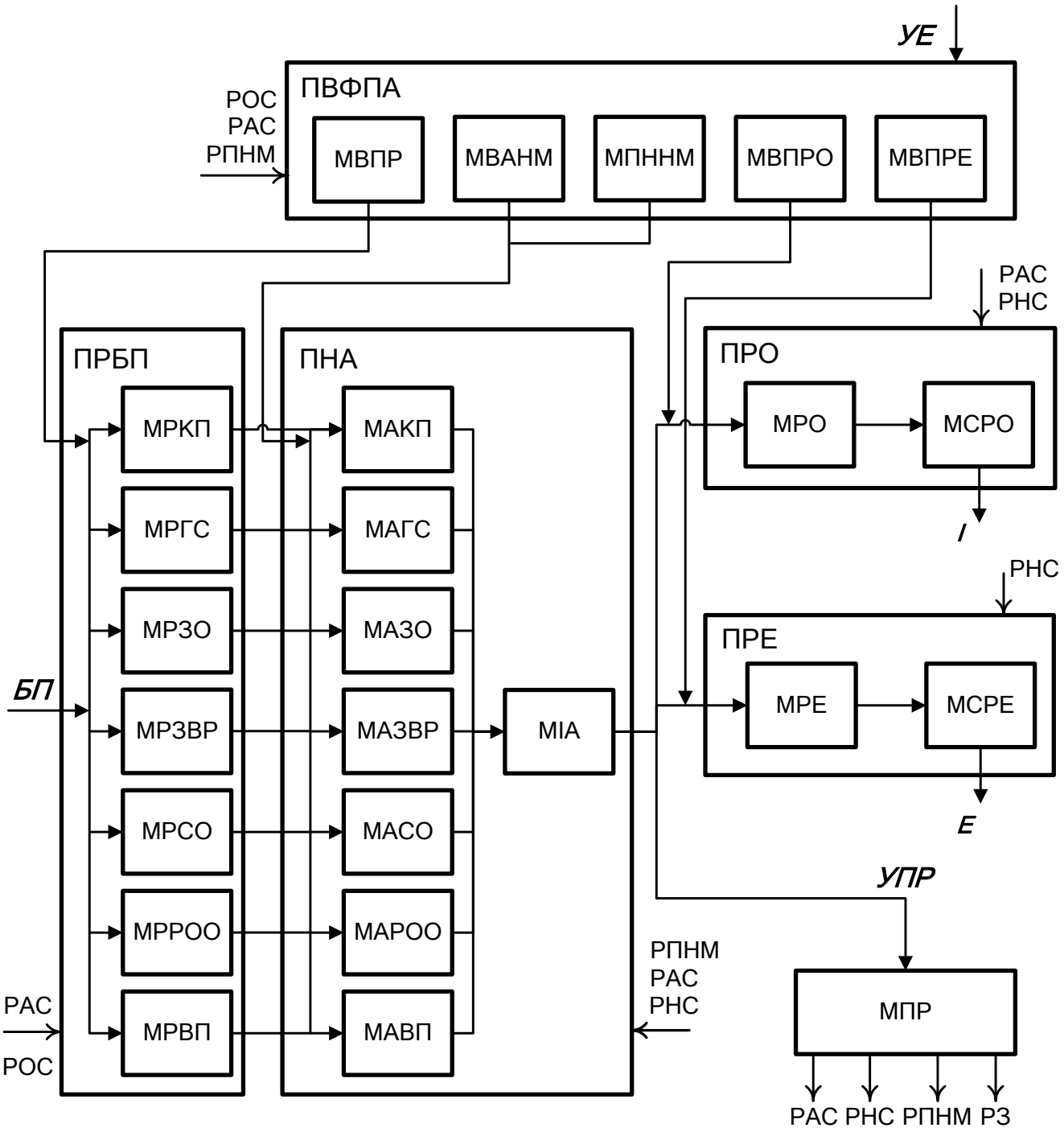

Рис. 5.3. Структура системи автоматизованого розпізнавання емоційного стану слухача СДН

Відповідно до наведених на рис. 5.1, рис. 5.2 та в табл. 5.1 варіантів використання, до складу САРЕС входить:

- ПВФПА – підсистема визначення функціональних параметрів аналізатора;
- ПРБП – підсистема реєстрації та первинної обробки БП;
- ПНА – підсистема нейромережевого аналізу;
- ПРО – підсистема розпізнавання особи;
- ПРЕ – підсистема розпізнавання емоційного стану.

Крім цього, в САРЕС для перемикання режимів функціонування використовується окремий модуль перемикання режимів (МПР).

Передбачено, що САРЕС може функціонувати в режимах:

- РАС – аутентифікації слухача;
- РНС – навчання слухача;
- РПНМ – визначення параметрів НММ;
- РЗ – зупинки.

Як показано на рис. 5.3, до складу ПВФПА входять модулі:

- МВМПР – визначення множини параметрів, що підлягають реєстрації;
- МВАНМ – визначення архітектури НММ;
- МПННМ – визначення параметрів навчання НММ;
- МВПРО – визначення параметрів розпізнавання особи;
- МВПРЕ – визначення параметрів розпізнавання емоцій.

До складу ПРБП входять модулі МРКП, МРГС, МРЗО, МРЗВР, МРСО, МРРОО, МРВП, функціональність яких пов’язана з реєстрацією та первинною обробкою таких БП, як КП, ГС, зображення обличчя, зображення вушної раковини, сітківка ока, райдужна оболонка ока, відбитки пальців.

Нейромережевий аналіз перерахованих БП реалізується в модулях ПНА, що призначені для аналізу:

- МАКП – КП;
- МАГС – ГС;

- МАЗО – зображення обличчя;

- МАЗВР – зображення вушної раковини;

- МАСО – сітківки ока;

- МАРОО – райдужної оболонки ока;

- МАВП – відбитків пальців.

Крім цього, до складу ПНА входить модуль інтегрованого аналізу (МІА), призначений для інтеграції результатів нейромережевого аналізу окремо кожного з перерахованих БП. Відзначимо, що, з позицій теорії технічної діагностики, впровадження в ПНА модуля МІА дозволяє підвищити точність розпізнавання емоційного стану і особи слухача, оскільки остаточна класифікація реалізується за рахунок узагальненої оцінки декількох гарантовано значущих діагностичних параметрів. Також використання модуля МІА дозволяє підвищити надійність результатів нейромережевого аналізу у випадку виникнення труднощів, пов'язаних з реєстрацією того чи іншого БП. В базовому варіанті функціональність модулю МІА співвідноситься з восьмим етапом методології розпізнавання емоційного стану та особи слухача СДН.

Функціональність підсистем ПРО і ПРЕ полягає в інтерпретації результатів нейромережевого аналізу БП у вигляді оцінок розпізнаних емоцій і особи слухача. До складу цих підсистем входять модулі:

- МРО – розпізнавання особи;

- МРЕ – розпізнавання емоцій та емоційного стану;

- МСРО – сигналізації про розпізнану особу;

- МСРЕ – сигналізації про розпізнані емоції та емоційний стан.

Відзначимо, що, відповідно до [97, 117, 223], представлені на рис. 5.1 і рис. 5.2 діаграми та показана на рис. 5.3 структурна схема можуть служити базовим варіантом опису САРЕС слухачів СДН як у випадку інтегрального використання декількох БП, так і у випадку ізольованого використання одного БП.

В якості прикладу адаптації розроблених архітектурних рішень до умов задачі розпізнавання особи і емоцій слухача за одним БП, розроблена архітектура нейромережевого аналізатора ГС. Побудовано діаграми варіантів використання, класів і компонентів, що показані на рис. 5.4-5.6.

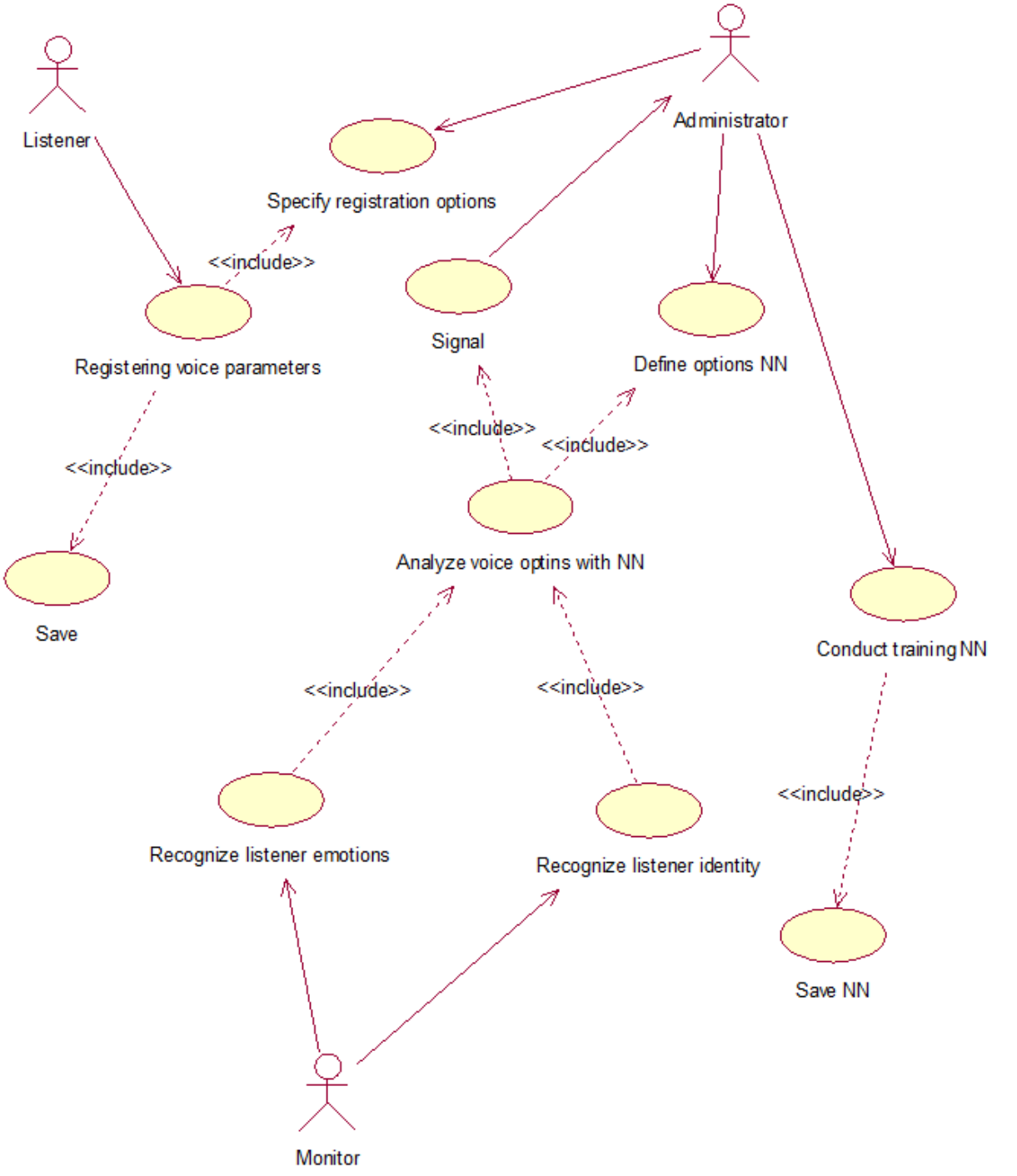

Рис. 5.4. Діаграма варіантів використання нейромережевого аналізатора ГС

До складу показаної на рис. 5.4 діаграми варіантів використання входять актори "Administrator", "Monitor" та "Listener".

Актор "Administrator" представляє собою адміністратора системи і відповідає за налаштування параметрів модуля нейромережевого аналізу ГС слухача СДН.

Актор "Listener" співвідноситься із слухачем СДН, який використовує голосовий інтерфейс і чий ГС підлягає нейромережевому аналізу.

Актор "Monitor" співвідноситься з блоком моніторингу слухача СДН.

Розроблена діаграма використання передбачає виконання наступних функціональних завдань:

- Register voice parameters – реалізація реєстрації параметрів ГС;
- Specify registration options – визначення параметрів запису ГС;
- Save – збереження зареєстрованих параметрів ГС;
- Recognize listener identity – розпізнавання особи слухача;
- Recognize listener emotions – розпізнавання емоцій слухача;
- Define options NN – визначення архітектурних параметрів НММ;
- Conduct training NN – навчання НММ;
- Save NN – збереження параметрів НММ;
- Analyze voice options with NN – аналіз зареєстрованих параметрів ГС за допомогою НММ;
- Signal – сигналізація про результати нейромережевого аналізу, що свідчить про результати розпізнавання особи і емоційного стану слухача.

Показана на рис. 5.5 діаграма класів визначає типи класів модуля нейромережевого аналізу ГС і зв'язки між цими класами. У загальному випадку діаграма класів – це граф, вершинами якого є елементи типу "класифікатор", які пов'язані між собою різними типами структурних відносин. Слід зауважити, що діаграма класів може також містити інтерфейси, об'єкти і зв'язки між ними. Вважається, що діаграма класів повинна відображати статичну структурну модель системи, що проєктується. Тому діаграму класів прийнято вважати графічним

представленням таких структурних взаємозв'язків логічної моделі системи, які не залежать від часу.

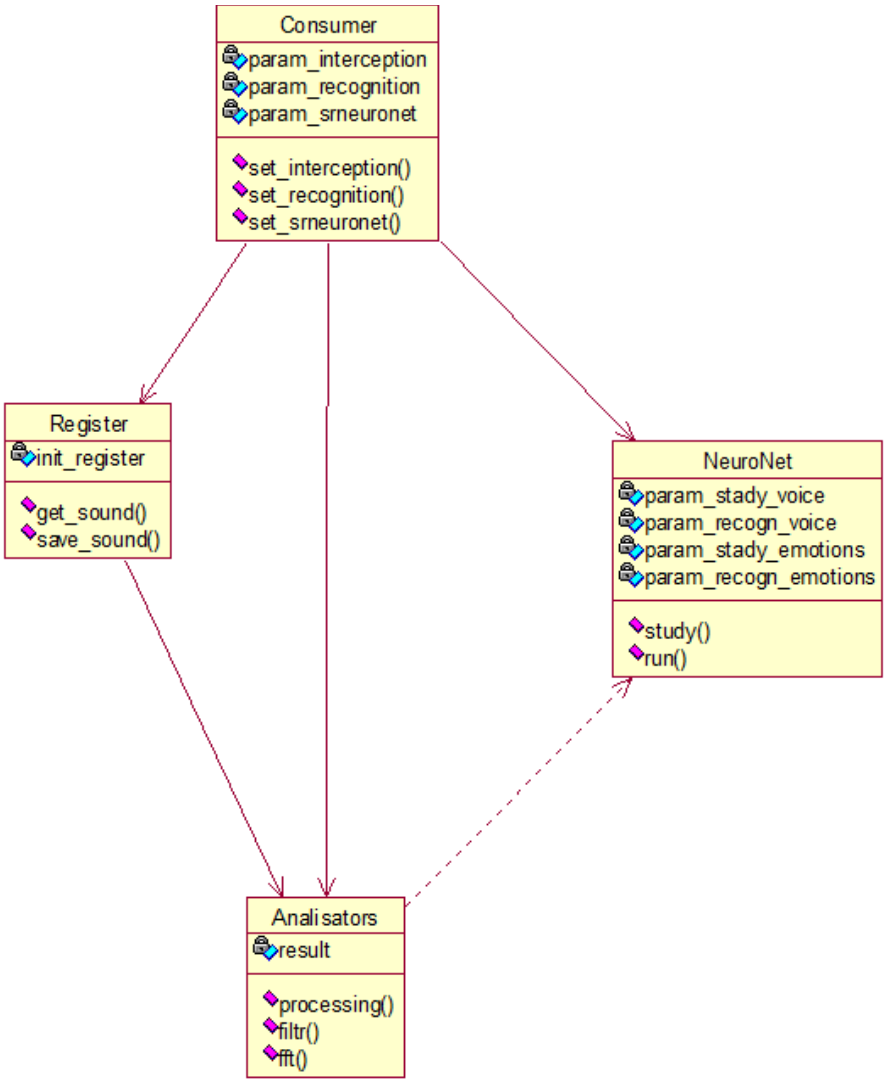

Рис. 5.5. Діаграма класів нейромережевого аналізатора голосу слухача СДН

На побудованій діаграмі класів передбачено відображати основні атрибути класів і обмеження, які накладаються на зв'язки між класами. До її складу входять класи, що призначені для:

- Register – реєстрації ГС;
- Consumer – обробки ГС;
- NeuroNet – навчання НММ;

- Analisators – нейромережевого аналізу ГС.

Показана на рис. 5.6 діаграма компонентів характеризує особливості фізичного представлення нейромережевого аналізатора і забезпечує візуалізацію структури програмної системи.

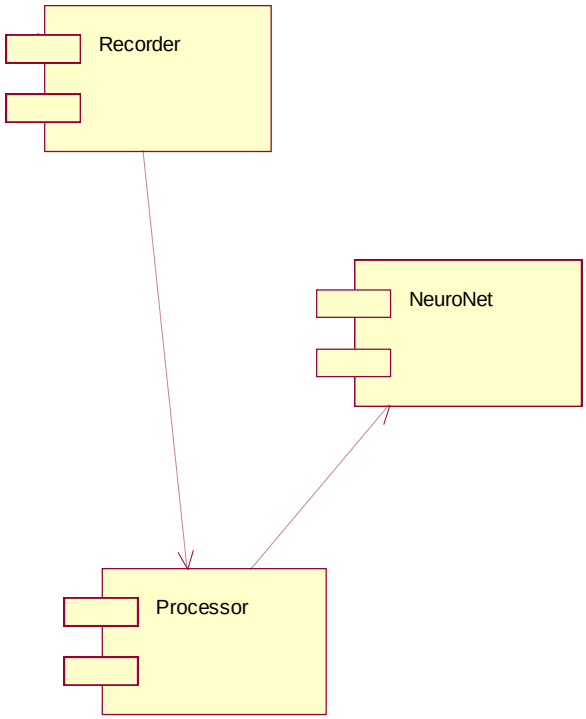

Рис. 5.6. Діаграма компонентів нейромережевого аналізатора ГС

Також діаграма компонентів дозволяє за рахунок формування залежностей між програмними компонентами спростити компоновку блоків, складових архітектуру модуля. Пунктирні стрілки, що з'єднують блоки даної діаграми, показують відношення взаємозалежності, аналогічні тим, які мають місце при компіляції вихідного коду. Основними графічними елементами показаної на рис. 5.6 діаграми є компоненти, інтерфейси і залежності між ними.

Передбачається, що нейромережевий аналізатор повинен складатися з трьох блоків: Recorder, Processor, NeuroNet. Зазначені блоки представляють інтерфейс отримання параметрів ГС слухача, обробку отриманих параметрів, а також розпізнавання особи і емоційного стану слухача СДН.

В доповнення до представлених діаграм розроблена показана на рис. 5.7 структурна схема нейромережевого аналізатора ГС слухача СДН.

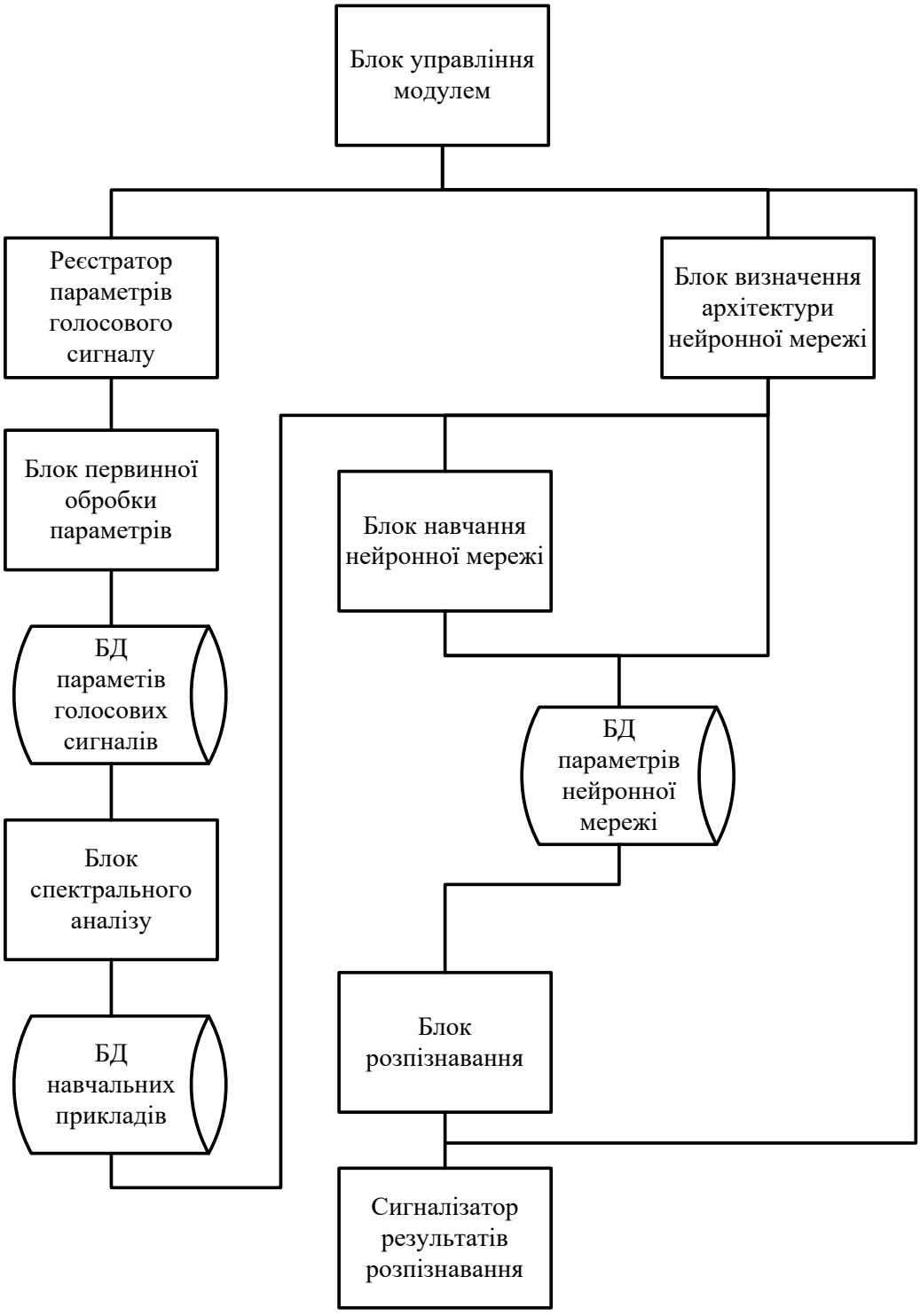

Рис. 5.7. Структурна схема нейромережевого аналізатора ГС

Як показано на рис. 5.7, аналізатор складається з 8 окремих блоків і 3 сховищ даних, призначених для зберігання параметрів голосу, навчальної вибірки і параметрів НМ.

Наведені архітектурні рішення послужили основою для розробки інструментальних засобів, призначених для розпізнавання емоційного стану та особи слухача СДН на основі нейромережевого аналізу кожного із основних та додаткових БП. Виходячи з позицій експериментального характеру вказаних інструментальних засобів та мінімізації ресурсів на розробку нейромережевих компонент, для написання програмного коду використано мову програмування Python та програмні бібліотеки TensorFlow і Keras. Лістинг програмного коду частково наведено в додатку Б.

Таким чином, в результаті проведених досліджень в термінах мови моделювання UML розроблена архітектура CAPEC слухачів СДН. Також наведено приклад адаптації отриманих архітектурних рішень до умов задачі розпізнавання особи і емоційного стану слухачів СДН на основі ГС. Отримані архітектурні рішення послужили основою для розробки відповідних інструментальних засобів. Крім того, показано, що напрямки подальших досліджень пов'язані з розробкою математичного забезпечення модуля узагальнення результатів нейромережевого аналізу зареєстрованих БП.

5.2. Оцінка ефективності засобів нейромережевого аналізу основних біометричних параметрів

Відповідно до заявленого функціоналу запропонованої методології нейромережевого розпізнавання емоційного стану та особи слухача СДН, проведені комп’ютерні експерименти, спрямовані на оцінювання ефективності НМЗ аналізу ГС, КП та зображення обличчя.

Відповідно до рекомендацій [91, 226, 265], для оцінки ефективності ЗНМ

використані такі параметри, як точність розпізнавання (A) та крос-ентропія (L) на навчальній, валідаційній та тестовій вибірці.

Для розрахунку значень цих показників використано вирази виду:

$$A = \frac{N_r}{N}, \tag{5.1}$$

$$L = \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{K} a_{n.k} ln\left(y_{n,k}\right), \tag{5.2}$$

де N_r - кількість правильно розпізнаних прикладів;

N - загальна кількість прикладів;

K - кількість класів, що мають бути розпізнані;

$y_{n,k}$ - значення вихідних сигналів НММ, що свідчать про віднесення деякого n-го прикладу до k-го класу.

Якщо n-ий приклад відноситься до k-го класу, то $a_{n.k}$=1. В протилежному випадку $a_{n.k}$=0.

Оцінка ефективності НМЗ аналізу ГС. При проведенні експериментів, пов'язаних з оцінкою ефективності НМЗ аналізу ГС, враховано широку варіативність параметрів записів емоційно забарвлених ГС, представлених в доступних БД, попередня обробка яких здійснюється при виконанні 2-5 етапів запропонованої методології. Вказана варіативність визначає можливість застосування в НМЗ декількох типів НММ: ДШП, БШП, класичної ЗНМ типу LeNet-5 та сучасних типів ЗНМ.

Зазначимо, що розраховані значення функції ефективності вказаних типів НММ відрізняються не більше, ніж на 3-5%.

У випадку використання ДШП для розпізнавання емоційного стану прийнято, що для формування тренувальної та тестової вибірки використовуються записи ГС, представлені в БД Emotion_Recognition (www.kaggle.com). З представленої БД використано записи голосу одного диктора. Диктор англійською мовою в студійних

умовах озвучував 19 текстів різного змісту, виявляючи при цьому чотири різні емоції (радість, нейтральність, смуток та злість). Для формування як навчальної, так і тестової вибірки використовувалися лише перші 7,008 с кожного із записів.

Обсяг сформованої тренувальної вибірки - 760 прикладів, а обсяг тестової вибірки - 76 прикладів. За аналогією з [275, 290, 323], як вхідні параметри ДШП передбачено використовувати мел-кепстральні коефіцієнти, розраховані за допомогою виразів (1.1-1.12) для кожного із квазістаціонарних фрагментів ГС. Відповідно до рекомендацій [280, 281, 290, 311], тривалість кожного із квазістаціонарних фрагментів прийнята рівною t_{st}=16 мс. Це пояснюється можливістю ефективного застосування швидкого перетворення Фур'є для розрахунку спектру квазістаціонарного фрагмента при частоті дискретизації 16000 Гц, характерної для записів ГС у доступних базах.

У першому наближенні прийнято, що кількість мел-кепстральних коефіцієнтів дорівнює 24. Ця цифра прийнята з міркувань простоти реалізації розрахунку з використанням апробованих програмних комплексів. Така кількість мел-кепстральних коефіцієнтів відповідає 21000 вхідних параметрів моделі. При цьому розрахунок кількості вхідних параметрів реалізований з урахуванням половинного перекриття квазістаціонарних фрагментів при тривалості ГС 7,008 с. Для розрахунку кількості схованих нейронів використано вирази (4.43-4.45).

Таким чином, у даній серії експериментів кількість вихідних нейронів N_{out}=4, а розрахована кількість схованих нейронів N_h=70.

Модель ДШП із сигмоїдальною функцією активації схованих та вихідних нейронів реалізована програмно з використанням мови програмування Python та бібліотеки Keras. Також при створенні програми використано бібліотеки Numpy – для обробки масивів, Soundfile – для обробки звукових файлів та Matplotlib – для візуалізації результатів.

Експерименти проводилися на персональному комп'ютері з процесором Intel Core i7-8700 (3.2 - 4.6 ГГц), об'ємом оперативної пам'яті 16 ГБ, відеокартою nVidia

GeForce GTX 1660Ti, який функціонував під керуванням операційної системи Microsoft Windows 10. Зазначимо, що вказане апаратне забезпечення використано у всіх експериментах, пов'язаних з оцінкою ефективності НМЗ аналізу основних БП.

Ефективність ДШП оцінювалася за допомогою показника точності розпізнавання на тренувальній та тестовій вибірці за 100 епох навчання. Вказані показники розраховані за допомогою виразів (5.1, 5.2).

На рис. 5.8 показані графіки залежності точності розпізнавання (Accuracy) на тренувальній (Train) та тестовій вибірці (Test) від кількості епох навчання (Epoch).

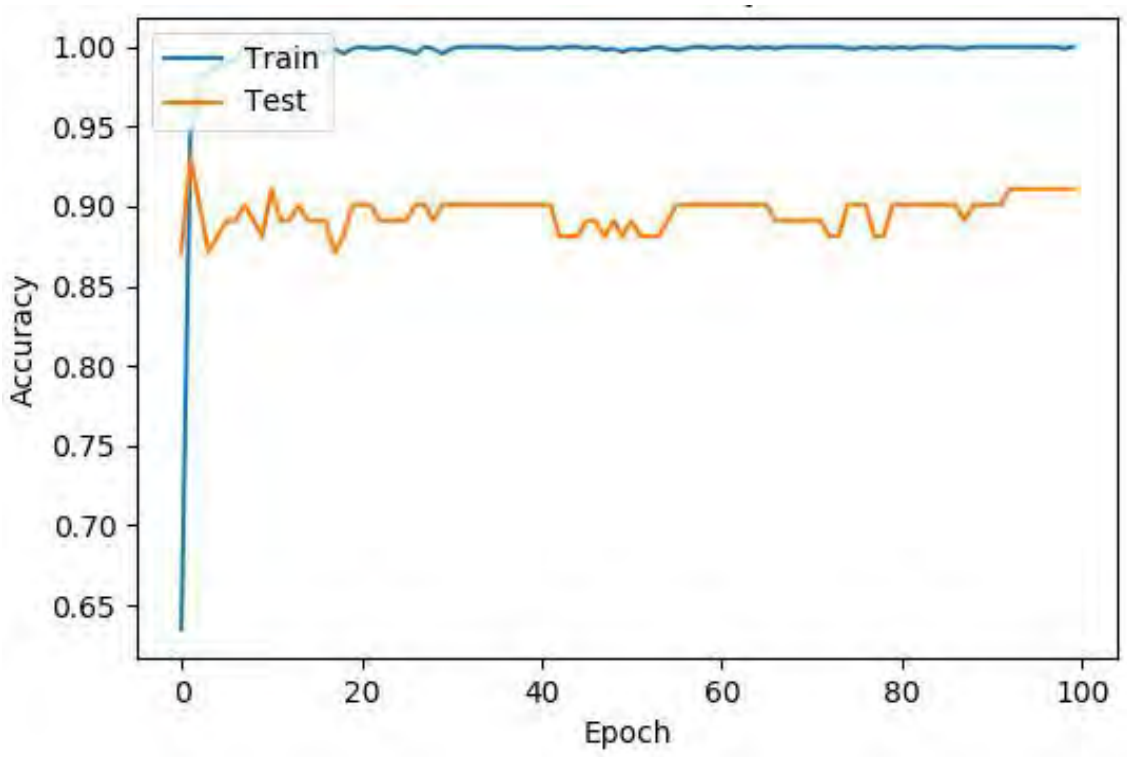

Рис. 5.8. Графік залежності точності розпізнавання емоційного стану від кількості епох навчання для тренувальних та тестових даних при використанні 24 мел-кепстральних коефіцієнтів

Слід зазначити, що досягнута за допомогою ДШП точність розпізнавання емоційного стану диктора на тестовій вибірці становить приблизно 0,94.

Наступний етап досліджень пов'язаний із оцінкою ефективності НМЗ на основі БШП. Для створення навчальної вибірки використано БД TESS, в якій представлено записи емоційно забарвлених ГС двох дикторів. Кількість представлених емоцій у цій БД дорівнює 7, а кількість записів, що використовується для формування навчальних прикладів, дорівнює 2800.

Використовуючи розроблений метод формування вхідного поля НММ, показано можливість використання квазістаціонарного фрагменту ГС тривалістю 0,011 мс. При цьому визначено доцільність проведення експериментальних досліджень, пов'язаних з розрахунком кількості мел-кепстральних коефіцієнтів, що характеризують кожен із квазістаціонарних фрагментів. Зазначимо, що кількість мел-кепстральних коефіцієнтів, які використовуються для опису одного квазістаціонарного фрагмента, є однією зі складових виразу (3.83), що застосовується для розрахунку розміру вхідного поля ЗНМ.

Для проведення експериментальних досліджень за допомогою мови програмування Python розроблено спеціалізоване програмне забезпечення.

Відповідно до [86, 187, 290, 323], доцільність використання тієї чи іншої кількості мел-кепстральних коефіцієнтів прийнято оцінювати за допомогою показників A та L на тестових даних, які не використовуються під час навчання. Оптимальним вважається така кількість мел-кепстральних коефіцієнтів, при якій точність розпізнавання максимізується, а втрати мінімізуються. Таким чином, план експериментів передбачав побудову залежностей виду:

$$A = f\left(K_{mel}, K_{epoch}\right), K_{mel} \in [1, 2, \ldots, 24], \tag{5.3}$$

$$L = f\left(K_{mel}, K_{epoch}\right), K_{mel} \in [1, 2, \ldots, 24], \tag{5.4}$$

де K_{epoch} - кількість епох навчання.

Для побудови залежностей (5.3, 5.4) на тестових даних необхідно попередньо навчити БШП, що передбачає побудову таких самих залежностей на тренувальних даних. Використовуючи рекомендації [70, 98], встановлено, що обсяг тестових

даних має становити 10% загального обсягу навчальної вибірки.

Наприклад, на рис. 5.8 та рис. 5.9 показані графіки залежності точності розпізнавання та крос-ентропії на тренувальній та тестовій вибірці від кількості епох навчання при описі одного квазістаціонарного фрагмента за допомогою дев'ятьох мел-кепстральних коефіцієнтів.

У процесі експериментів встановлено, що величини зазначених показників A та L стабілізуються після 300 епох навчання. Тому у виразах (5.3, 5.4) прийнято, що $K_{epoch} = 300$.

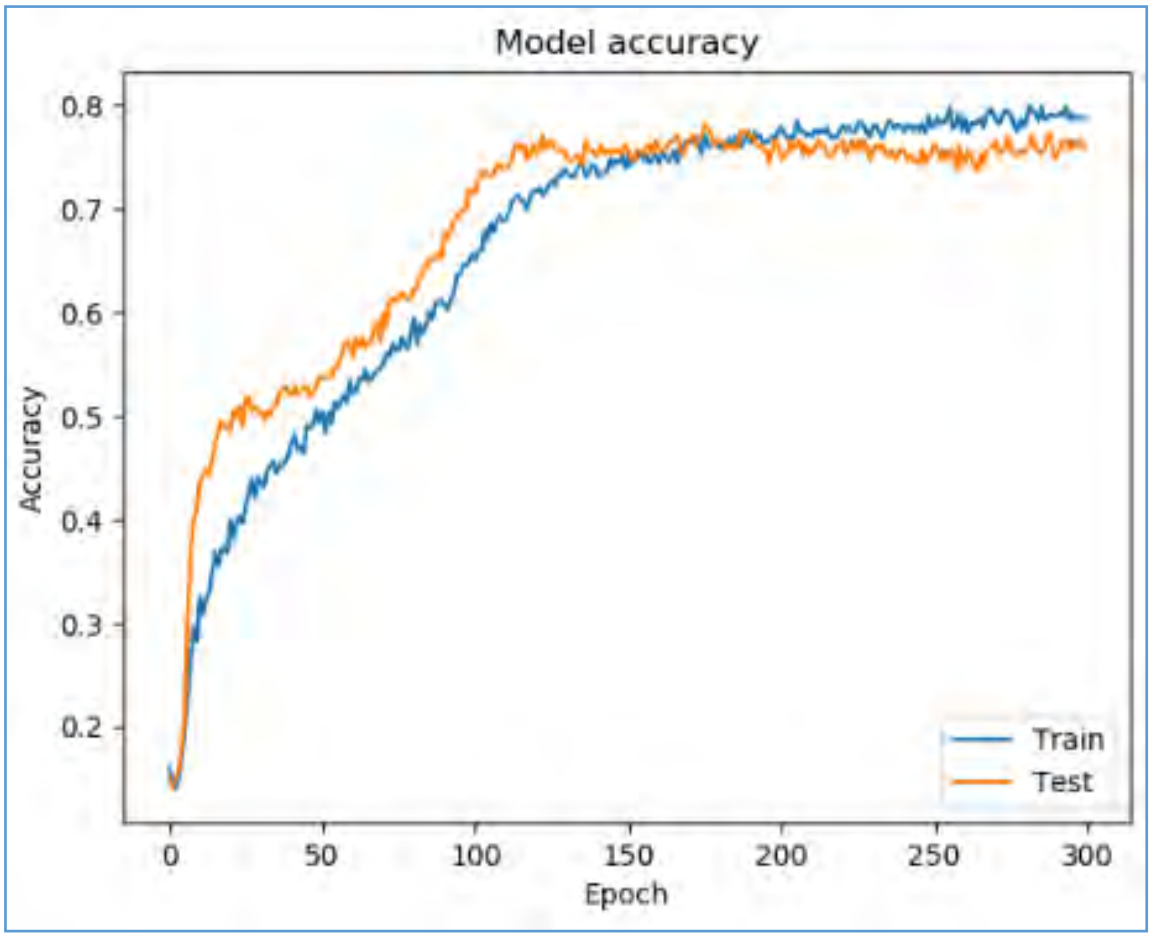

Рис. 5.9. Графіки залежності точності розпізнавання від кількості епох навчання для тренувальних та навчальних даних при використанні 9 мел-кепстральних коефіцієнтів

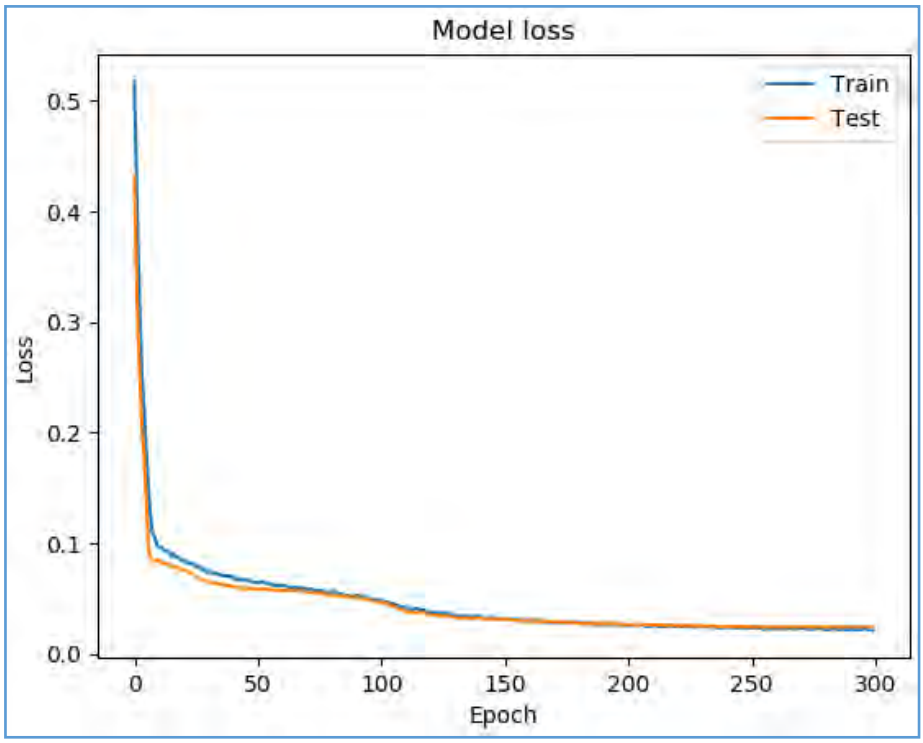

Рис. 5.10. Графіки залежності показника крос-ентропії від кількості епох навчання при використанні 9 мел-кепстральних коефіцієнтів

В результаті проведених експериментів побудовано показані на рис. 5.11 та рис. 5.12 графіки залежностей точності розпізнавання та крос-ентропії на тестових даних від кількості мел-кепстральних коефіцієнтів, що використовуються для опису одного квазістаціонарного фрагмента ГС. Як свідчать результати експериментів, максимальна точність розпізнавання та мінімальне значення крос-ентропії при використанні БШП досягаються у випадку опису одного квазістаціонарного фрагмента 20 мел-кепстральними коефіцієнтами. У цьому випадку досягається точність розпізнавання емоційного стану диктора близько 0,94. За такої кількості мел-кепстральних коефіцієнтів кількість вхідних параметрів БШП дорівнює 10220. Також результати експериментів показали, що збільшення числа мел-кепстральних коефіцієнтів понад 20 не призводить до збільшення точності розпізнавання, проте через збільшення кількості вхідних нейронів підвищує ресурсоємність БШП.

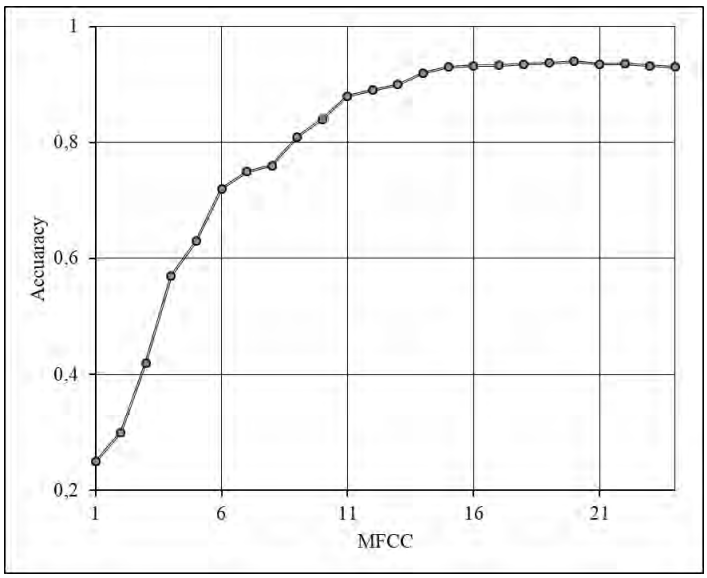

Рис. 5.11. Графік залежності точності розпізнавання емоційного стану диктора від кількості мел-кепстральних коефіцієнтів

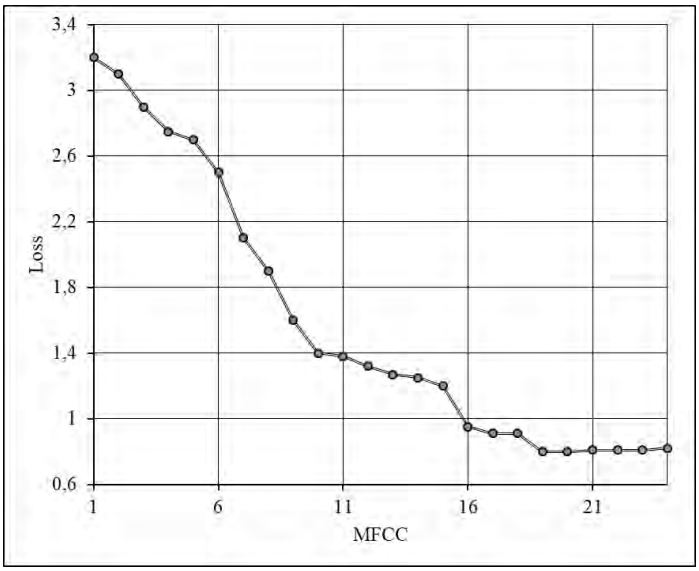

Рис. 5.12. Графік залежності крос-ентропії для HMM розпізнавання емоційного стану диктора від кількості мел-кепстральних коефіцієнтів

Також проведена апробація застосування НМЗ на основі найбільш апробованого типу ЗНМ – LeNet-5. Адаптація ЗНМ до поставленої задачі розпізнавання емоційного стану диктора реалізована на основі рішень,

використаних в запропонованій методології автоматизованого розпізнавання емоційного стану слухача СДН. Формування вхідного поля ЗНМ реалізоване з використанням наведеного в п. 3.4 відповідного методу, з урахуванням представлення фіксованого фрагменту ГС у вигляді квадратного зображення з одним кольоровим каналом. Структурні параметри ЗНМ, адаптованої до завдання розпізнавання емоційного стану диктора на фіксованих ділянках ГС із БД TESS, представлені в табл. 5.2.

Таблиця 5.2

Значення структурних параметрів згорткової нейронної мережі, адаптовані до задачі розпізнавання емоційного стану за голосом

Назва параметру	**Значення параметру**
Розмір вхідного поля	$(a_0 \times a_0) = (1140 \times 1140)$
Кількість шарів згортки	$K_{ls} = 5$
Кількість карт згортки в і-му шарі згортки	$L_{h,1} = 36, L_{h,2} = L_{h,3} = L_{h,4} = L_{h,5} = 256$
Розмір карти згортки в і-му шарі	$(a_1 \times a_1) = (378 \times 378), (a_2 \times a_2) = (182 \times 182), (a_3 \times a_3) = (84 \times 84), (a_4 \times a_4) = (34 \times 34), (a_5 \times a_5) = (10 \times 10)$
Розмір ядра згортки	$b = 9$
Зсув ядра згортки в і-му шарі згортки	$d_1 = 3, d_2 = d_3 = d_4 = d_5 = 1$
Кількість шарів субдискретизації	$K_{ld} = 5$
Масштабний коефіцієнт	$m = 2$
Кількість нейронів у повнозв'язному шарі	$L_f = 1000$
Кількість вихідних нейронів	$L_{out} = 7$

Зазначимо, що адаптація параметрів ЗНМ реалізована відповідно до запропонованого в п. 4.2 методу розробки архітектури ЗНМ. Розроблена модель ЗНМ реалізована програмно з використанням мови програмування Python та бібліотеки TensorFlow, що дозволило провести комп'ютерні експерименти, спрямовані на верифікацію запропонованих рішень щодо ефективності застосування такої моделі для розпізнавання емоційного стану на основі ГС. План експериментів передбачав визначення точності розпізнавання у разі зміни конструктивних параметрів ЗНМ. Наприклад, на рис. 5.13 показано залежність точності розпізнавання емоційного стану від кількості епох навчання для різної кількості карт ознак у першому згортковому шарі. Кількість карт ознак в інших згорткових шарах дорівнювала 256.

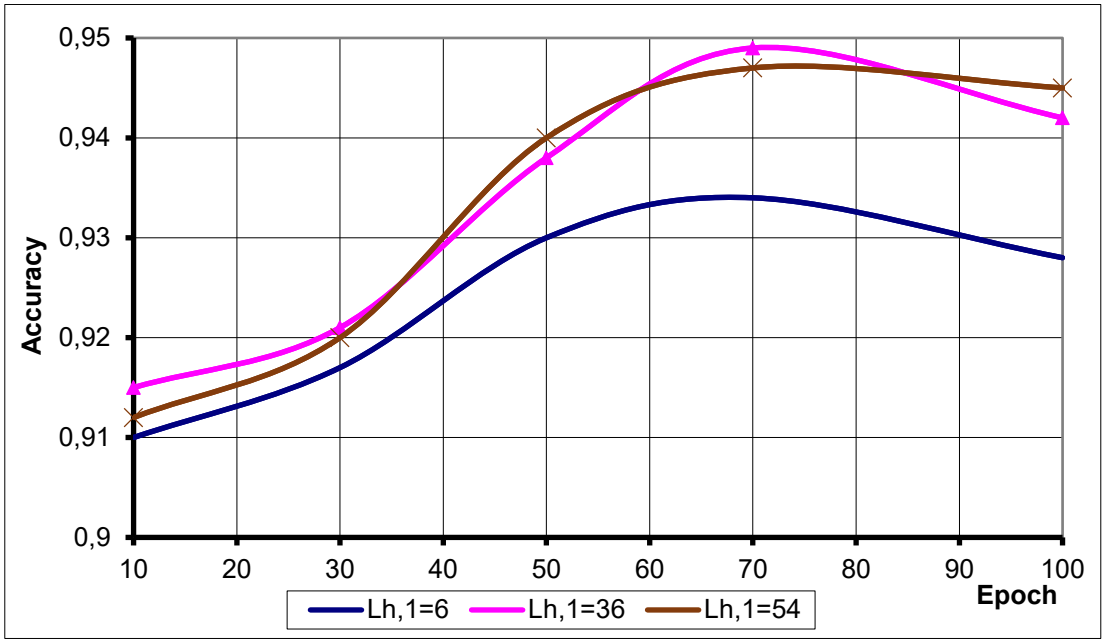

Рис. 5.13. Залежність точності розпізнавання від кількості епох навчання при різних значеннях $L_{h,1}$

Максимальна точність розпізнавання емоційного стану диктора, що дорівнює близько 0,95, досягнута при значеннях параметрів ЗНМ, що наведені у табл. 5.2. В порівнянні з [193, 225, 227, 247], використання запропонованого методу розробки

архітектури ЗНМ дозволило приблизно в 1,2 рази зменшити обсяг експериментів, спрямованих на визначення структурних параметрів.

На заключному етапі експериментальних досліджень, пов’язаних з нейромережевим аналізом ГС для розпізнавання емоційного стану, проведені експерименти, спрямовані на апробацію ЗНМ типу SqueezeNet. Характерні особливості SqueezeNet полягають у відносно невисокій ресурсоємності та достатньої точності розпізнавання, що у ряді випадків може мати вирішальне значення для систем моніторингу емоційного стану слухача СДН.

Крім власне апробації сучасного типу ЗНМ, ще одним завданням експериментів була перевірка гіпотези про можливість перенесення знань в НММ. Тому хоча структура використаної SqueezeNet була адаптована до задачі розпізнавання емоційного стану за ГС, першочерговий розподіл вагових коефіцієнтів цієї НММ був отриманий внаслідок навчання SqueezeNet на прикладах зображень із БД ImageNet (https://image-net.org/).

Інші умови проведення експерименту відповідають умовам, що використані при апробації ЗНМ типу LeNet-5. Як і для LeNet-5, вхідне поле SqueezeNet асоційоване з множиною мел-кепстральних коефіцієнтів, що характеризують піддослідний ГС. Так як тривалість ГС у wav-файлах БД TESS варіюється в межах 1,5-2 с, з урахуванням можливості зміни тривалості квазістаціонарного фрагменту в межах від 0,01 до 0,02 с, доцільності половинного перекриття квазістаціонарних фрагментів, особливостей реалізації швидкого перетворення Фур'є, специфіки SqueezeNet, прийнято, що розмір вхідного поля НММ дорівнює 227х227. Через те, що кількість мел-кепстральних коефіцієнтів дорівнює 24, то, починаючи з 25-го стовпця, стовпці вхідного поля мережі заповнюються нулями. Оскільки у початковому варіанті мережа SqueezeNet налаштована для класифікації зображень на 1000 категорій, а в поставленій задачі аналізу ГС передбачено розпізнавання 7 емоцій, останні 5 шарів мережі були адаптовані під цю умову.

Структура модифікованої мережі SqueezeNet показана на рис. 5.14.

Рис. 5.14. Структура нейронної мережі типу SqueezeNet

Мережа SqueezeNet реалізована за допомогою засобів MATLAB R2018b, що дозволило провести комп’ютерні експерименти, спрямовані на визначення

залежності значень показників точності розпізнавання та крос-ентропії від кількості епох навчання.

Графіки залежності точності розпізнавання та крос-ентропії від кількості епох навчання показано на рис. 5.15 та на рис. 5.16.

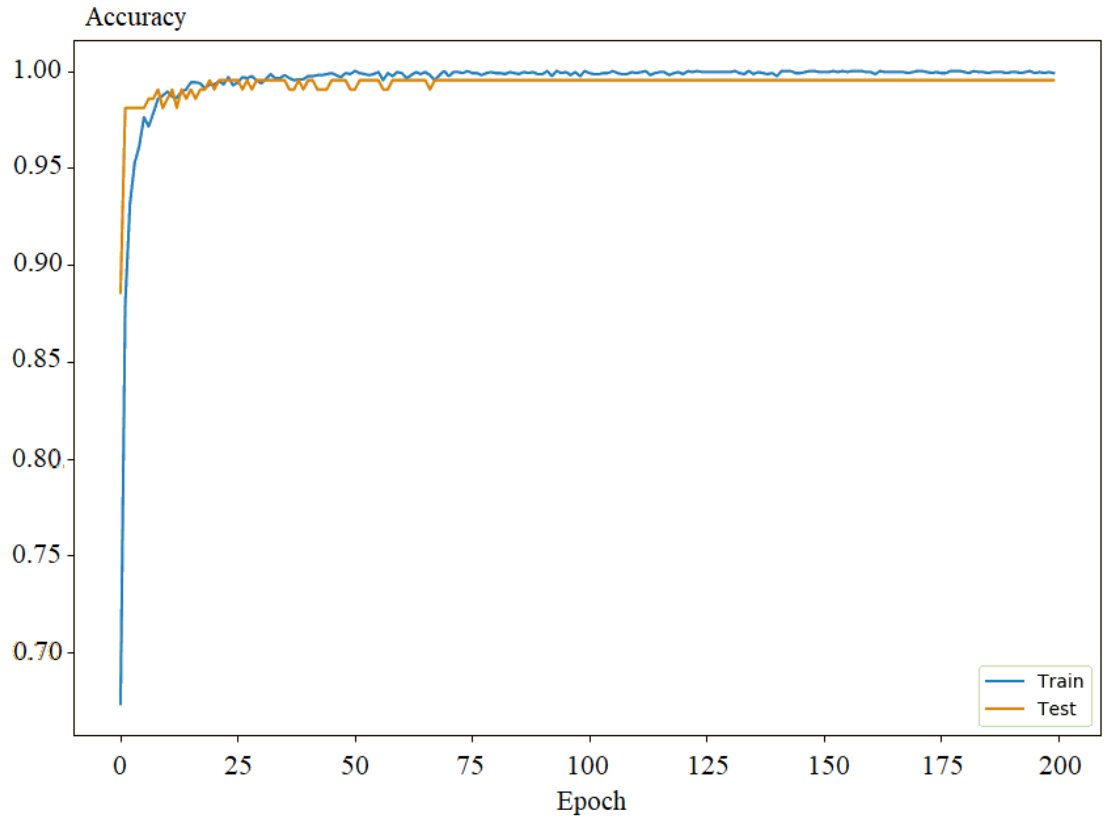

Рис. 5.15. Графіки залежності точності розпізнавання від кількості епох навчання для SqueezeNet

Як випливає з аналізу зазначених графіків, величини точності та крос-ентропії стабілізуються приблизно після 80-ої епохи навчання. При цьому точність розпізнавання на тестових прикладах становить близько 0,98. Зазначимо, що точність розпізнавання на тестовій вибірці, яка не використовувалась для навчання склала близько 0,95, що приблизно в 1,1-1,15 рази перевищує точність найкращих НМС розпізнавання аналогічного призначення [193, 227, 247].

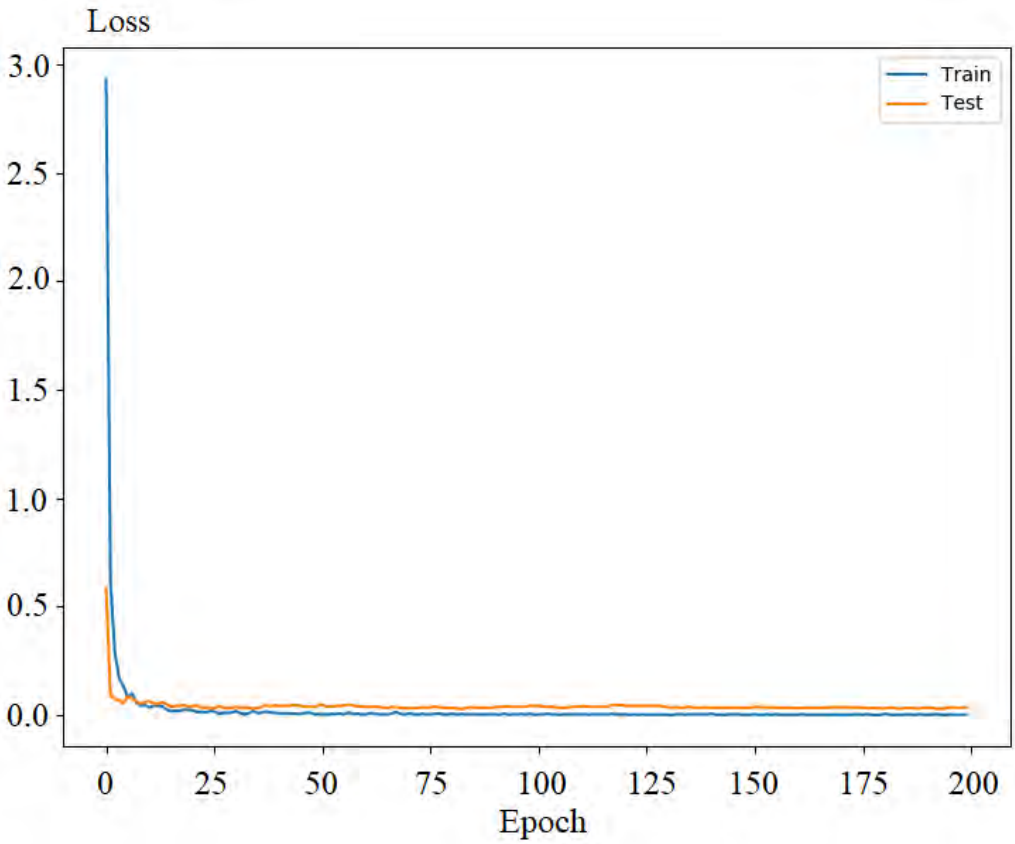

Рис. 5.16. Графіки залежності крос-ентропії від кількості епох навчання для SqueezeNet

Таким чином, результати експериментів вказують на доцільність використання НММ типу SqueezeNet для аналізу ГС для розпізнавання емоцій диктора. Слід зазначити, що ГС, що представлені у БД TESS, були записані в студійних умовах. При цьому на практиці викликає інтерес розпізнавання емоцій диктора за умов впливу різноманітного шуму.

Для оцінки впливу шуму на точність розпізнавання проведено експерименти, у яких, за допомогою описаної вище мережі SqueezeNet, проводилося розпізнавання зашумлених навчальних прикладів. Наприклад, на рис. 5.17 та на рис. 5.18. показані графіки залежності точності розпізнавання та крос-ентропії для тренувальних та тестових прикладів залежно від кількості епох навчання при впливі на тестові приклади шуму гучністю 10 дБ.

В результаті експериментів визначено, що при впливі білого шуму гучністю близько 20 дБ, точність розпізнавання емоцій знизилася і складає приблизно 0,6.

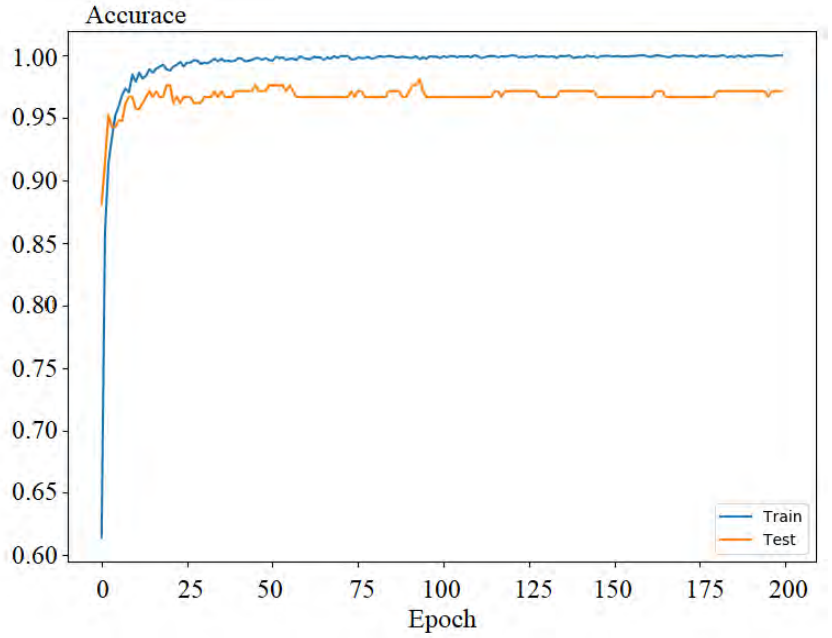

Рис. 5.17. Графіки залежності точності розпізнавання від кількості епох навчання при впливі шуму

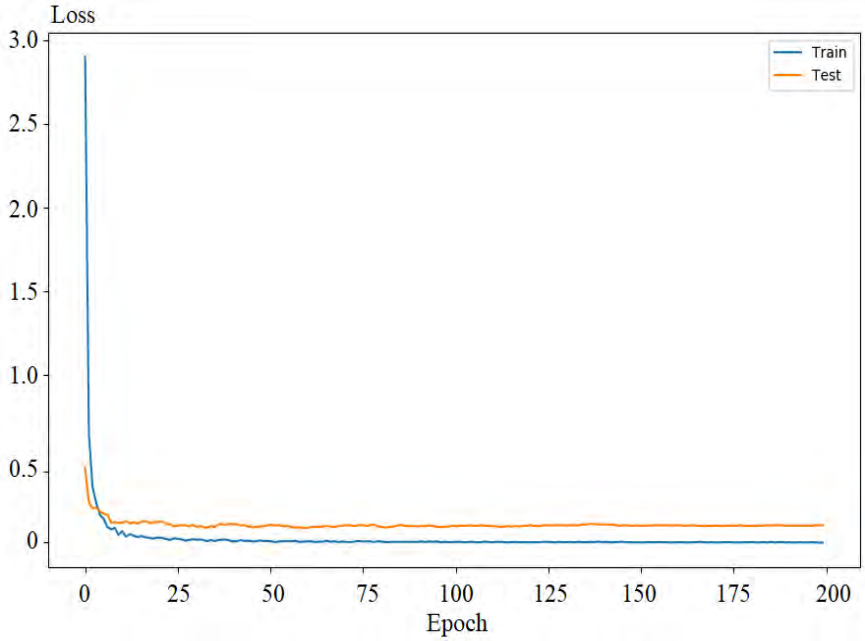

Рис. 5.18. Графіки залежності крос-ентропії від кількості епох навчання при впливі шуму

Зважаючи на отримані результати, шляхи подальших досліджень доцільно співвіднести з адаптацією нейромережевих рішень до розпізнавання емоцій диктора за умов реалізації різноманітних шумових перешкод.

Також, базуючись на загальноприйнятій методології підвищення ефективності нейромережевих засобів діагностування [42, 94, 100], можна припустити перспективність досліджень у напрямку формування репрезентативної навчальної вибірки для створення дикторонезалежних систем розпізнавання емоцій. Крім цього, враховуючи тісний зв'язок задачі розпізнавання емоцій диктора із завданням розпізнавання особи користувача за голосом, доцільно розробити метод адаптації архітектурних параметрів SqueezeNet до умов задачі аналізу ГС для одночасного розпізнавання особи та емоцій диктора.

Порівняння отриманих експериментальних результатів з даними [184, 193, 247, 257] вказують на те, що точність розпізнавання емоційного стану диктора за допомогою побудованих ДШП, БШП, ЗНМ типу LeNet-5 і SqueezeNet знаходиться на рівні найкращих сучасних систем аналогічного призначення. Однак широке впровадження засобів нейромережевого розпізнавання емоцій диктора однозначно спричинить суттєве збільшення обсягів навчальних вибірок ГС. В цьому випадку критичним фактором може стати висока ресурсоємність засобів розпізнавання на базі ДШП і БШП. Перспективним шляхом усунення цього недоліку являється використання ЗНМ, яка є менш ресурсоємним типом ГНМ.

Також проведено експериментальні дослідження, спрямовані на оцінку ефективності НМЗ розпізнавання особи диктора. Використано НМЗ на базі БШП. Базуючись на результатах [7, 58, 98], прийнято такі умови використання системи розпізнавання диктора:

- Кількість дикторів, що мають бути розпізнані - 10.
- Частота дискретизації ГС дорівнює 16000 Гц.
- Тривалість ГС становить 7,008 с.
- Тривалість квазістаціонарного фрагменту дорівнює 0, 016 с.

- Для опису одного квазістаціонарного фрагменту використовується 20 мел-кепстральних коефіцієнтів з половинним перекриттям.

- Для формування тренувальної та тестової вибірки використана БД Speaker_recognition, що розміщена на сайті www.kaggle.com.

Відповідно до наведених умов, кількість вхідних нейронів БШП дорівнює 17500, кількість вихідних нейронів – 10. Обсяг сформованої тренувальної вибірки становив 900 прикладів, а обсяг тестової вибірки – 100 прикладів. Фрагменти ГС кожного прикладу є унікальними.

Лістинг коду розробленого програмного застосунку для розпізнавання особи диктора представлений у додатку Б.

При підготовці експериментів враховані результати [126, 194], в яких показано, що для розпізнавання особи слухача за ГС можливо використовувати БШП, який містить як мінімум два сховані шари нейронів. При цьому в першому схованому шарі має бути 256 нейронів. Також зазначено, що тривалість навчання не повинна перевищувати 1000 епох.

Тому план досліджень передбачав проведення двох серій експериментів, спрямованих на:

- визначення оптимальної кількості нейронів у другому схованому шарі тришарового персептрону;

- визначення доцільності збільшення глибини БШП.

Використовуючи дані [42, 70], прийнято припущення, що діапазон пошуку оптимальної кількості нейронів у другому схованому шарі обмежений величинами 10 і 256. Тому в першій серії експериментів досліджувався тришаровий персептрон, для якого кількість нейронів у першому схованому шарі H_1=256, а кількість нейронів у другому схованому шарі H_2 змінювалась від 10 до 256. Наприклад, на рис. 5.19 та рис. 5.20 показані графіки залежності точності розпізнавання та крос-ентропії від кількості епох навчання. В процесі проведення експериментів визначено, що точність навчання досягає свого максимуму за 200 епох.

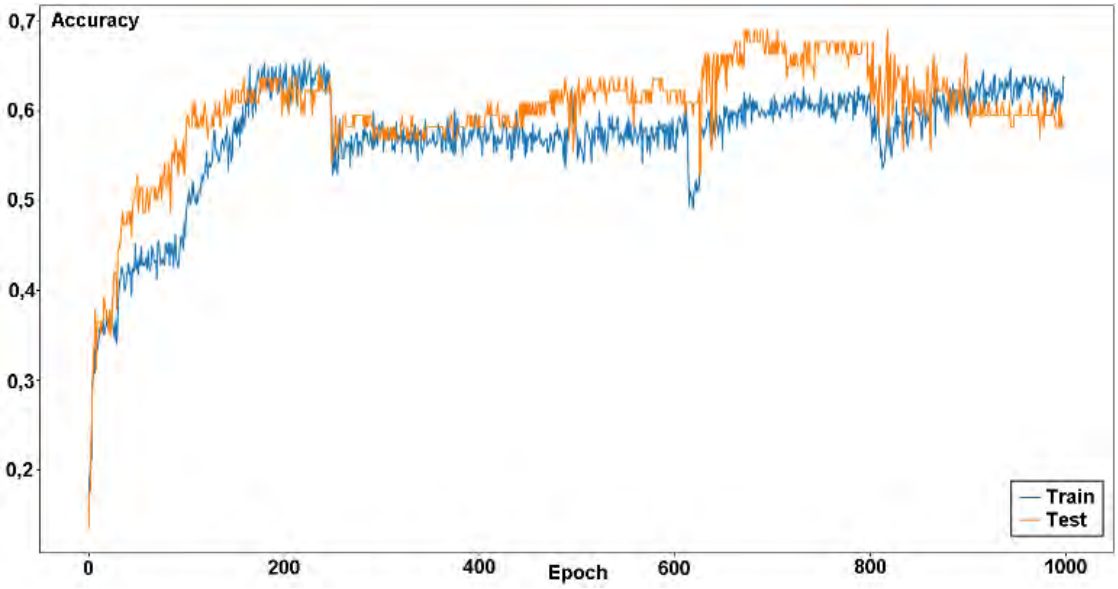

Рис. 5.19. Графік залежності точності розпізнавання від кількості епох навчання для тришарового персептрона при H_1=256 та H_2=10

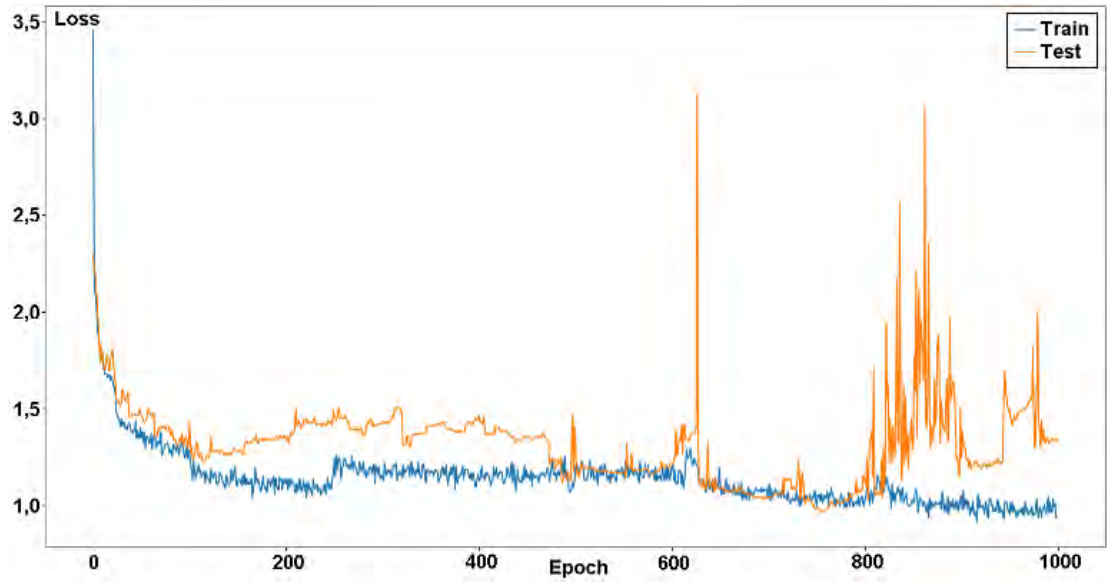

Рис. 5.20. Графік залежності крос-ентропії від кількості епох навчання для тришарового персептрона при H_1=256 та H_2=10

Надалі значення крос-ентропії на тестовій вибірці змінюється в негативний бік, що можна пояснити виникненням явища «перенавчання» НММ. Тому в експериментах кількість епох навчання обмежена 200. В результаті першої серії

експериментів отримано залежність точності та втрат при розпізнаванні тестових даних від кількості нейронів у другому схованому шарі. Відповідні графіки показано на рис. 5.21 та рис. 5.22.

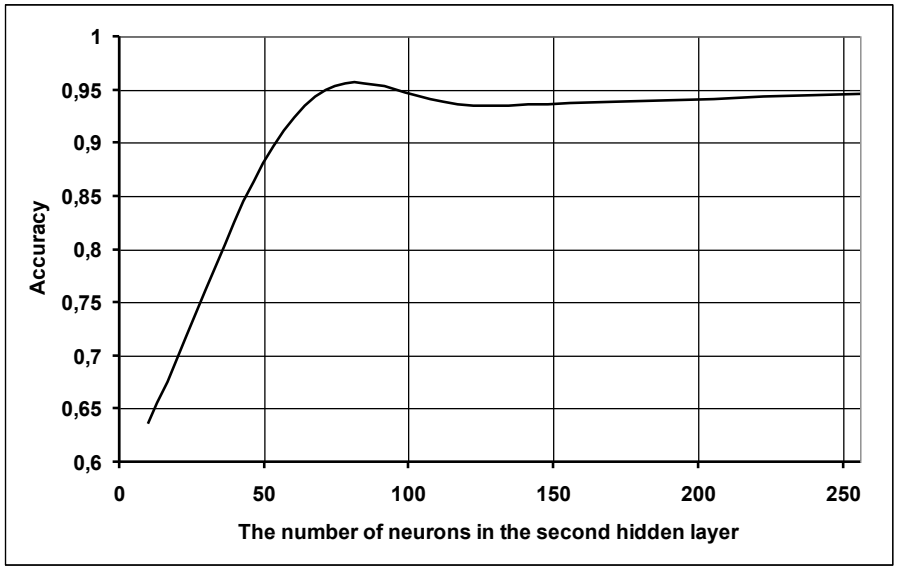

Рис. 5.21. Графік залежності точності розпізнавання тестових даних від кількості нейронів у другому схованому шарі

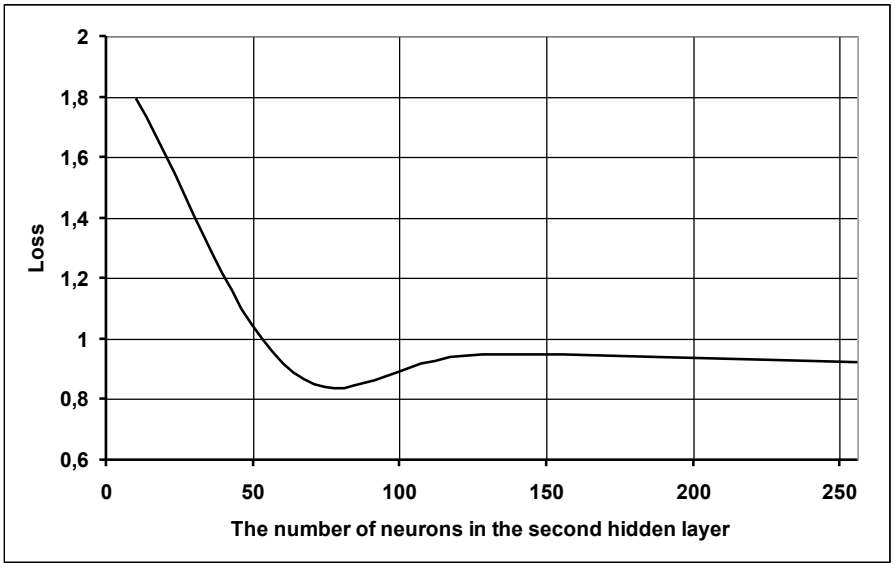

Рис. 5.22. Графік залежності крос-ентропії при розпізнаванні тестових даних від кількості нейронів у другому схованому шарі

Як випливає з аналізу цих графіків, точність розпізнавання стабілізується при кількості нейронів у другому схованому шарі, що дорівнює 80. Подальше збільшення кількості нейронів у другому схованому шарі не призводить до помітного збільшення точності, збільшуючи при цьому ресурсоємність НММ. Таким чином, результати першої серії експериментів вказують на доцільність використання у другому схованому шарі 80 нейронів. При цьому точність розпізнавання перевищує 0,95, а крос-ентропія дорівнює 0,84, що можна порівняти з показниками сучасних засобів розпізнавання диктора [129, 136, 205].

У другій серії експериментів було досліджено декілька БШП з різною структурою: (256-64-64), (128-64-64), (64-64-64-64), (64-56- 48-40-24). Зауважимо, що в описі структури БШП перша цифра позначає кількість нейронів у першому схованому шарі нейронів, друга – у другому і т. д. Наприклад, (256-64-64) описує БШП із трьома схованими шарами. У першому схованому шарі міститься 256 нейронів, у другому – 64, а у третьому – 64. Результати другої серії експериментів у вигляді діаграм представлені на рис. 5.23 та рис. 5.24.

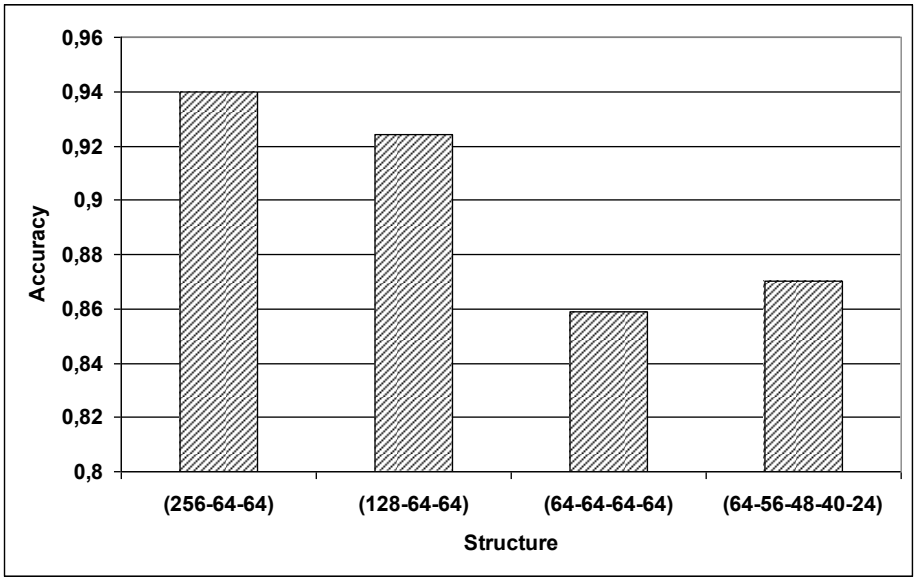

Рис. 5.23. Діаграма точності розпізнавання тестових даних для БШП з різною структурою

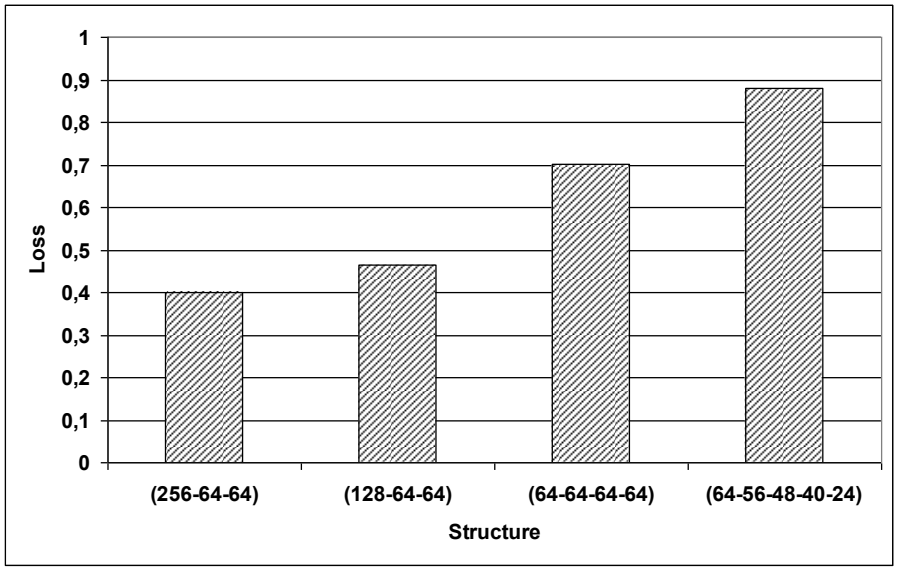

Рис. 5.24. Діаграма крос-ентропії при розпізнавання тестових даних для БШП з різною структурою

Аналіз зазначених діаграм, а також порівняння результатів першої та другої серії експериментів дозволяють стверджувати, що при приблизно рівній кількості схованих нейронів збільшення кількості схованих шарів не призводить до збільшення точності розпізнавання.

Таким чином, результати експериментів вказують на недоцільність використання БШП з кількістю схованих шарів нейронів більшою, ніж 2. Разом з тим, отримані дані свідчать, що при допустимому рівні ресурсоємності розроблена НММ дозволяє досягти точності розпізнавання близько 0,96, що можна порівняти з найбільш сучасними засобами біометричної аутентифікації на основі ГС.

Оцінка ефективності НМЗ аналізу КП. Результат виконання 2-5 етапів методології нейромережевого розпізнавання емоційного стану слухача СДН вказує на доцільність застосування для аналізу КП НМЗ на основі ЗНМ різних типів. Відповідно, проведення експериментальних досліджень розділено на два етапи, що співвідносяться із застосуванням класичної ЗНМ типу LeNet-5 та із застосуванням ЗНМ типу AlexNet.

Прийнято такі умови застосування ЗНМ типу LeNet-5:

- Кількість слухачів СДН, що мають бути розпізнані – 8.

- Розпізнаванню підлягають три емоції – нейтральна, радість та страх.

- Текст може складатися з великих літер англійського алфавіту та 6 знаків пунктуації (всього 32 символи).

- КП аналізується під час набору текстів, довжина яких становить 28 символів.

- У кожний фіксований момент часу параметрами КП, які підлягають аналізу є символ, що відповідає клавіші, а також значення параметрів ТУК і ТМК.

Зазначені умови застосування визначені з позицій оціночного характеру експериментальних досліджень, спрощення формування навчальної вибірки та можливості коректного представлення вхідного поля ЗНМ.

Архітектурні параметри використаної ЗНМ визначені за допомогою запропонованого в п. 4.2 методу розробки архітектури та мають наступні значення: розмір вхідного поля - 32х32, кількість вихідних нейронів - 11, кількість шарів згортки - 2, кількість шарів субдискретизації - 2, масштабний коефіцієнт – 0,5, кількість повнозв'язних шарів - 2, розмір ядер згортки - 5х5, кількість карт згортки в першому шарі - 6, кількість карт згортки у другому шарі – 16, кількість нейронів у першому повнозв'язному шарі – 120, кількість нейронів у другому повнозв'язному шарі – 84.

ЗНМ була реалізована за допомогою пакету прикладних програм MATLAB 2018b. Для її навчання використано БД відфільтрованих зразків КП, що відповідають трьом зазначеним емоціям для 8 осіб. Кожна емоція представлена 80 записами (10 записів на людину) КП для одного і того ж самого тексту. Виклик певної емоції у випробуваних осіб реалізовано з допомогою перегляду відповідного мультимедійного контенту.

Зразок КП для емоції кожного користувача записаний у форматі: введений символ-ТУК-ТМК. 90% записів БД були використані для формування навчальної вибірки, решта 10% - для тестової.

Побудована за результатами проведених експериментів діаграма точності розпізнавання особи та емоційного стану слухача СДН, показана на рис. 5.25.

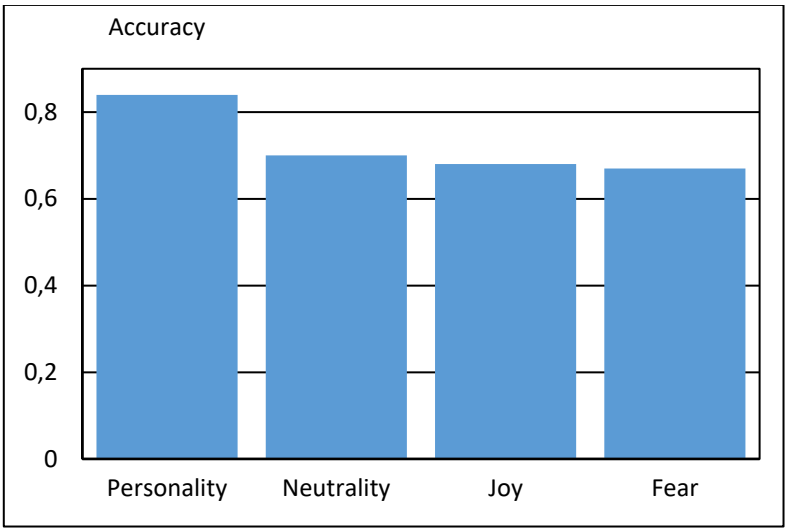

Рис. 5.25. Діаграма точності розпізнавання особи та емоційного стану за клавіатурним почерком

Зазначимо, що хоча під час проведення експериментів у цьому дослідженні використані класичні ЗНМ, отримані похибки співвідносяться з похибками відомих систем аналогічного призначення [169, 219, 252, 306]. При цьому на точність розпізнавання задіяної при експериментах НМЗ негативно вплинув відносно невеликий обсяг навчальної вибірки, формування якої є досить ресурсоємною процедурою. Це дозволяє стверджувати, що достатньо низьку помилку розпізнавання досягнуто завдяки використанню запропонованого методу формування вхідного поля ЗНМ, призначеної для аналізу КП. Крім цього, проведені експерименти показали можливість інтегрального розпізнавання емоцій та особи слухача СДН.

Другий етап експериментів присвячений оцінці ефективності НМЗ на базі ЗНМ типу AlexNet у задачі розпізнавання особи слухача СДН за його КП. За виключенням розпізнавання емоцій умови застосування НМЗ на базі AlexNet не

відрізнялись від умов застосування НМЗ на базі LeNet-5.

Оскільки класичний варіант AlexNet передбачає можливість розпізнавання 1000 класів, використано модифікований варіант AlexNet, адаптований до розпізнавання особи 8 слухачів СДН. Адаптація реалізована за рахунок модифікації двох останніх шарів НММ. Структура адаптованої мережі AlexNet, візуалізована за допомогою засобів MATLAB R2018b, показана на рис. 5.26.

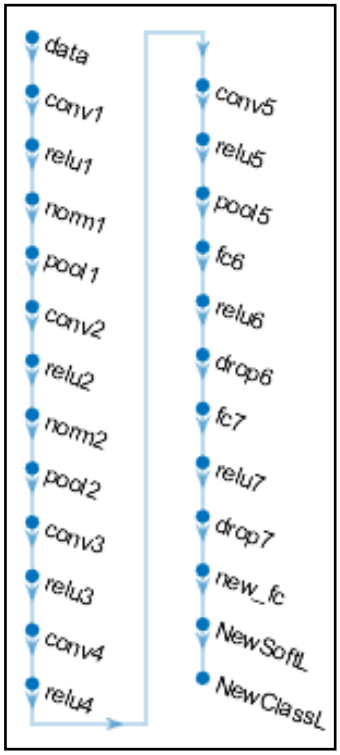

Рис. 5.26. Структура модифікованої нейронної мережі AlexNet

План другого етапу експериментів передбачав визначення показників точності розпізнавання та крос-ентропії на навчальній та валідаційній вибірці у випадках формування вхідного поля AlexNet з використанням: введеного символу та ТУК, введеного символу та ТМК, введеного символу, ТУК та ТМК.

Так, на рис. 5.27 и на рис. 5.28 показані графіки залежностей показників точності (Accuracy) та крос-ентропії (Loss) від кількості епох навчання. З метою усунення шумової складової процесу навчання для тренувальної вибірки також

розраховувалися згладжені значення цих показників Accuracy та Loss, графіки яких також показано на рис. 5.27 та рис. 5.28.

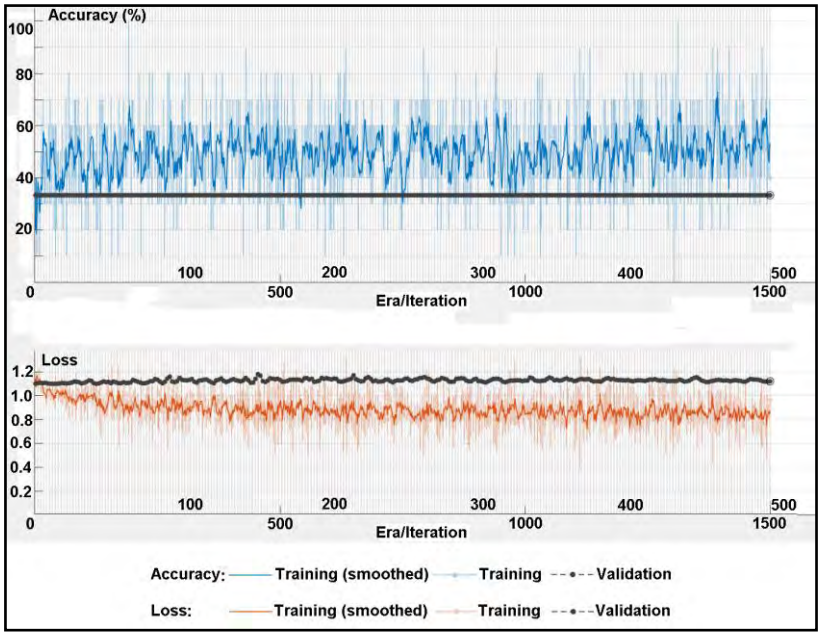

Рис. 5.27. Графіки залежності точності розпізнавання та крос-ентропії від кількості епох навчання при використанні ТУК

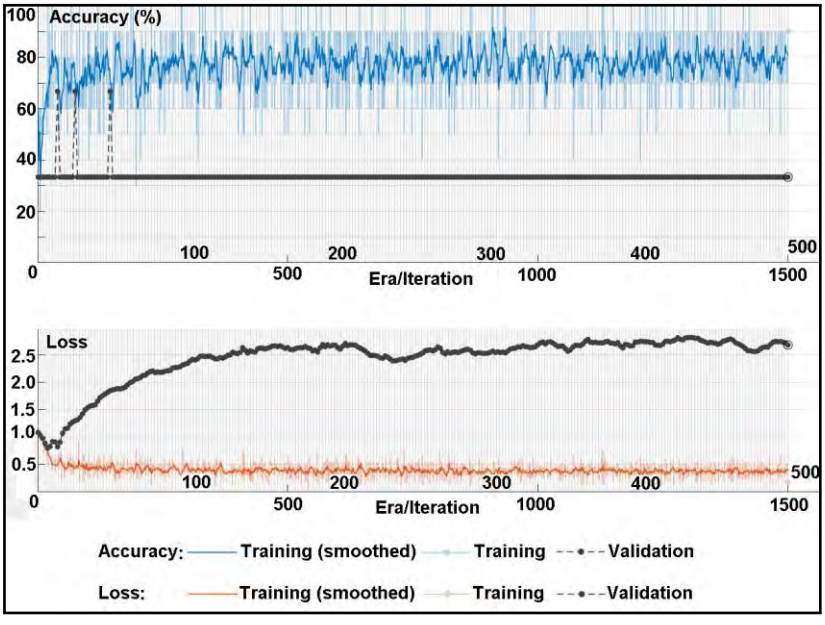

Рис. 5.28. Графіки залежності точності розпізнавання та крос-ентропії від кількості епох навчання при використанні ТУК та ТМК

Результати експериментів показують, що:

- Вхідне поле AlexNet доцільно формувати з одночасним використанням введеного символу та показників ТУК і ТМК.

- Ефективність розпізнавання AlexNet валідаційних прикладів дуже низька.

- На тренувальній вибірці показники точності та крос-ентропії стабілізуються приблизно після 50 епохи навчання. При цьому у випадку використання ТУК і ТМК Accuracy перевищує 80%, а значення Loss становить біля 0,2.

Аналіз отриманих результатів та їх порівняння з [122, 130, 182, 203] дозволяють стверджувати, що на тренувальних даних за показником Accuracy адаптована до поставленої задачі НММ типу AlexNet приблизно на 5% перевищує НММ типу LeNet та співмірна з моделлю SVM. У той же час при співмірній точності термін навчання AlexNet приблизно в 20 разів менший, ніж термін навчання вказаних моделей та моделей на основі рекурентних нейронних мереж [210]. Вкрай негативні результати розпізнавання валідаційних прикладів при розпізнаванні особи слухача СДН можна пояснити низькою репрезентативністю та недостатнім обсягом навчальної вибірки, що використовується, а також недостатньою адаптацією параметрів AlexNet до умов поставленої задачі. Тому шляхи подальших досліджень необхідно співвіднести з формуванням навчальної вибірки, що забезпечує якісне навчання НММ. Крім цього, доцільно розробити метод адаптації архітектурних параметрів AlexNet до умов завдання аналізу КП для розпізнавання особи та емоцій користувача.

Оцінка ефективності НМЗ аналізу ображення обличчя. Оскільки основною відмінністю запропонованих НММ аналізу зображення обличчя для розпізнавання емоційного стану являється використання капсульної НМ, адаптованої до розпізнавання емоцій (CapsNetEm), то експерименти акцентовано на оцінку ефективності саме цього типу ЗНМ. Зазначимо, що створення НММ є результатом реалізації 2-5 етапів запропонованої у п. 4.4 методології. Для навчання

та тестування НММ використано набори даних, які були сформовані на основі БД Fer2013-images, доступної на сайті www.kaggle.com. У цій БД представлені 35494 jpg-файли із фотографіями обличчя людей, що виражають 6 базових емоцій та нейтральний стан. Є фотографії, на яких особа зафіксована у фронтальній проекції, а також фотографії, на яких особа зафіксована у поверненому стані. Кут повороту від -45 ° до +45 °. Кожна фотографія є зображенням у відтінках сірого з роздільною здатністю 48×48 пікселів.

На першому етапі досліджень проведено експерименти по розпізнаванню емоцій на основі аналізу геометрії фронтального, добре освітленого зображення обличчя людини. Другий етап досліджень пов'язаний з розпізнаванням емоцій на основі аналізу зображень повернутого обличчя. Експерименти реалізовані за допомогою спеціально розробленої комп'ютерної програми, що базувалася на математичному апараті (2.72-2.81). Отримані графіки точності розпізнавання базових емоцій на фронтальному та поверненому зображенні обличчя показані на рис. 5.29 та рис. 5.30.

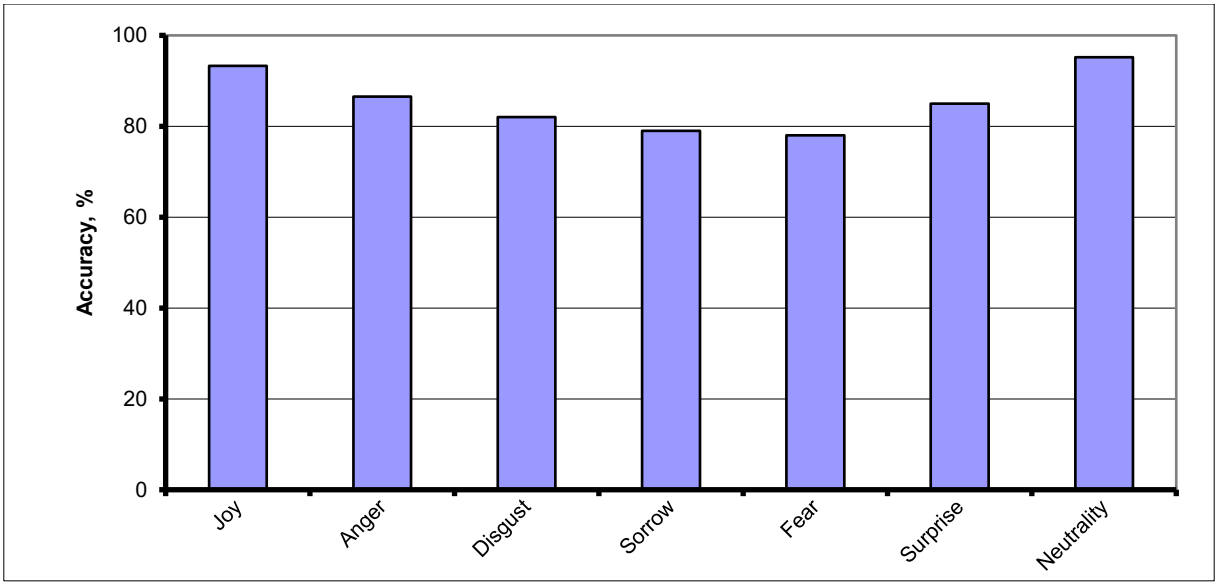

Рис. 5.29. Точність розпізнавання капсульною НМ різних емоцій на фронтальному зображенні обличчя

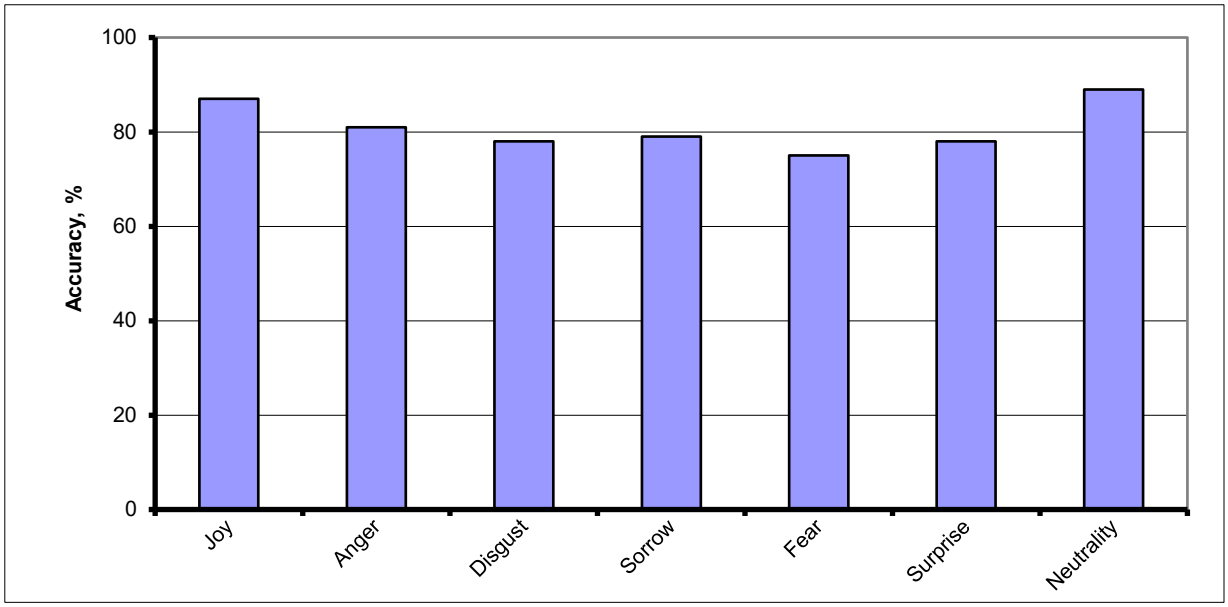

Рис. 5.30. Точність розпізнавання капсульною НМ різних емоцій на повернутому зображенні обличчя

Аналіз графіків, показаних на рис. 5.29 та на рис. 5.30, вказує на те, що при використанні фронтального зображення обличчя найменш точно розпізнаються емоції смуток (sorrow) і страх (fear). У випадку використання повернутих зображень - найгіршим чином розпізнаються емоції страх (fear) і здивування (surprise). Це можна пояснити не пропорційним представленням відповідних зображень обличчя у БД Fer2013-images. Також для перевірки гіпотези про достатню ефективність капсульної НМ в умовах, характерних для СДН, проведено порівняння точності розпізнавання побудованої НММ з: моделлю на базі нечіткої логіки (Fuzzy Logic) [192, 272]; моделлю на базі машини опорних векторів (SVM) [137, 223, 287]; НММ AUMPNET, що побудована на основі класичної ЗНМ типу LeNet [141]; рекурентною НМ на основі БШП (RNN) [148]; гібридної НММ на основі класичної ЗНМ типу LeNet та рекурентної НМ типу BLSTM (CNN and BLSTM) [183]. Результати порівняння зведені у табл. 5.3. Дані табл. 5.3 свідчать про те, що точність розпізнавання емоцій на основі і фронтальних і повернутих зображень обличчя капсульною НМ помітно вища, ніж точність розпізнавання НММ, побудованих з

використанням класичних ЗНМ на основі LeNet, може бути співставлена з точністю розпізнавання моделей на основі рекурентної НМ, співмірна з моделями на базі машини опорних векторів і нечіткої логіки, та співмірна з точністю відомих комплексів розпізнавання емоцій.

Таблиця 5.3

Середня точність розпізнавання базових емоцій

Тип моделі	**Середня точність**	
	Для фронтальних зображень	Для повернутих зображень
Капсульна НМ	0,85	0,81
SVM	0,78-0,9	0,7-0,73
AUMPNET	0,72	0,68
RNN	0,82-0,93	0,78
CNN and BLSTM	0,74-0,77	0,7-0,73
Fuzzy Logic	0,94	0,8

Також можна відзначити, що в цілому точність розпізнавання НММ повернутих зображень приблизно на 4% нижча, ніж точність розпізнавання фронтальних зображень. При використанні інших моделей точність погіршується приблизно на 7-14%. Це свідчить про перспективність використання капсульної НМ для розпізнавання емоцій за повернутим зображенням обличчя. Подальші дослідження можуть бути пов'язані з удосконаленням модифікованої капсульної НМ у напрямку зменшення кількості навчальних ітерацій при забезпеченні прийнятної точності. Також проведені експерименти для оцінки ефективності НМЗ розпізнавання особи слухача СДН за зображенням обличчя. Для цього використано БД Yale Faces (www.kaggle.com). Побудована з використанням описаного в п. 4.2 методу ЗНМ показала точність розпізнавання тестових прикладів на рівні 0,95, що дозволяє стверджувати про ефективність запропонованих дисертаційних рішень.

5.3. Оцінка ефективності засобів нейромережевого аналізу додаткових біометричних параметрів

На даному етапі досліджень проведені комп'ютерні експерименти, пов'язані з оцінкою ефективності НМЗ, що призначені для аналізу додаткових БП з метою розпізнавання особи слухача СДН. Зауважимо, що результат такого розпізнавання можливо використовувати в ЗАРЕС для підвищення точності розпізнавання емоційного стану, оскільки в більшості випадків точність знеособленого розпізнавання нижча, ніж точність розпізнавання емоційного стану слухача, особа якого встановлена.

Як встановлено в другому розділі даної дисертаційної роботи, в ЗАРЕС СДН до додаткових БП відносяться відбитки пальців, вушна раковина, райдужна оболонка ока та сітківка ока. Характерною особливістю кожного із вказаних БП являється можливість асоціації такого параметру з зображенням, що відповідно до результатів четвертого розділу, зумовлює реалізацію нейромережевого аналізу за допомогою ЗНМ.

Для оцінки ефективності засобів нейромережевого аналізу відбитків пальців проведені експерименти по розпізнаванню відбитку вказівного пальця 10 користувачів. Для формування навчальної та тестової вибірки використано зображення відбитків, представлені в БД FVC2000.

Розмір аналізуємих відбитків пальців становив 300×300 пікселів. Отже, $a_0 \times a_0 = 300 \times 300$, $L_{in} = 89401$, $L_{out} = 10$. Використовуючи розроблену методологію, шляхом тривіальних підстановок у вирази (4.41, 4.42), розраховані такі параметри ЗНМ: $a_1 = 98$, $a_2 = 30$. При розрахунках враховано, що у розробленій ЗНМ вже визначено такі параметри: $(b \times b)_1 = (b \times b)_1 = (5 \times 5)$, $d_1 = d_2 = 3$, $K_{ls,min} = 2$, $L_{h,1} = 5$, $L_{h,2} = 48$, $r_1 = 1$, $r_2 = 0$. При цьому розрахунок параметрів r_1 і r_2 базувався на тому, що визначений за допомогою виразу (4.41)

розмір карти ознак $H_{CNN,k} = a_k$ у k-му згортковому шарі має бути цілим числом.

Побудова ЗНМ реалізована, виходячи з необхідності зведення двовимірної карти ознак розміру $a_2 \times a_2$ до вектору ознак. Використана аналогія з класичною ЗНМ. В результаті мережа була доповнена двома згортковими шарами та двома шарами підвибірки. Визначені параметри шарів: $(b \times b)_3 = (5 \times 5)$, $(b \times b)_4 = (4 \times 4)$, $d_3 = d_4 = 1$, $r_3 = r_4 = 0$, $m_1 = m_2 = 2$, $L_{h,3} = 48$, $L_{h,4} = 32$. При цьому відповідно до виразу (4.41) $a_3 = 26$, $a_4 = 10$, а розмір карт ознак першого та другого шару підвибірки $c_1 = 13$, $c_2 = 5$.

Зазначимо, що накладання п'ятого ядра згортки розміром $(b \times b)_5 = (5 \times 5)$ на карти четвертого шару підвибірки дозволяє зробити остаточну згортку вхідного поля до вектору ознак. Використовуючи вирази (4.43-4.45), отримано $L_{f,min} = 66$, $L_{f,max} = 1280$. Остаточно кількість нейронів у повнозв'язному шарі $L_{f,max} = 512$ вибрано за аналогією з [21], виходячи з балансу необхідних обчислювальних ресурсів та обсягу пам'яті мережі. Структурні параметри побудованої ЗНМ представлені в табл. 5.4.

Таблиця 5.4.

Структурні параметри згорткової нейронної мережі, адаптованої до розпізнавання відбитків пальців

Параметр	**Значення параметру**
1	*2*
Розмір вхідного поля	$(a_0 \times a_0) = (300 \times 300)$
Кількість вхідних нейронів	$L_{in} = 8940$
Кількість вихідних нейронів	$L_{out} = 10$
Кількість нейронів у повнозв'язному шарі	$L_f = 512$
Кількість згорткових шарів	$K_{ls} = 4$

Таблиця 5.4. (продовження)

1	*2*
Кількість карт ознак у першому, другому, третьому та четвертому згорткових шарах	$L_{h,1} = 5;\quad L_{h,2} = 48;\quad L_{h,3} = 48;$ $L_{h,4} = 32$
Кількість шарів підвибірки	$K_{ld} = 2$
Розмір першого, другого, третього та п'ятого ядра згортки	$(b \times b)_1 = (b \times b)_2 = (b \times b)_3 =$ $= (b \times b)_5 = (5 \times 5)$
Розмір четвертого ядра згортки	$(b \times b)_4 = (4 \times 4)$
Зміщення рецептивного поля при виконанні першої та другої процедури згортки	$d_1 = d_2 = 3$
Зміщення рецептивного поля при виконанні третьої та четвертої процедури згортки	$d_1 = d_2 = 1$
Розмір карти ознак для першого, другого, третього та четвертого згорткового шару	$(a_1 \times a_1) = (98 \times 98);(a_2 \times a_2) =$ $(30 \times 30);\ \ (a_3 \times a_3) = (26 \times 26);$ $(a_4 \times a_4) = (10 \times 10)$
Кількість доповнюючих нулів для першого, другого, третього та четвертого згорткового шару	$r_1 = 1; r_2 = r_3 = r_4 = 0$
Масштабний коефіцієнт для першого та другого шару підвибірки	$m_1 = m_2 = 2$
Розмір карт для першого та другого шару підвибірки	$c_1 = 13\,;\ c_2 = 5$

Побудована ЗНМ стала основною для розробки програмного комплексу Fingerprint Recognition, призначеного для розпізнавання відбитків пальців. Програма написана мовою програмування Python за допомогою бібліотеки TensorFlow. Система розпізнавання навчалася та тестувалася на вибірці з бази

даних FVC2000 DB1, що знаходиться у вільному доступі за посиланням http://bias.csr.unibo.it/fvc2000/db1.asp.

В експериментах розраховувалася точність розпізнавання тестових прикладів при різних кількостях епох навчання (t_e) та кількостях карт ознак в 1, 2, 3 і 4 згорткових шарах ($L_{h,1}$, $L_{h,2}$, $L_{h,3}$, $L_{h,4}$). Досліджено набір архітектур, що визначався комбінаціями параметрів $L_{h,1} \in [4,5,6]$, $L_{h,2} \in [34,48,58]$, $L_{h,3} \in [34,48,58]$, $L_{h,4} \in [22,32,42]$, $t_e \in [10, 30, 50, 70, 100]$. Для експериментів використаний комп'ютер із процесором Intel Core i7-7700 (3.6 – 4.2 ГГц) та обсягом оперативної пам'яті 16 ГБ, який функціонував під керуванням операційної системи Windows 10. Результати експериментів частково показані на рис. 5.31 та на рис. 5.32.

Максимальна точність розпізнавання, що дорівнює близько 0,95, досягнута за наведених у табл. 5.4 параметрах конфігурації ЗНМ. Отримані результати загалом відповідають найкращим показникам подібних систем [6-16] і вказують на перспективність подальших досліджень у цьому напрямку.

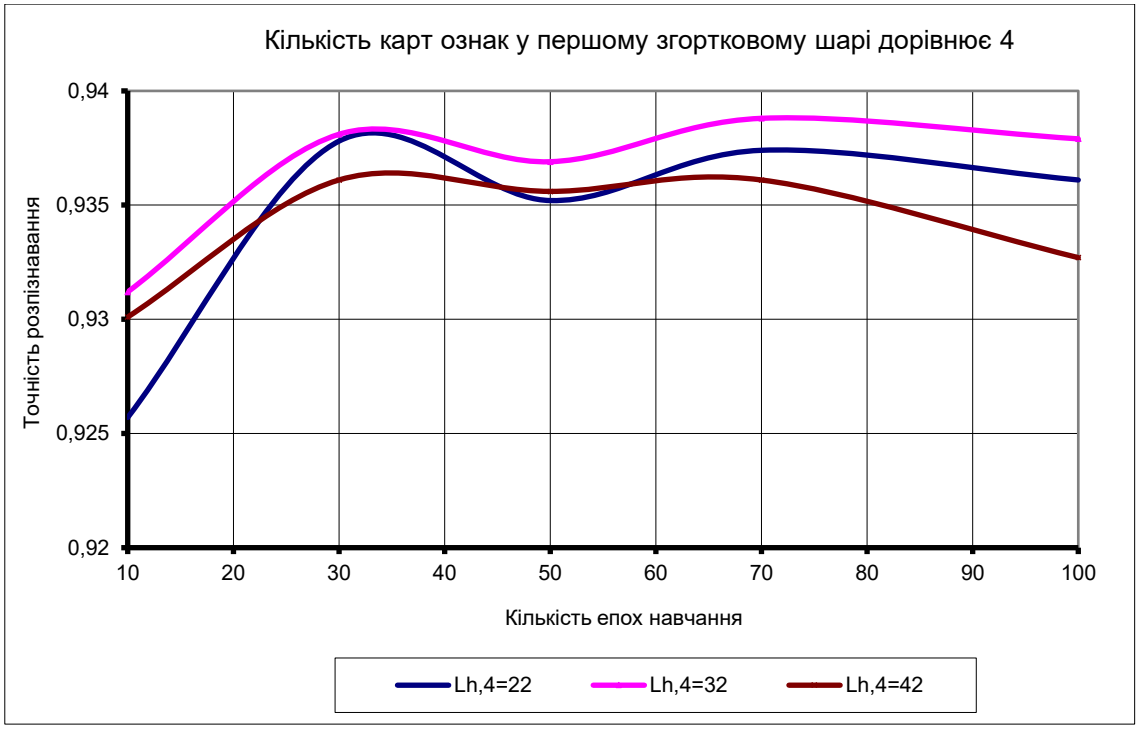

Рис. 5.31. Залежність точності розпізнавання від кількості епох навчання при $L_{h,1} = 4$, $L_{h,2} = 48$, $L_{h,3} = 48$ і різних значеннях $L_{h,4}$

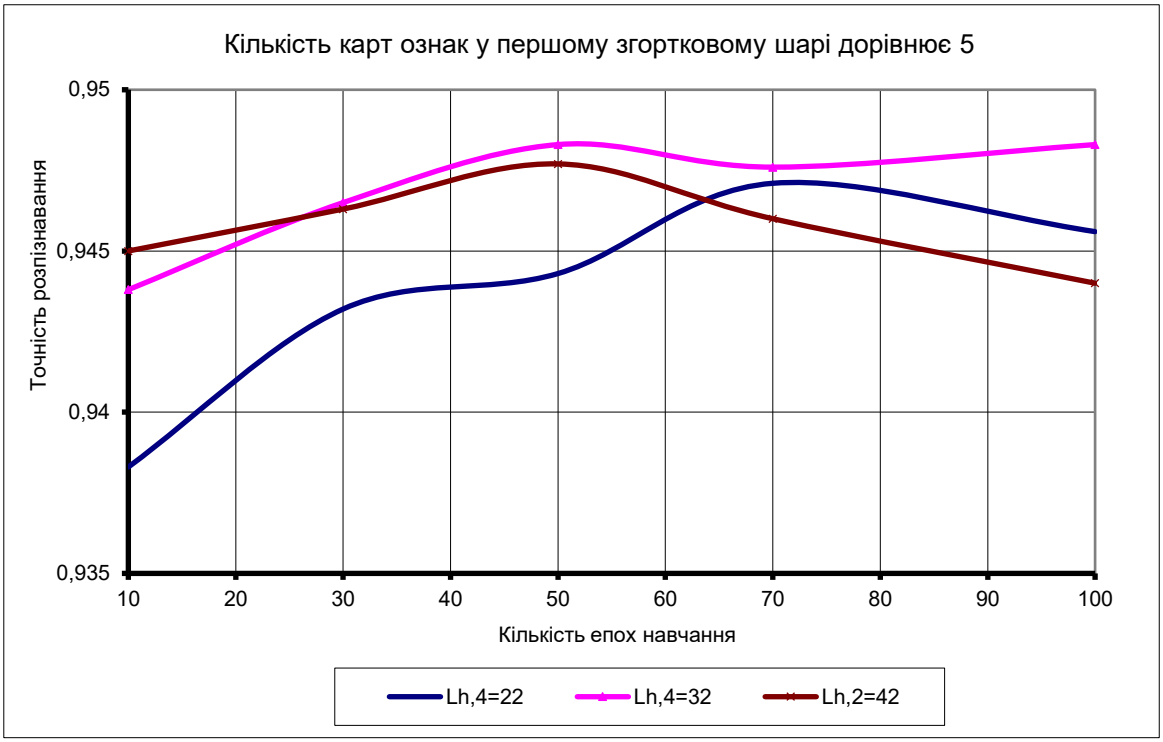

Рис. 5.32. Залежність точності розпізнавання від кількості епох навчання при $L_{h,1} = 5, L_{h,2} = 48, L_{h,3} = 48$ і різних значеннях $L_{h,4}$

На наступному етапі досліджень проведені експерименти, спрямовані на оцінку ефективності НМЗ розпізнавання особи слухача на основі нейромережевого аналізу зображення ушної раковини. Розмір аналізованих зображень вушних раковин становив 300×300 пікселів. Отже, $a_0 \times a_0 = 300 \times 300$, $L_{in} = 89401$, $L_{out} = 5$. Використовуючи запропонований в п. 4.2 метод розробки архітектури ЗНМ, призначеної для аналізу БП з використанням виразів (4.41, 4.42) розраховані такі параметри ЗНМ: $a_1 \times a_1 = 98 \times 98$, $a_2 \times a_2 = 30 \times 30$, $(b \times b)_1 = (b \times b)_2 = (5 \times 5)$, $d_1 = d_2 = 3$, $K_{ls,min} = 2$, $L_{h,1} = 5$, $L_{h,2} = 48$, $r_1 = 1$, $r_2 = 0$. Подальша побудова ЗНМ реалізована, виходячи з необхідності зведення двовимірної карти ознак до вектору ознак. Для цього створена мережа була доповнена двома шарами згортками і двома шарами підвибірки. Використавши запропоновані в п. 2.2 принципи адаптації ЗНМ до умов застосування, обрано наступні параметри шарів ЗНМ: $(b \times b)_3 = (5 \times 5)$, $(b \times b)_4 = (4 \times 4)$, $d_3 = d_4 = 1$, $r_3 = r_4 = 0$, $m_1 = m_2 =$

$2, L_{h,3} = 48, L_{h,4} = 32$. При цьому, відповідно до виразів (4.41, 4.42), визначено: $a_3 = 26, c_1 = 13, a_4 = 10, c_2 = 5$. Використавши (4.45), отримано $L_{f,min} = 66$. Остаточно кількість нейронів у повнозв'язному шарі $L_f = 300$ обрано, виходячи з балансу необхідних обчислювальних ресурсів та обсягу пам'яті мережі. Побудована ЗНМ стала основою програмного комплексу Ear Detection, призначеного для розпізнавання особи за зображенням вушної раковини. Система розпізнавання навчалася та тестувалася на вибірці, що була сформована автором. В експериментах розраховувалася точність розпізнавання тестових прикладів при різній кількості епох навчання (t_e) та різній кількості карт ознак ($L_{h,1}, L_{h,2}, L_{h,3}, L_{h,4}$). Використано АПЗ, як у випадку експериментів з розпізнаванням відбитків пальців. Результати експериментів частково показані на рис. 5.33 та на рис. 5.34.

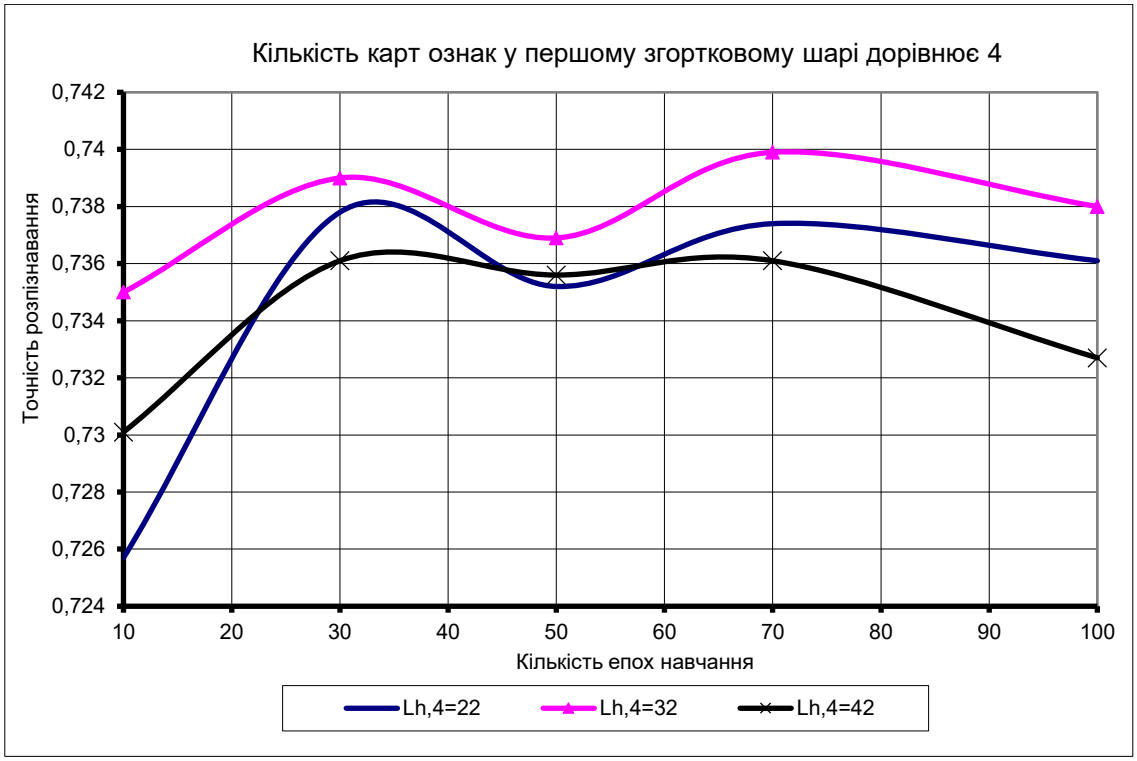

Рис. 5.33. Залежність точності розпізнавання від кількості епох навчання при $L_{h,1} = 4, L_{h,2} = 38, L_{h,3} = 38$ і різних значеннях $L_{h,4}$

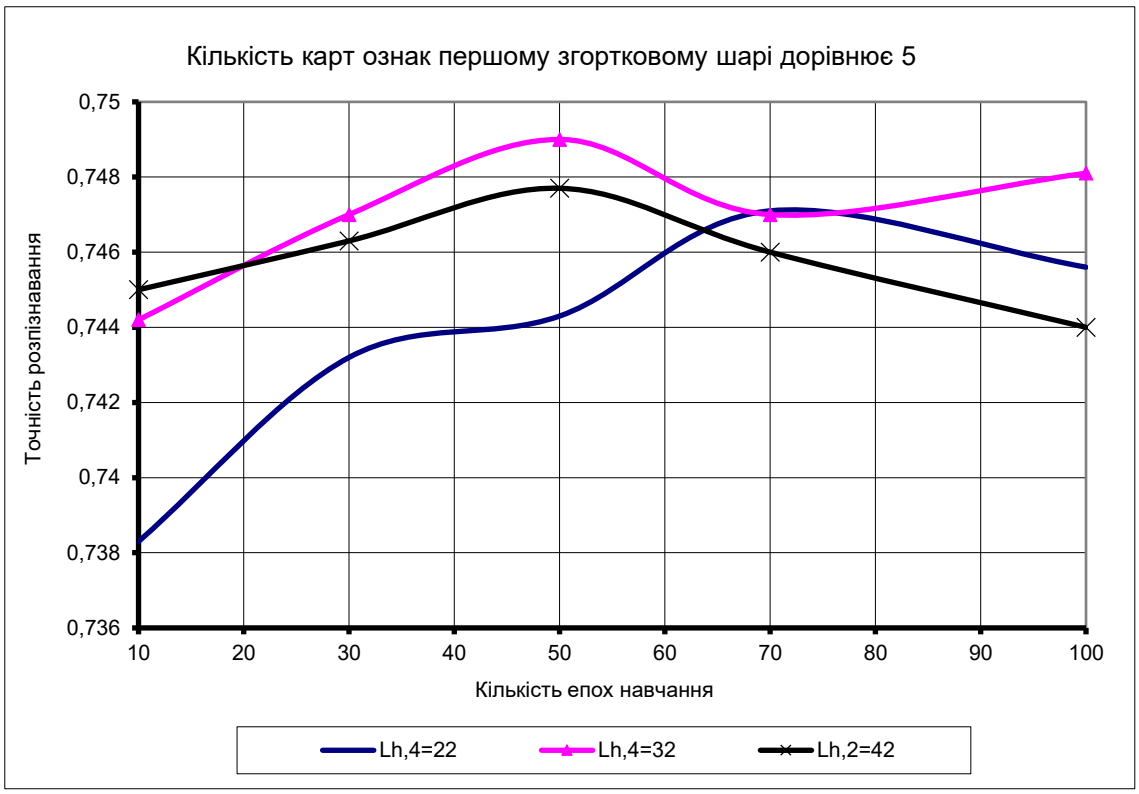

Рис. 5.34. Залежність точності розпізнавання від кількості епох навчання при $L_{h,1} = 5$, $L_{h,2} = 48$, $L_{h,3} = 48$ і різних значеннях $L_{h,4}$

Максимальна точність розпізнавання, що дорівнює близько 0,75, досягнута при наступних параметрах конфігурації ЗНМ: $L_{h,1} = 5$, $L_{h,2} = 48$, $L_{h,3} = 48$, $L_{h,4} = 32$, t_e=50. Отримані результати дозволяють підтвердити перспективи використання в системах біометричної аутентифікації додаткового модуля розпізнавання користувача за геометрією вушної раковини.

На заключному етапі досліджень, присвячених оцінці ефективності НМЗ аналізу додаткових БП, проведені експерименти, спрямовані на розпізнавання особи слухача СДН на основі райдужної оболонки ока та сітківки ока. Оскільки підходи до розробки та застосування засобів нейромережевого аналізу таких БП досить подібні, але для реєстрації райдужної оболонки ока достатньо загальновживаних відеокамер, а для реєстрації сітківки ока необхідні спеціалізовані засоби, то основна увага була зосереджена на розпізнаванні особи слухача СДН на

основі райдужної оболонки ока. Дане дослідження орієнтоване на нейромережеве розпізнавання райдужної оболонки ока 100 користувачів, 10000 зображень очей яких представлені в БД CASIA-Iris-Interval. Вказані зображення очей мають роздільну здатність 320х280 пікселів. На практиці отримане зображення райдужної оболонки ока буде частково обрізаним і викривленим внаслідок впливу типових завад, описаних в першому та другому розділах дисертаційної роботи. У базовому варіанті передбачається обрізані та викривлені частини зображення райдужної оболонки ока заповнити фоновим кольором.

Аналіз вказаних зображень дозволяє підтвердити, що зображення райдужки вписується в квадрат розміром 280х280 пікселів. Такий же розмір прийнятий і для вхідного поля ЗНМ, тобто $a_0 \times a_0 = 280 \times 280$. Таким чином, кількість вхідних параметрів ЗНМ $L_{in} = 78400$. Відповідно до визначених в п.2.2 принципів адаптації ЗНМ до умов застосування, шляхом експертного оцінювання визначено, що кількість рівнів розпізнавання райдужної оболонки дорівнює 4, що відповідає кількості шарів згортки $K_{ls} = 4$.

Також, відповідно до вказаних у п.2.2 принципів 2 і 3, проаналізувавши зображення райдужної оболонки, представлені в БД CASIA-Iris-Interval, визначено, що на першому рівні розпізнавання, за аналогією з даними [55, 101], доцільно використовувати 5 елементарних ознак розміром 5×5. Таким чином, $L_{h,1} = 5$, $(b \times b)_1 = 5 \times 5$, $d_1 = 1$. При цьому розмір ядра згортки може бути однаковим для всіх рівнів розпізнавання.

Прийняте припущення про відповідність карт ознак другого згорткового шару з 5 елементарними типами ознак, що характерні райдужній оболонці та по-різному орієнтовані у просторі. Враховуючи стійкість ЗНМ до розпізнавання зображень з кутом повороту ±15°, прийнято $L_{h,2} = (360/15) \times 5 = 120$.

Кількість карт ознак третього та четвертого згорткового шару визначено, виходячи з передумови необхідності розпізнавання чотирьох по-різному

орієнтованих комбінованих ознак райдужної оболонки.

Враховуючи стійкість ЗНМ до розпізнавання зображень з кутом повороту ±15°, $L_{h,3} = L_{h,4} = (360/15) \times 4$=96. Для всіх шарів субдискретизації обрано масштабний коефіцієнт $m = 4$. Інші представлені в табл. 5.5 конструктивні параметри ЗНМ розраховані відповідно до запропонованого методу розробки архітектури ЗНМ, призначеної для аналізу БП, за допомогою виразів (4.41-4.45) з урахуванням отриманих значень $a_0, L_{h,1}, L_{h,2}, L_{h,3}, L_{h,4}, b_1, b_2, b_3, b_4, \ d, m$.

Таблиця 5.5

Структурні параметри згорткової нейронної мережі

Параметр	**Значення параметру**
Розмір вхідного поля	280x280
Кількість вхідних нейронів	78400
Кількість вихідних нейронів	100
Кількість нейронів у повнозв'язному шарі	1000
Кількість згорткових шарів	4
Кількість карт ознак у першому, другому, третьому та четвертому згорткових шарах	5; 120; 96; 96
Кількість шарів підвибірки	4
Розмір ядра згортки для першого, другого та третього згорткового шару	5x5
Розмір ядра згортки для четвертого згорткового шару	4x4
Зміщення рецептивного поля	1
Розмір карти ознак для першого, другого, третього та четвертого згорткового шару	276; 68; 16; 16
Кількість повнозв'язних шарів нейронів	1
Масштабний коефіцієнт для шарів підвибірки	4

Побудована ЗНМ стала основною для розробки програмного комплексу Iris Recognition. Програма написана мовою програмування Python за допомогою бібліотеки TensorFlow. В експериментах розраховувалася точність розпізнавання для тестових прикладів для різних конфігурацій ЗНМ, що навчались на протязі 300 епох. Експерименти проводилися з використанням тих самих апаратних засобів, що використовувалися для оцінки ефективності НМЗ аналізу інших додаткових БП. Подібна ЗНМ використана в НМЗ аналізу сітківки ока. Максимальна точність розпізнавання тестових прикладів, що дорівнює близько 0,95, досягнута за наведених у табл. 5.5 параметрів конфігурації ЗНМ. Отримані результати точності розпізнавання особи слухача за райдужною оболонкою ока та за сітківкою ока в цілому співвідносяться з заявленою точністю розпізнавання особи подібних відомих систем біометричної аутентифікації користувачів ІС загального призначення, що свідчить про ефективність запропонованих рішень.

5.4. Оцінка ефективності системи автоматизованого розпізнавання емоційного стану та особи слухача

На заключному етапі досліджень реалізована оцінка ефективності розробленої системи автоматизованого розпізнавання емоційного стану та особи слухача СДН. Для цього використано апробований в [42, 58, 59] підхід до оцінки ефективності нейромережевих систем розпізнавання та визначені в першому розділі параметри оцінки ефективності ЗАРЕС. Відповідно до вказаного підходу, розрахувати інтегральну оцінку ефективності нейромережевих ЗАРЕС можливо за допомогою виразу:

$$E_{\Sigma} = \beta_A E_A + \beta_B E_B + \beta_C E_C + \beta_D E_D, \quad (5.5)$$

де E_{Σ} – інтегральний критерій ефективності ЗАРЕС;

E_A, E_B, E_C, E_D – критерії, що визначають ефективність ЗАРЕС щодо групи вимог A, B, C, D, відповідно;

$\beta_A, \beta_B, \beta_C, \beta_D$ – вагові коефіцієнти критеріїв E_A, E_B, E_C, E_D.

При цьому $\beta_A + \beta_B + \beta_C + \beta_D = 1$.

Розрахунок значень критеріїв E_A, E_B, E_C, E_D, які визначають ефективність ЗАРЕС відповідно до вимог кожної із груп A-D, можливо реалізувати так:

$$E_A = \sum_{i=1}^{I} \gamma_{i,A} E_{i,A} = \sum_{i=1}^{I} \left(\gamma_{i,A} \sum_{n=1}^{N_A} \alpha_{A,i,n} A_{i,n} \right), \quad (5.6)$$

$$E_B = \sum_{i=1}^{I} \gamma_{i,B} E_{i,B} = \sum_{i=1}^{I} \left(\gamma_{i,B} \sum_{n=1}^{N_{B,i}} \alpha_{B,i,n} B_{i,n} \right), \quad (5.7)$$

$$E_C = \sum_{i=1}^{I} \gamma_{i,C} E_{i,C} = \sum_{i=1}^{I} \left(\gamma_{i,C} \sum_{n=1}^{N_{C,i}} \alpha_{C,i,n} C_{i,n} \right), \quad (5.8)$$

$$E_D = \sum_{i=1}^{I} \gamma_{i,D} E_{i,D} = \sum_{i=1}^{I} \left(\gamma_{i,D} \sum_{n=1}^{N_{D,i}} \alpha_{D,i,n} D_{i,n} \right), \quad (5.9)$$

де $E_{i,A}$, $E_{i,B}$, $E_{i,C}$, $E_{i,D}$ – критерії ефективності, що стосується і-го БП, розраховані на основі групи параметрів A-D відповідно;

$\alpha_{A,i,n}$, $\alpha_{B,i,n}$, $\alpha_{C,i,n}$, $\alpha_{D,i,n}$ – вагові коефіцієнти n-го параметру ефективності, що відносяться до групи A-D при використанні і-го БП;

$\gamma_{i,Z}$ – ваговий коефіцієнт z-го критерію ефективності при використанні і-го БП, де $z \in \{A, B, C, D\}$;

N_A – кількість параметрів ефективності в групі А;

$N_{B,i}, N_{C,i}, N_{D,i}$ – кількість параметрів ефективності в кожній із груп B-C при використанні і-го БП;

$A_{i,n}, B_{i,n}, C_{i,n}, D_{i,n}$ – n-ий параметр ефективності однієї із груп A-D при використанні і-го БП.

Зазначимо, що $N_A = 6$, а кількість параметрів ефективності, що групах B,C,D залежить від виду БП, що підлягає аналізу.

У виразах (5.6-5.9) кількість БП дорівнює 3. При цьому номер БП, що асоціюються з зображенням, дорівнює 1, номер БП, що асоціюється з голосом слухача, дорівнює 2, а номер БП, що асоціюється з КП, дорівнює 3.

Таким чином, з урахуванням (5.5-5.9), вираз для оцінки ефективності ЗАРЕС в розгорнутому вигляді записується так:

$$E_{\Sigma} = \beta_A \sum_{i=1}^{3} \left(\gamma_{i,A} \sum_{n=1}^{N_A} \alpha_{A,i,n} A_{i,n} \right) + \beta_B \sum_{i=1}^{3} \left(\gamma_{i,B} \sum_{n=1}^{N_{B,i}} \alpha_{B,i,n} B_{i,n} \right) + \\ + \beta_C \sum_{i=1}^{3} \left(\gamma_{i,C} \sum_{n=1}^{N_{C,i}} \alpha_{C,i,n} C_{i,n} \right) + \beta_D \sum_{i=1}^{3} \left(\gamma_{i,D} \sum_{n=1}^{N_{D,i}} \alpha_{D,i,n} D_{i,n} \right). \quad (5.10)$$

При цьому для відомих ЗАРЕС значення параметрів ефективності, що відносяться до групи А, наведено в табл. 1.3, 1.5, 1.7. Для запропонованої нейромережевої САРЕС отримані за допомогою методів експертного оцінювання значення параметрів ефективності групи А (базові параметри), показані в табл. 5.6.

Таблиця 5.6

Оцінка значень базових параметрів ефективності для запропонованої ЗАРЕС

Номер БП	**Значення базового параметру**					
	A_1	A_2	A_3	A_4	A_5	A_6
1	1	1	1	0,9	0,9	0,9
2	1	1	1	0,8	0,9	1
3	1	1	1	0,6	0,9	0,6

Розраховані значення параметрів ефективності групи B-D, що характеризують особливості розпізнавання, особливості технічної реалізації та сервісні можливості ЗАРЕС, які базуються на різних БП відомих та розробленої ЗАРЕС показані в табл. 5.7-5.11. Зазначимо, що перелік та пояснення параметрів

оцінки ефективності, які використано в табл. 5.7 наведені в табл. 1.2. При цьому в табл. 5.2 використана наступна специфікація ЗАРЕС:

- ZO_1 – Affectiva Affdex;
- ZO_2 – Pupeteer;
- ZO_3 – CLMtrackr;
- ZO_4 – Oxford;
- ZO_5 – FaceReader;
- ZO_6 – Emovu;
- ZO_7 – NuraLogix;
- ZO_8 – eMotion Software;
- ZO_9 – MMER_FEASy;
- ZO_{10} – FaceSecurity;
- ZA – запропонована СЗРЕС.

Аналіз вказаних ЗАРЕС наведено в п.1.2.

Таблиця 5.7

Значення додаткових параметрів оцінки ефективності ЗАРЕС за зображенням обличчя (БП=1)

Пара-метр	**Специфікація ЗАРЕС за зображенням обличчя**										
	ZO_1	ZO_2	ZO_3	ZO_4	ZO_5	ZO_6	ZO_7	ZO_8	ZO_9	ZO_{10}	**ZA**
1	2	3	4	5	6	7	8	9	10	11	12
$B_{1,1}$	0,8	0,7	0,7	0,8	0,8	0,7	0,8	0,7	0,7	0,7	0,8
$B_{1,2}$	0,7	0,7	0,5	0,8	0,3	0,6	0,7	0,5	0,6	0,4	0,8
$B_{1,3}$	0,2	0,2	0,5	0,4	0,5	0,3	0,3	0,3	0,3	0,3	0,8
$B_{1,4}$	0,2	0,3	0,5	0,5	0,5	0,3	0,7	0,3	0,4	0,4	0,5
$B_{1,5}$	0,6	0,8	0,5	0,6	0,6	0,4	0,7	0,4	0,5	0,4	0,6
$B_{1,6}$	0,6	0,8	0,5	0,6	0,6	0,4	0,7	0,4	0,5	0,4	0,6
$B_{1,7}$	0,3	0,3	0,3	0,9	0,3	0,3	0,2	0,5	0,6	0,6	0,5

Таблиця 5.7(продовження)

1	2	3	4	5	6	7	8	9	10	11	12
$B_{1,8}$	1	1	1	1	1	1	0,5	0,8	1	1	1
$B_{1,9}$	0,8	0,8	0,9	0,9	0,9	0,8	0,8	0,9	0,8	0,9	0,9
$B_{1,10}$	0,8	0,8	0,9	0,9	0,8	0,7	0,7	0,7	0,7	0,7	0,9
$B_{1,11}$	0,8	0,8	0,9	0,9	0,8	0,7	0,7	1	0,8	0,8	0,9
$B_{1,12}$	1	1	1	1	1	1	1	0,5	1	1	1
$B_{1,13}$	0	0	0	0	0	0,4	0	1	0	0	0
$B_{1,14}$	0,7	0,7	0,8	0,9	0,8	0,7	0,7	0,7	07,	0,7	0,8
$C_{1,1}$	0	0	0	1	1	1	0	0	1	1	1
$C_{1,2}$	0,5	0,5	1	1	1	0,6	0,1	0	1	1	1
$C_{1,3}$	0	0	0	1	1	1	0	0	0	0	1
$C_{1,4}$	1	1	1	1	1	0	0	1	1	1	1
$D_{1,1}$	0	0,9	0	0,9	0,9	0	0	0	0,9	0	0,9
$D_{1,2}$	0	0	0	0	0	0	0	1	0	0	0
$D_{1,3}$	0	0	0	0	0,9	0	0	0	0,8	0	0
$D_{1,4}$	0	0	0	0	0,9	0	0	0	0,8	0	0

Аналіз даних табл. 5.7 свідчить про те, що можливості запропонованої САРЕС знаходяться на рівні найкращих рішень в області розпізнавання емоцій за зображенням обличчя, що пояснюється застосуванням запропонованих рішень для обробки зображення обличчя та адаптацією НММ до очікуваних умов застосування.

Перелік параметрів оцінки ефективності, які використано в табл. 5.8 наведені в табл. 1.4. В табл. 5.8 використана наступна специфікація ЗАРЕС:

- $ZГ_1$ – комп'ютерна система Вищої школи економіки;
- $ZГ_2$ – Beyond Verbal;

- $ZГ_3$ – Xpression;
- $ZГ_4$ – системи розпізнавання емоцій в мові людини;
- $ZГ_5$ – програма для вимірювання емоцій абонента;
- $ZГ_6$ – EmoWatch;
- ZA – запропонована САРЕС.

Аналіз вказаних ЗАРЕС наведено в п.1.3.

Таблиця 5.8

Значення додаткових параметрів оцінки ефективності ЗАРЕС за голосом (БП=2)

№ параметру	**Специфікація ЗАРЕС за голосом**						
	$ZГ_1$	$ZГ_2$	$ZГ_3$	$ZГ_4$	$ZГ_5$	$ZГ_6$	**ZA**
$B_{2,1}$	0,2	0,5	0,4	0,4	0,2	0,9	0,9
$B_{2,2}$	0,5	0,5	0,6	0,4	0,3	0,5	0,5
$B_{2,3}$	0,5	0,5	0,5	0,4	0,3	0,5	0,5
$B_{2,4}$	0,4	0,7	0,6	0,5	0,4	0,4	0,9
$B_{2,5}$	0,7	0,9	0,5	0,6	0,5	0,6	0,9
$B_{2,6}$	0,5	0,7	0,4	0,4	0,2	0,4	0,5
$B_{2,7}$	0,4	0,8	0,3	0,2	0,8	0,8	0,7
$B_{2,8}$	0,4	0,7	0,6	0,4	0,7	0,8	0,7
$B_{2,9}$	0	0	0	1	1	0	0
$C_{2,1}$	0,2	0,7	0,%	0,3	0,3	0,7	0,9
$C_{2,2}$	0,5	0,7	0,6	0,2	0,4	0,7	0,9
$C_{2,3}$	0,5	0,7	0,2	0,4	0,3	0,7	0,9
$D_{2,1}$	0	0,8	0,8	0	0	0,8	0,8
$D_{2,2}$	0	0	0,5	0	0	0	0
$D_{2,3}$	0	0	0	0	0	0,8	0
$D_{2,4}$	0	0	0,5	0	0	0,4	0

Перелік та пояснення параметрів оцінки ефективності, які використано в табл. 5.9 наведені в табл. 1.6. При цьому в табл. 5.9 використана наступна специфікація ЗАРЕС:

- ZK_1 - метод репрезентативного нейромережевого розпізнавання емоцій;
- ZK_2 - метод двосторонньої валентності;
- ZK_3 - лінійної регресивної моделі;
- ZK_4 - метод розпізнавання емоційного контексту;
- ZK_5 - метод аналізу КП за допомогою LSTM;
- ZA - запропонована САРЕС.

Аналіз вказаних ЗАРЕС наведено в п.1.4.

Таблиця 5.9

Значення додаткових параметрів оцінки ефективності ЗАРЕС за клавіатурним почерком (БП=3)

№ параметру	**Специфікація ЗАРЕС за клавіатурним почерком**					
	ZK_1	ZK_2	ZK_3	ZK_4	ZK_5	**ZA**
$B_{3,1}$	0,5	0,3	0,4	0,4	0,5	0,6
$B_{3,2}$	0,5	0,3	0,4	0,4	0,5	0,7
$B_{3,3}$	0	0	0	0	0	0
$B_{3,4}$	0	0	0	0	0	0
$C_{3,1}$	0,4	0,4	0,4	0,4	0,4	0,9
$C_{3,2}$	0,5	0	0	0	0	0,9
$D_{3,1}$	0	0	0	0	0	0,7

Визначення, показаних в табл. 5.6-5.9 показників ефективності дозволило перейти до розрахунку значень критеріїв ефективності. Для цього використано вирази (5.5-5.10). Так в табл. 5.10-5.12 показано значення критеріїв $E_{1,A}$, $E_{2,A}$, $E_{3,A}$,

що відображають ефективність відомих та розробленої ЗАРЕС при використанні кожного із основних БП відповідно.

Таблиця 5.10

Значення базового критерію ефективності при використанні БП=1

Критерій	**Специфікація ЗАРЕС за зображенням обличчя**										
	ZO_1	ZO_2	ZO_3	ZO_4	ZO_5	ZO_6	ZO_7	ZO_8	ZO_9	ZO_{10}	**ZA**
$E_{1,A}$	0,52	0,9	0,57	0,60	0,58	0,92	0,73	0,95	0,58	0,95	0,95

Таблиця 5.11

Значення базового критерію ефективності при використанні БП=2

Критерій	**Специфікація ЗАРЕС за голосом**						
	$ZГ_1$	$ZГ_2$	$ZГ_3$	$ZГ_4$	$ZГ_5$	$ZГ_6$	**ZA**
$E_{2,A}$	0,73	0,85	0,55	0,67	0,67	0,7	0,95

Таблиця 5.12

Значення базового критерію ефективності при використанні БП=3

Критерій	**Специфікація ЗАРЕС за клавіатурним почерком**					
	$ZК_1$	$ZК_2$	$ZК_3$	$ZК_4$	$ZК_5$	**ZA**
$E_{3,A}$	0,57	0,57	0,4	0,4	0,78	0,85

Зазначимо, що дані табл. 5.10-5.12 отримані за умови однакових значень вагових коефіцієнтів параметрів ефективності ЗАРЕС. Аналіз отриманих даних вказує на те, що ефективність розробленої САРЕС знаходиться приблизно на однаковому рівні із відомими ЗАРЕС на основі зображення обличчя, та приблизно в 1,1 рази перевищує ефективність ЗАРЕС на основі голосу та КП.

Використавши дані табл. 5.10-5.12, для відомих та запропонованої ЗАРЕС

розраховано значення базового критерію ефективності щодо групи основних вимог (вимог групи А). Для розрахунку використано вираз (5.6) з урахування того, що $\gamma_{1,A} = 0{,}5, \gamma_{2,A} = 0{,}3, \gamma_{3,A} = 0{,}2$.

Вказані величини вагомості критеріїв ефективності визначені з позицій апробованості ЗАРЕС на основі відповідних БП. Отримані значення критерію E_A різних ЗАРЕС за умови одночасного використання всіх основних БП наведені в табл. 5.13.

Таблиця 5.13

Значення базового критерію ефективності при одночасному використанні основних БП

ЗАРЕС	ZO_1	ZO_2	ZO_3	ZO_4	ZO_5	ZO_6	ZO_7	ZO_8	ZO_9	ZO_{10}	$ZГ_1$
E_A	0,26	0,45	0,29	0,30	0,29	0,46	0,37	0,48	0,29	0,48	0,22
ЗАРЕС	$ZГ_2$	$ZГ_3$	$ZГ_4$	$ZГ_5$	$ZГ_6$	$ZК_1$	$ZК_2$	$ZК_3$	$ZК_4$	$ZК_5$	**ZA**
E_A	0,26	0,17	0,20	0,20	0,21	0,11	0,11	0,08	0,08	0,16	0,98

Дані табл. 5.13 свідчать про те, що величина базового критерію ефективності розробленої САРЕС приблизно в 2 рази вища, ніж у найкращих відомих ЗАРЕС. Це пояснюється тим, що у всіх проаналізованих ЗАРЕС розпізнавання емоцій реалізується на основі одного БП. При цьому слід зазначити, що у відкритих джерелах згадується декілька ЗАРЕС, що реалізують розпізнавання емоцій на основі аналізу множини БП, до якої, крім інших, входять зображення обличчя та голос слухача СДН.

Найбільш відомими ЗАРЕС такого типу являються HireVue (www.hirevue.com) та Pymetrics (www.pymetrics.ai). Однак вказані ЗАРЕС мають закритий характер, використані в них методи розпізнавання емоцій не описані, а відповідно, не доступні для застосування у вітчизняних СДН. Разом з тим, вказані

ЗАРЕС призначені для використання в жорстко заданих умовах офісу для розпізнавання емоційного стану співробітників при проходженні ними анкетування. Відповідно слід очікувати, що ефективність HireVue та Pymetrics в умовах СДН не буде перевищувати значення $E_{1,A}$ та $E_{2,A}$ для запропонованої САРЕС. При цьому $E_{3,A}$ для цих систем дорівнює 0.

Оскільки для розробленої САРЕС $E_{3,A} = 0{,}85$, а $\gamma_{3,A} = 0{,}2$, можливо стверджувати, що для авторської САРЕС значення базового критерію ефективності приблизно в 1,16 разів перевищує аналогічне значення для найкращих ЗАРЕС подібного призначення.

Використавши дані табл. 5.7-5.10 за допомогою виразів (5.5-5.10) розраховано значення критеріїв ефективності ЗАРЕС для кожної із груп параметрів B-D. Аналіз цих результатів показує, що ефективність розробленої САРЕС щодо забезпечення кожної із додаткових вимог знаходиться на рівні найкращих відомих ЗАРЕС.

Підставивши отримані значення в вираз (5.5) визначено, що для розробленої ЗАРЕС $E_{\Sigma} = 0{,}82$, що приблизно в 1,1 рази перевищує E_{Σ} для інших відомих ЗАРЕС на основі БП, котрі можуть бути зареєстровані за допомогою розповсюдженого комп'ютерного забезпечення. Застосування подібного підходу дозволило визначити, що з позицій забезпечення можливості розпізнавання особи слухача ефективність розробленої нейромережевої системи практично не поступається відомим засобам біометричної аутентифікації користувачів ІС загального призначення.

Крім того, аналіз визначених показників та критеріїв ефективності дозволяє стверджувати, що перспективи вдосконалення розробленої САРЕС, в першу чергу, співвідносяться з впровадженням методів нейромережевого аналізу ГС та КП, що дозволяють реалізувати розпізнавання емоційного стану знеособленого слухача СДН.

5.5. Висновки до п’ятого розділу

В даному розділі вирішувалось науково-практичне завдання розроблення архітектури системи автоматизованого розпізнавання емоційного стану та проведення експериментальних досліджень, спрямованих на оцінку ефективності дисертаційних рішень. Основні результати розділу наступні:

1. На основі запропонованої методології розроблено архітектура системи автоматизованого розпізнавання емоційного стану та особи слухача СДН, в якій на відміну від відомих передбачено використання модулю визначення множини параметрів, що підлягають реєстрації, модулю визначення архітектури НММ, модулів, які дозволяють реалізувати інтегральний нейромережевий аналіз основних та додаткових біометричних параметрів, що забезпечує достатню точність розпізнавання і адаптацію до очікуваних умов застосування. Отримані архітектурні рішення послужили основою для розробки інструментальних засобів, призначених для нейромережевого аналізу основних та додаткових біометричних параметрів слухача системи дистанційного навчання.

2. Проведені експериментальні дослідження нейромережевих засобів аналізу основних біометричних параметрів дозволяють стверджувати, що використання запропонованих рішень дозволяє забезпечити точність розпізнавання емоційного стану ідентифікованого слухача СДН за голосом близько 0,95, а за клавіатурним почерком близько 0,76. Точність розпізнавання емоційного стану неідентифікованого слухача за фронтальним зображенням обличчя становить близько 0,85, а у випадку використання повернутого зображення близько 0,81. По відношенню до відомих, використання розроблених нейромережевих засобів дозволяє в 1,1-1,2 рази підвищити точність розпізнавання емоційного стану слухача. Точність розпізнавання особи слухача за допомогою засобів нейромережевого аналізу основних біометричних параметрів знаходиться в межах 0,85-0,95. Отримані результати свідчать, що точність розроблених нейромережевих

засобів є достатньою для їх впровадження в СДН. Винятком являються лише засоби для розпізнавання емоційного стану за клавіатурним почерком. Також експериментально доведено, що застосування запропонованого методу розробки архітектури ЗНМ дозволяє приблизно в 1,2 рази зменшити обсяг експериментів, пов'язаних з адаптацією архітектури ЗНМ до умов аналізу БП слухача СДН.

3. Проведені експериментальні дослідження нейромережевих засобів аналізу таких додаткових біометричних параметрів, як відбиток пальця, зображення вушної раковини, райдужної оболонки ока та сітківки ока дозволяють стверджувати що точність розпізнавання особи слухача СДН окремо за кожним із перерахованих параметрів становить 0,92-0,98, що відповідає точності розпізнавання найкращих аналогічних засобів біометричної аутентифікації користувачів інформаційних систем загального призначення. Комплексне використання засобів аналізу основних та додаткових біометричних параметрів забезпечує можливість резервування засобів аутентифікації слухачів СДН.

4. За допомогою методів експертного оцінювання визначено значення основних та додаткових параметрів оцінки ефективності систем розпізнавання емоційного стану та особи слухача СДН, що дозволило реалізувати розрахунок відповідних критеріїв ефективності.

5. Результати оцінювання ефективності запропонованої системи автоматизованого розпізнавання емоційного стану та особи слухача системи дистанційного навчання вказують на те, що при використанні всіх основних біометричних параметрів за критерієм відповідності до основних вимог розроблена система приблизно в 1,16 разів перевищує ефективність найкращих подібних систем.

ВИСНОВКИ

У ході виконання дисертаційної роботи вирішена актуальна науково-прикладна проблема розробки методології розпізнавання емоційного стану слухачів системи дистанційного навчання, що базується на нейромережевих методах і моделях аналізу біометричних параметрів, та орієнтована на створення відповідних інструментальних засобів, які розширюють функціональні можливості системи дистанційного навчання.

Проведені дослідження дозволяють зробити наступні висновки:

1. Виконано аналіз передумов та перспектив впровадження технології розпізнавання емоційного стану в системи дистанційного навчання. Аналіз довів актуальність науково-прикладної проблеми створення методології розпізнавання емоційного стану слухачів системи дистанційного навчання, вирішення якої має значний практичний ефект.

2. Побудовано концептуальну модель розпізнавання емоційного стану слухачів системи дистанційного навчання, оригінальність якої забезпечується конкретизацією операцій нейромережевого аналізу біометричних параметрів, а також обґрунтованим переліком параметрів оцінки ефективності розпізнавання.

3. Удосконалено підходи до визначення конструктивних параметрів, навчання та мінімізації ресурсоємності нейромежевої моделі, що забезпечило можливість розробки ефективних моделей і методів обробки біометричних параметрів та методів нейромережевого розпізнавання емоційного стану та особи слухача системи дистанційного навчання.

4. Побудовано моделі обробки біометричних параметрів, асоційованих з зображеннями, які за рахунок інтегрованого опису процесів центрування, масштабування, бінаризації, виділення контуру, фільтрації і виділення контрольних точок дозволяють нівелювати типові завади і викривлення та забезпечують теоретичний базис розробки методу формування вхідного поля нейромережевих моделей аналізу біометричних параметрів, асоційованих з

зображеннями.

5. Розроблено метод застосування вейвлет-перетворень для фільтрації вхідного поля нейромережевої моделі аналізу біометричних параметрів, асоційованих з зображеннями, в якому передбачено інтеграцію вейвлет-коефіцієнтів, розрахованих з використанням ефективного типу базисного вейвлету, для послідовно зареєстрованих зображень, що забезпечує можливість фільтрації завад, типових для системи дистанційного навчання.

6. Створено комплекс методів формування вхідного поля нейромережевих моделей аналізу біометричних параметрів, який передбачає використання процедур подання біометричних параметрів у вигляді придатному до обробки за допомогою згорткової нейронної мережі, що забезпечує можливість ефективного розпізнавання емоцій та особи слухача системи дистанційного навчання.

7. Розроблено метод формування навчальних прикладів, який дозволяє врахувати в очікуваному вихідному сигналі близькість еталонів заданих класів, що дозволяє в 1,3-1,5 рази зменшити кількість навчальних ітерацій для досягнення допустимої помилки навчання.

8. Побудовано метод розробки архітектури згорткової нейронної мережі, який дозволяє приблизно в 1,2 рази зменшити обсяг експериментів, пов’язаних з адаптацією значень конструктивних параметрів мережі до умов системи дистанційного навчання.

9. Розроблено метод нейромережевого розпізнавання емоцій слухача системи дистанційного навчання за одним із основних біометричних параметрів, який шляхом застосування запропонованих рішень щодо формування вхідного поля нейромережевих моделей для аналізу зображення обличчя, голосового сигналу та клавіатурного почерку і рішень щодо відображення у вихідному полі нейронної мережі близькості еталонів емоцій, дозволяє підвищити точність розпізнавання.

10. Розроблено методологію автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання, яка

забезпечуючи формалізоване використання запропонованих методів формування вхідного поля нейронної мережі аналізу голосового сигналу, клавіатурного почерку, біометричних параметрів, асоційованих з зображеннями, методів застосування вейвлет-перетворень для фільтрації вхідного поля, формування навчальних прикладів, розробки архітектури згорткової нейронної мережі та методу нейромережевого розпізнавання емоцій, що дозволяє створювати ефективні засоби автоматизованого розпізнавання емоційного стану слухача, які розширюють функціональні можливості системи дистанційного навчання.

11. Побудовано архітектуру системи автоматизованого розпізнавання емоційного стану слухача системи дистанційного навчання, ефективність якої приблизно в 1,16 разів перевищує ефективність найкращих подібних систем.

12. Проведені експериментальні дослідження нейромережевих засобів аналізу біометричних параметрів дозволяють стверджувати, що використання запропонованих рішень забезпечує можливість в 1,1-1,2 рази підвищити точність розпізнавання емоційного стану слухача відносно випадку використання відомих засобів аналогічного призначення.

Методологія розпізнавання емоційного стану слухачів системи дистанційного навчання створює сучасний науково-методологічний базис вдосконалення інструментарію дистанційного навчання в організаціях освітньої сфери України. Тим самим закладається фундамент забезпечення ефективного розпізнавання емоційного стану та особи слухача системи дистанційного навчання за рахунок нейромережевого аналізу біометричних параметрів, які реєструються за допомогою загальнопоширеного апаратно-програмного забезпечення, що повністю підтвердило гіпотезу роботи.

Запропоновані методологія, моделі, методи та засоби впроваджені в освітніх закладах України та можуть бути використані в практичній діяльності в контурах моніторингу психоемоційного стану операторів інформаційних систем різноманітного призначення.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Абашин В. Г. Автоматизация процесса определения психофизиологического состояния оператора автоматизированного рабочего места в АСУТП : автореф. дис. … канд. техн. наук : 05.13.06. Орел, 2007. 20 с.

2. Арсентьев Д. А. Разработка алгоритмов интеллектуализации прикладных задач при построении автоматизированных систем анализа и обработки результатов обучения : дис. … канд.. техн. наук : 05.13.06. Москва, 2013. 145 с.

3. Биков В. Ю., Буров О. Ю., Дементієвська Н. П. Кібербезпека в цифровому навчальному середовищі. *Інформаційні технології і засоби навчання*. 2019. Т. 70, № 2. С. 313-331.

4. Биков В. Ю., Буров О. Ю., Гуржій А. М., Жалдак М. І., Лещенко М. П., Литвинова С. Г., Луговий В. І., Олійник В. В., Спірін О. М., Шишкіна М. П. Теоретико-методологічні засади інформатизації освіти та практична реалізація інформаційно-комунікаційних технологій в освітній сфері України : монографія. Київ : Компринт, 2019. 214 с.

5. Биков В. Ю., Литвинова С. Г., Тукало С. М. Концептуальні підходи до проєктування цифрового портфоліо наукових і науково-педагогічних працівників. *Сучасні інформаційні технології та інноваційні методики навчання у підготовці фахівців: методологія, теорія, досвід, проблеми*. 2021. № 60. С.9-16.

6. Биков В. Ю., Кучанський О. Ю., Білощицький А. О., Андрашко Ю. В., Діхтяренко О. В., Буднік С. В. Розробка інформаційної технології комплексного оцінювання діяльності закладів вищої освіти. *Інформаційні технології і засоби навчання. Електронне наукове фахове видання*. 2019. Т. 73, № 5. С. 293-306. DOI: 10.33407/itlt.v73i5.

7. Биков М. М., Ковтун В. В. Оцінювання надійності автоматизованих систем розпізнавання мовців критичного застосування. *Вісник Вінницького політехнічного інституту.* 2017. № 2. С. 70-76.

8. Бобе А. С., Конышев Д. В., Воротников С. А. Система распознавания базовых эмоций на основе анализа двигательных единиц лица. *Инженерный журнал: наука и инновации*. 2016. Вып. 9. С. 1-16.

9. Бредихин А. И., Иванович А. В. Распознавание эмоций в человеческой речи с использованием вейвлетов : материалы XII Междунар. научн.-практ. конф. Новосибирск, 2018. С.30-42.

10. Брилюк Д. В., Старовойтов В. В. Распознавание человека по изображению лица нейросетевыми методами. Минск, 2002. 54 с. (Препринт. Ин-т техн. кибернетики НАН Беларуси; № 2).

11. Буй Т. Ч., Спицын В. Г. Разложение изображений с помощью двумерного дискретного вейвлет-преобразования и быстрого преобразования Хаара. *Проблемы информатики*. 2011. № 2. С. 11–15.

12. Бушуєв Д. А. Імунні механізми управління проектами розвитку організацій в умовах кризи : автореф. дис. … докт. техн. наук : 05.13.22. Київ, 2020. 44 с.

13. Бушуев С. Д., Бушуева Н. С., Бабаев И. А., Яковенко В. Б., Гриша Е. В., Дзюба С. В., Войтенко А. С. Креативные технологии управления проектами и программами : монография. Киев : Самит-Книга, 2010. 768 с.

14. Бушуєв С. Д., Бушуєв Д. А., Бушуєва В. Б., Бойко О. О. Концепція онтологій у моделюванні системи управління міською логістичною інфраструктурою. *Управління розвитком складних систем*. 2020. № 41. С. 14- 20.

15. Бушуев С. Д., Бушуева Н. С. Методология управления проектами как универсальная модель знаний. *Управління проектами та розвиток*

виробництва : зб. наук. пр. Луганськ: СНУ ім. В. Даля, 2003. № 4(8). С. 5-12.

16. Бушуєв С. Д., Ярошенко Р. Ф. Методології управління проектами на моделях класу "рушійні сили – опори". *Управління розвитком складних систем*. 2010. Вип. 2. С. 11-14.

17. Бушуєв С.Д., Козир Б.Ю. Гібридизація методологій управління інфраструктурними проектами та програмами. *Вісник Одеського національного морського університету*. 2020. №1 (61). С. 187-207.

18. Бушуєв С. Д., Русан Н. І. Емоційний інтелект керівника проекту. *Управління розвитком складних систем*. 2016. Вип. 28. С. 47-56.

19. Вартанов, А.В. Антропоморфный метод распознавания эмоций в звучащей речи. *Национальный психологический журнал*. 2013. №2(10). С. 69-79.

20. Веренич О.В. Управління ментальним простором проектів та програм : автореф. дис. … докт. техн. наук : 05.13.22. Київ, 2019. 41 с.

21. Гонзалес Р., Вудс Р. Цифровая обработка изображений. Москва: Техносфера, 2006. 1072 с.

22. Гришенкова Н.П., Лавров Д.Н. Обзор методов идентификации человека по радужной оболочке глаза. *Математические структуры и моделирование.* 2014. Вып. 1(29). С. 43–64.

23. Грищук Т. В., Биков М. М. Моделювання процесу аналізу і класифікації голосових команд : монографія. Вінниця : УНІВЕРСУМ-Вінниця, 2008. 146 с.

24. Гудков В. Ю., Дорофеев К. А. Применение нейронной сети для распознавания частных признаков дактилоскопических изображений. *Вестник Челябинского государственного университета*. 2012. Вып.13, № 14 (268). С. 98-100.

25. Демин А. А. Адаптивная обработка каллиграфической информации, представленной в виде рукописных символов: дис. … канд..

техн. наук : 05.13.01. Москва, 2014. 182 с.

26. Дичка І., Терейковський І., Терейковська Л., Радченко К. Метод визначення ефективного типу базисного вейвлету для застосування в шаблонах нормальної поведінки веб-сервера. *Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні*. 2018. Вип. 2(36). С. 46-55.

27. Добеши И. Десять лекций по вейвлетам. Ижевск : НИЦ «Регулярная и хаотическая динамика». 2001. 464 с.

28. Дорошенко А. Ю., Шаронова Н. В., Єна Б. О., Янголенко О. В. Розробка програмних компонентів інформаційної системи екстракції фактографічних даних з веб-ресурсів. *Проблеми інформаційних технологій*. 2018. № 1(23). С. 27–35.

29. Друки А. А. Система поиска, выделения и распознавания лиц на изображениях. *Известия Томского политехнического университета*. 2011. Т. 318, № 5. С.64-70.

30. Егоров Е. А. Неотложная офтальмология: учебное пособие для вузов. Москва : Геотар Медиа, 2005. 118 с.

31. Заболеева-Зотова А. В., Орлова Ю. А., Розалиев В. Л., Бобков А. С. Задача создания системы автоматизированного распознавания эмоций. *OSTIS-2012* : материалы Междунар. науч.-техн. конф. Минск, 2012. С. 347-350.

32. Ибрагимов В. В., Арсентьев Д. А. Алгоритмы и методы распознавания личности в условиях современных информационных технологий. *Вестник МГУП*. 2015. №1. С. 67-69.

33. Иванов А. И. Нейросетевые алгоритмы биометрической идентификации личности : монография. Москва : Радиотехника, 2004. 144 с.

34. Кафтанников И. Л., Парасич А. В. Проблемы формирования обучающей выборки в задачах машинного обучения. *Вестник ЮУрГУ. Серия*

«Компьютерные технологии, управление, радиоэлектроника». 2016. Т. 16, №3. С. 15-24.

35. Кемайкин В. К., Арнольдова Е. А. Алгоритм прямого вейвлет-преобразования изображения. *Программные продукты, системы и алгоритмы*. 2016. №2. С. 1-4.

36. Козуля Т. В., Шаронова Н. В., Козуля М. М. Формування знання-орієнтованого інформаційного забезпечення досліджень складних систем. *Системні дослідження та інформаційні технології*. 2017. № 3. С. 63-72.

37. Корепанов А. О. Обнаружение границ радужной оболочки с использованием преобразования Хоу. *Вестник Самарского государственного аэрокосмического университета*. 2008. № 2. С. 235-239.

38. Корж Р., Пелещишин А., Мастикаш О. Побудова програмного комплексу управління ІД ВНЗ у соціальних середовищах Інтернету. *Вісник Національного університету «Львівська політехніка», Серія: «Інформатизація вищого навчального закладу»*. 2017. № 879. С. 64-76.

39. Корж Р., Пелещишин А., Мастикаш О. Інформаційна система взаємодії ВНЗ з соціальними середовищами Інтернету. *Вісник національного університету «Львівська політехніка», Серія: «Інформаційні системи та мережі»*. 2017. № 872. С. 69-83.

40. Корж Р. О. Формальний опис агресії в соціальних середовищах інтернету. *Вісник інженерної академії України*. 2018. Вип. 1. С. 70-74.

41. Корж Р. О. Формування кількісних характеристик та оцінювання інформаційного образу закладу вищої освіти. *Науковий журнал «Вчені записки ТНУ імені В.І. Вернадського. Серія Технічні науки»*. 2018. Т. 29(68), №2. С. 142-148.

42. Корченко А., Терейковский И., Карпинский Н., Тынымбаев С. Нейросетевые модели, методы и средства оценки параметров безопасности

интернет-ориентированных информационных систем : монография. Київ : Наш Формат, 2016. 273 с.

43. Костюк Г. С., Балла Г. А. Учебный материал и учебные ситуации. Психологические аспекты. Київ : Рад. шк., 1986. 143 с.

44. Кошевая Н. А., Мазниченко Н. И. Подход к повышению надежности идентификации пользователей компьютерных систем по динамике написания паролей. *Системи обробки інформації*. 2014. Вип. 6 (122). С. 140-146.

45. Кулаков Ю., Терейковська Л., Терейковський І. Спосіб застосування згорткової нейронної мережі для розпізнавання особи і емоцій користувача за клавіатурним почерком. *Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні*. 2019. Вип. 2 (38). С. 9-17.

46. Лавров О. Ю. Метод формування навчальної вибірки зображень для навчання підсистеми автоматизованого розпізнавання об’єктів повітряної розвідки. *Системи обробки інформації*. 2016. Вип. 8 (145). С. 29-32.

47. Лебедева Н. Н., Каримова Е. Д. Акустические характеристики речевого сигнала как показатель функционального состояния человека. *Успехи физиологических наук*. 2014. №1. С. 57-95.

48. Леонтьєв О. М. Діяльність, свідомість, особистість. Москва: Політвидав, 1975. 304 с.

49. Ляхов П. А. Валуева М. В. Применение сглаживающих фильтров для очистки от шума изображений в оттенках серого. *Наука. Инновации. Технологии*. 2015. № 3. С. 37-50.

50. Минакова Н. Н., Петров И. В. Параметризация структуры радужной оболочки глаза с использованием вейвлет-преобразования. *Известия Алтайского государственного университета*. 2011. №1. С.105-107.

51. Михайлов В. О. Программная реализация измерения эмоций

абонента мобильного телефона. *Современные проблемы науки и образования*. 2012. № 5. С. 45-62.

52. Михайленко В. М., Терейковская Л. А. Обзор средств распознавания эмоционального состояния человека по геометрии лица. *Управління розвитком складних систем*. 2019. № 37. С. 178-184.

53. Михайленко В. М., Терейковская Л. А., Легеза В. П. Адаптация параметров двухслойного персептрона, предназначенного для биометрической аутентификации пользователей по отпечаткам пальцев. *Управління розвитком складних систем*. 2018. № 34. С.130-138.

54. Михайленко В. М., Терейковская Л. А. Концептуальная модель нейросетевого распознавания эмоционального состояния слушателей системы дистанционного обучения. *Управління розвитком складних систем*. 2019. № 38. С. 193-199.

55. Михайленко В. М., Терейковская Л. А. Парадигмы нейросетевого распознавания эмоций. *Управління розвитком складних систем*. 2019. №39. С. 179-186.

56. Мищенко В. А. Алгоритм распознавания графических образов. *Вестник Воронежского государственного технического университета*. 2009. Т. 5., № 12. С. 103–105.

57. Мищенко В. А. Модели и алгоритмы распознавания графических образов на основе нечетких нейронных сетей : автореф. дис. … канд. техн. наук : 05.13.17. Воронеж, 2013. 16 с.

58. Міхайленко В. М., Терейковська Л. О. Оцінка ефективності нейромережевих моделей, призначених для розпізнавання емоцій диктора. *Управління розвитком технологій* : збірник тез доповідей VII Міжнародної науково-практичної конференції. Київ, 2020. С. 91-92.

59. Міхайленко В. М., Терейковська Л. О., Терейковський І. А.,

Ахметов Б. Б. Нейромережеві моделі та методи розпізнавання фонем в голосовому сигналі в системі дистанційного навчання : монографія. Київ : Компринт, 2017. 252 с.

60. Міхайленко В. М., Терейковська Л. О., Терейковський О. І. Нейромережеве розпізнавання відбитків пальців. *Сучасні методи, інформаційне, програмне та технічне забезпечення систем керування організаційно-технічними та технологічними комплексами* : матеріали V Міжнар. наук.-техн. Internet-конференції. 22-23 листопада 2018 р. Київ. С. 179-180.

61. Одиноких Г. А. Методы и алгоритмы биометрического распознавания человека по радужной оболочке глаза на мобильном устройстве : дис. … канд.. техн. наук : 05.13.17. Москва, 2019. 109 с.

62. Переберин А. В. О классификации вейвлет-преобразований. *Вычислительные методы и программирование*. 2001. Т. 2. Вып.3, С. 15–40.

63. Попова А. С., Рассадин А. Г., Пономаренко А. А. Детектирование эмоций в мультимедиа контенте. *Информационные системы и технологии* : материалы Междунар. науч.-техн. конф. Москва, 2017. С. 852-857.

64. Прохоров В. Г. Использование сверточных сетей для распознавания рукописных символов. *Прикладне програмне забезпечення*. 2008. №2-3. С. 669-674.

65. Путятин Е. П., Кобылин О. А. Использование вейвлет-преобразования для описания контура изображения. *Системи обробки інформації*. 2004. Вип. 1. С. 51-54.

66. Ребрун И. А., Сидоров К. В., Филатова Н. Н., Ханеев Д. М. Модель проявления эмоций в естественной речи. *КИИ-2014* : материалы XIV национальной конференции по искусственному интеллекту. Казань, 2014. С. 112-121.

67. Романько О. В. Особенности применения пакета WAVELET TOOLBOX для спектрального анализа сигналов. *Системи обробки інформації*. 2005. Вип. 2. С. 119-126.

68. Рубан И. В., Хижняк И. А. Анализ подходов к идентификации лиц в системах контроля доступа. *Наука і техніка Повітряних Сил Збройних Сил України.* 2014. № 4. С. 67-70.

69. Рубинштейн С. Основы общей психологии. СПб : Издательство Питер, 2016. 720 с.

70. Руденко О. Г., Бодянський Є. В. Штучні нейронні мережі : навч. посіб. Харків : ТОВ "Компанія СМІТ", 2006. 404 с.

71. Русан Н. І., Бушуєв С. Д. Застосування емпатії та емоційного інтелекту для різних видів команд в умовах пандемії. *Вісник Національного технічного університету «ХПІ». Серія: Стратегічне управління, управління портфелями, програмами та проектами*. 2021. № 1(3). С. 3-8.

72. Рыканов А. С. Анализ методов распознания отпечатков пальца. *Системи обробки інформації*. 2010. Вип. 6 (87). С.164-171.

73. Савинов А. Н. Методы, модели и алгоритмы распознавания клавиатурного почерка в ключевых системах : автореф. дис. … канд. техн. наук: 05.13.19. Санкт-Петербург, 2013. 19 с.

74. Савинов А. Н., Сидоркина И. Г. Математическая модель механизма распознавания клавиатурного почерка на основе Гауссовского распределения. *Известия Кабардино-Балкарского научного центра РАН*. Нальчик, 2013. Вып. I. С. 26-32.

75. Сидоров К. В. Биотехническая система мониторинга эмоций человека по речевым сигналам и электроэнцефалограммам : дис. …канд. техн. наук : 05.11.17. Тверь, 2015. 182 с.

76. Силантьев И. О. Нейросетевые модели биометрической

идентификация и аутентификация пользователей по отпечаткам пальцев : материалы XXXVI Междунар. науч.-техн. конф. Новосибирск : изд. «СибАК», 2014. № 7 (32). С 43-47.

77. Скринникова А. В. Изменение индивидуальной динамики манипуляций устройствами управления курсором под влиянием эмоций страха и радости. *Известия ЮФУ. Технические науки*. 2013. №5. С. 246-251.

78. Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в MATLAB. Москва : ДМК Пресс, 2014. 628 с.

79. Смульсон М. Л., Машбиць Ю. І., Жалдак М. І. та ін. Дистанційне навчання: психологічні засади : монографія. Кіровоград : Імекс-ЛТД, 2012. 240 с.

80. Солдатова О. П., Гаршин А. А. Применение сверточной нейронной сети для распознавания рукописных цифр. *Компьютерная оптика*. 2010. №2. С.252-258.

81. Сыропятов А. А. Метод мониторинга трафика защищенных высокоскоростных коммерческих сетей нового поколения. *Наукові записки УНІДЗ*. 2009. № 2(1) С. 65-73.

82. Тарасенко В., Корченко О., Терейковський І. Метод застосування продукційних правил для подання експертних знань в нейромережевих засобах розпізнавання мережевих атак на комп'ютерні системи. *Безпека інформації*. 2013. Т. 19, № 3. С. 168-174.

83. Терейковська Л. О. Аналіз засобів розпізнавання емоційного стану людини за голосом. *Вчені записки Таврійського національного університету імені В. І. Вернадського, серія «Технічні науки»*. 2019. Том 30 (69). Ч. 1, №1. С.120-124.

84. Терейковська Л. О. Застосування нейронних мереж в програмних системах, що призначені для розпізнавання емоцій диктора. *Розподілені*

програмні системи і технології : збірник тез доповідей І науково-практичної конференції. Київ, 2020. С. 64.

85. Терейковська Л. О. Нейромережева модель розпізнавання емоційного стану операторів автоматизованих робочих місць за клавіатурним почерком. *Вчені записки Таврійського національного університету імені В. І. Вернадського, серія «Технічні науки»*. 2019. Том 30 (69). Ч. 1, №4. С. 129-133.

86. Терейковська Л. О. Нейромережева модель розпізнавання емоцій по голосу. *Сучасні методи, інформаційне, програмне та технічне забезпечення систем керування організаційно-технічними та технологічними комплексами:* матеріали VII Міжнар. наук.-техн. Internet-конференції. Київ, 26 листопада 2020. С.124-125.

87. Терейковська Л. Метод нейромережевого аналізу клавіатурного почерку. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2019. Вип. 37. С. 53-59.

88. Терейковська Л. О. Метод визначення виду згорткової нейронної мережі для аналізу параметрів клавіатурного почерку. *Вчені записки Таврійського національного університету імені В. І. Вернадського, серія «Технічні науки».* 2019. Том 30 (69). Ч. 1, № 6. С. 164-168.

89. Терейковська Л. О. Метод нейромережевого аналізу голосового сигналу. *Кібербезпека: освіта, наука, техніка*. 2020. № 3 (7). С. 31-42.

90. Терейковська Л. Метод фільтрації біометричних параметрів на основі вейвлет-перетворень. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2021. №42. С. 95-103.

91. Терейковська Л. О. Модель формування навчальних прикладів нейронної мережі, що призначена для аналізу клавіатурного почерку. *Кібербезпека: освіта, наука, техніка*. 2020. № 1 (9). С. 104-114.

92. Терейковська Л. Метод нейромережевого розпізнавання емоцій по

зображенню обличчя. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2020. Вип. 40. С. 146-152.

93. Терейковська Л. О. Проблема створення ефективних засобів розпізнавання голосових сигналів, що адаптовані до умов вітчизняних Інтернет-орієнтованих інформаційних систем. *Актуальні питання забезпечення кібербезпеки та захисту інформації* : збірник тез доповідей III Міжнар. наук.-практ. конф. Верхнє Студене, 22-25 лютого 2017 р. С. 163-165.

94. Терейковська Л. О., Іванченко Є. В., Погорелов В. В. Метод адаптації глибокої нейронної мережі до розпізнавання комп'ютерних вірусів. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2019. Вип. 35. С. 198-205.

95. Терейковська Л. О., Терейковський О. І. Нейромережева модель розпізнавання емоцій по зображенню обличчя. *Вчені записки Таврійського національного університету імені В. І. Вернадського, серія «Технічні науки».* 2019. Том 30 (69). Ч. 1, №2. С. 209-213.

96. Терейковская Л. А., Терейковский И. А. Применение сверточной нейронной сети для анализа биометрических параметров. *Вчені записки Таврійського національного університету імені В.І. Вернадського, серія «Технічні науки»*. 2020. Том 31 (70), № 5. С. 124-128.

97. Терейковская Л. А. Архитектура нейросетевого анализатора биометрических параметров слушателей системы дистанционного обучения. *Управління розвитком складних систем*. 2020. № 44. С. 91-99.

98. Терейковская Л. А. Нейросетевая модель распознавания эмоций по голосу. *Управління розвитком складних систем*. 2020. №41. С. 95-100.

99. Терейковський І. А., Заріцький О. В., Терейковська Л. О., Погорелов В. В. Метод розробки архітектури глибокої нейронної мережі, призначеної для розпізнавання комп'ютерних вірусів. *Захист інформації*.

2018. Том 20, № 3. С. 188-199.

100. Терейковский И., Корченко А., Ахметов Б., Бапиев И., Терейковська Л. Определение эффективных видов нейросетевых моделей распознавания кибератак на сетевые ресурсы. *Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні*. 2016. Вип. 2(32). С.56-63.

101. Терейковський І. А., Кулаков Ю. О., Терейковська Л. О., Терейковський О. І. Параметри оцінки ефективності засобів розпізнавання особи за райдужною оболонкою та сітківкою ока. *Вчені записки Таврійського національного університету імені В.І.Вернадського, серія «Технічні науки»*. 2022. Том 33 (72), № 1. С. 195-199.

102. Терейковський І. А., Терейковська Л. О. Застосування експертних знань для навчання нейронних мереж. *Сучасні методи, інформаційне, програмне та технічне забезпечення систем управління організаційно-технічними та технологічними комплексами* : матеріали III Міжнар. наук.-техн. Internet-конф. Київ, 23 листопада 2016. С. 147.

103. Терейковський І. А., Терейковська Л. О. Метод кодування вихідного сигналу нейромережевих моделей. *Наукові вісті НТУУ «КПІ»: міжнародний науково-технічний журнал*. 2017. № 5(115). С. 59-64.

104. Терейковський І. А., Терейковська Л. О., Ахметов Б. Б. Адаптація багатошарового персептрону до умов застосування в системах голосової аутентифікації. *ABIA-2017* : матеріали XIII міжнар. наук.-техн. конф. Київ, 2017. С.1.1-1.4.

105. Терейковський І. А., Терейковська Л. О., Терейковський О. І. Визначення ознак деструктивного впливу в невербальному контенті звукових рядів Інтернет-орієнтованих засобів масової інформації. *Кібербезпека в Україні: правові та організаційні питання : матеріали міжн. наук. практ.*

конф. м. Одеса. 26 листопада 2020 р. С.76-79.

106. Терейковський І., Терейковська Л., Терейковський О. Нейромережеві технології біометричної ідентифікації. *Безпека соціально-економічних процесів в кіберпросторі* : матеріали Всеукр. наук.-практ. конф. м. Київ. 27 березня 2019 р. С. 101-103.

107. Терейковский И. А., Терейковская Л. А. Многослойный персептрон в задачах оценки параметров технических систем. *Інтелектуальний аналіз інформації* : збірка праць XVII Міжнар. наук.конф. імені Т. А. Таран. м. Київ. 17-19 травня 2017 р. С.233-238.

108. Терейковский И. А., Терейковская Л. А., Картбаев Т. С., Бекетова Г. С. Способ нейросетевого распознавания личности пользователя по радужной оболочке глаза. *Энергетика, инфокоммуникационные технологии и высшее образование :* материалы XI Междунар. науч.-техн. конф. Алматы. 16-18 октября 2020. С.297-300.

109. Токмолдин С. Ж., Клименов В. В., Тулеуов З. Ж., Шакенов К. К. Обработка сигналов с помощью быстрого вейвлет-преобразования на суперкомпьютере в системе MATLAB. *Вестник КазНУ. Серия математика, механика, информатика*. 2012. №3(74). С. 69-80.

110. Третьяков И. Н., Минакова Н. Н. Алгоритм разграничения доступа по радужной оболочке глаза для решения задач контроля доступа к информационным ресурсам. *Доклады Томского государственного университета систем управления и радиоэлектроники*. 2010. №1, Т. 1. С. 100-102.

111. Умяров Н. Х., Костецкая Г. Ю., Федяев О. И. Программная модель сверточной нейронной сети. *Інформаційні управляючі системи та комп'ютерний моніторинг* : збірка праць III Всеукр. наук.-практ. конф. студентів, аспірантів та молодих вчених. Донецьк. 2011. С. 343-347.

112. Фам Д. Т. Некоторые вопросы применения методов сравнения отпечатков пальцев для биометрических систем идентификации личности *Международный научно-исследовательский журнал*. 2015. № 7(38) С. 127- 130.

113. Фан Н. Х., Спицын В. Г. Алгоритмы для классификации отпечатков пальцев на основе применения фильтра Габора, вейвлет-преобразования и многослойной нейронной сети. *Известия Томского политехнического университета*. 2012. Т. 320, № 5. С.60-64.

114. Фисенко В. Т., Фисенко Т. Ю. Компьютерная обработка и распознавание изображений. СПб : СПбГУ ИТМО, 2008. 192 с.

115. Хуршудов А. А. Разработка системы распознавания визуальных образов в потоке данных : дис. … канд.. техн. наук: 05.13.01. Краснодар, 2015. 130 с.

116. Цюцюра М. І. Інформаційні технології гармонізації зрівноваженого освітнього простору : автореф. дис. … докт. техн. наук : 05.13.06. Київ, 2020. 48 с.

117. Чернышев Д.О., Михайленко В.М., Терейковская Л.А. Разработка и исследование инструментальных средств для нейросетевого анализа голоса слушателей системы дистанционного обучения. *Управління розвитком складних систем*. 2020. № 43. С. 123-130.

118. Шокуров А.В., Михалев А.В. Оптимальное использование вейвлет-компонент. *Успехи математических наук*. 2007. Т. 62, № 4. С. 171-172.

119. Ягупов В. В., Петренко Л. М., Кравець С. Г. та ін. Дистанційне навчання в системі професійно-технічної освіти : монографія. Житомир: «Полісся», 2019. 234с.

120. Abate A. Two-Tier Image Features Clustering for Iris Recognition on Mobile. *Lecture Notes in Computer Science*. 2017. Vol. 10147. P. 260-269.

121. Abdel-Mottaleb M., Zhou J. Human Ear Recognition from Face Profile Images. *In: Zhang, D., Jain, A.K. (eds) Advances in Biometrics. ICB 2006*. Lecture Notes in Computer Science. Vol 3832. Springer, Berlin, Heidelberg. DOI:10.1007/11608288_105.

122. Abisado M. B., Gerardo B. D., Fajardo A. C. Towards keystroke analysis using neural network for multi-factor authentication of learner recognition in on-line examination. *Manila International Conference on Trends in Engineering and Technology*. 2017. P. 71-74.

123. Aim A., Bill F. Ocular circulation. *Adler's Physihology of the Eye*. Ed by Moses and Hart, 1987. P. 183-203.

124. Aitkhozhayeva Y., Tynymbayev S., Seilova N., Tereikovska L., Imanbayev A. Method and device for modulus reduction. *Bulletin of the National Academy of Sciences of the Republic of Kazakhstan*. 2019. Vol. 2, No 378. P. 220-225.

125. Ajinkya N. Jadhav, Nagaraj V. Dharwadkar. A Speaker Recognition System Using Gaussian Mixture Model, EM Algorithm and K-Means Clustering. *International Journal of Modern Education and Computer Science*. 2018. Vol.10, No.11. P. 19-28.

126. Akhmetov B., Tereykovsky I., Doszhanova A., Tereykovskaya L. Determination of input parameters of the neural network model, intended for phoneme recognition of a voice signal in the systems of distance learning. *International Journal of Electronics and Telecommunications*. 2018. Vol. 64, No 4. P. 425-432.

127. Akhmetov B., Tereikovskyi I., Tereikovska L., Adranova A. Neural Network User Authentication by Geometry of the Auricle. *Advances in Intelligent Systems and Computing*. 2018. Springer, Cham. Vol. 836. P. 11-19.

128. Akkermans A.H., Kevenaar T. A., Schobben D. W. Acoustic ear

recognition for person identification. *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*. 2005. P. 219-223.

129. Alam M., Kenny P., Gupta V. Tandem Features for Text-dependent Speaker Verification on the RedDots Corpus. INTERSPEECH, 2016. P. 420-424.

130. Alghamdi S., Elrefaei L. Dynamic user verification using touch keystroke based on medians vector proximity. *IEEE 7th International Conference on Computational Intelligence, Communication Systems and Networks*. 2015. P.121-126.

131. Ali M., Thakur K., Tappert C. User authentication and identification using neural network. *I-manager's Journal on Pattern Recognition*. 2015. No 2. P. 28-39.

132. Al-Ammouri A., Al-Ammori H., Klochan A., Degtiarova A. Probabilistic models reliability of information and control systems. Journal of Sustainable Development of Transport and Logistics. 2018. No 3(1). P. 60-69.

133. Al-Ammori A., Dmytrychenko A., Al-Ammori H. Probabilistic-mathematical models for formation of information flows in aircraft fire alarm system. Journal of Automation and Information Sciences. 2019. No 51(7). P. 67-80.

134. Al-Ammouri A., Dmytrychenko A., Al-Ammori H., Kharuta V. Development of structures of the aircraft fire alarm system by means of nested modules. Eastern-European Journal of Enterprise Technologies. 2019. No 2(9-98). P. 14-23.

135. Al-Raisi A. N., Al-Khouri A. M. Iris recognition and challenge of homeland and border control security in UEA. *Telematics and Informatics*. 2008. Vol. 25. P. 117-132.

136. Altincay H. Speaker identification by combining multiple classifiers using Dempster-Shafer theory of evidence. *Speech Communication*. 2003. Vol. 41, No 4. P. 531-547.

137. Anderson K., McOwan P. W. A real-time automated system for the

recognition of human facial expressions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*. 2006. Vol.36, No 1. P. 96-105.

138. Arsentyev D. A., Biryukova T. S. Method of flexible comparison on graphs as algorithm of images recognition. *Bulletin of Ivan Fedorov MGUP*. 2015. No. 6. P. 74-75.

139. Bapiyev I. M., Aitchanov B. H., Tereikovskyi I. A., Tereikovska L. A., Korchenko A. A. Deep neural networks in cyber attack detection systems. *International Journal of Civil Engineering and Technology*. November 2017. Vol, 8. Is. 11. P. 1086-1092.

140. Barmak O., Krak I., Mazurets O., Pavlov S., Smolarz A. Research of efficiency of information technology for creation of semantic structure of educational materials. *Advances in Intelligent Systems and Computing*. 2020. Vol. 1020. P. 554-569.

141. Batista J. C., Albiero V., Bellon O. R., Silva L. AUMPNet: Simultaneous Action Units Detection and Intensity Estimation on Multipose Facial Images Using a Single Convolutional Neural Network. *12th IEEE International Conference on Automatic Face & Gesture Recognition*. 2017. P. 866-871.

142. Baxter G. M., Williamson T. H. Color Doppler imaging of the eye: normal ranges, reproducibility and observer variation. *J. Ultrasound Med*. 1995. No 14(2). P. 91-96. DOI:10.7863/jum.1995.14.2.91.

143. Bello J., Pickens J. A Robust Mid-Level Representation for Harmonic Content in Music Signals. *6th International Conference on Music Information Retrieval*. London, UK. September 11-15, 2005. P. 304-311.

144. Bo C., Zhang L., Jung T., Han J., Li X.-Y., Wang Y. Continuous user identification via touch and movement behavioral biometrics. *IEEE 33rd International Performance Computing and Communications Conference*. 2014. P. 1-8.

145. Boulanger-Lewandowski N., Bengio Y., Vincent P. Audio chord

recognition with recurrent neural networks. *14th International Society for Music Information Retrieval Conference*. November 4-8, 2013. P. 21-29.

146. Bojorges-Valdes E. R. Scaling patterns of heart rate variability data. *Physiol Meas.* 2007. No 28(6). P. 721-730. DOI:10.1088/0967-3334/28/6/010.

147. Bowyer K.W. Identity in the Information Society. *Journal on Advances in Signal Processing*. 2009. Vol 2, No 3. P. 327-343.

148. Bryliuk D., Starovoitov V. Application of Recirculation Neural Network and Principal Component Analysis for Face Recognition. *The 2nd International Conference on Neural Networks and Artificial Intelligence*. Minsk, Belarus. October 2-5, 2001. P. 136-142.

149. Burge M., Burger W. Ear biometrics. *Biometrics. Springer*. Boston, 1996. P. 273-285. DOI:10.1007/0-306-47044-6_13.

150. Campbell W., Sturim D., Reynolds D. (2006). Support vector machines using GMM supervectors for speaker verification. *IEEE Signal Processing Letters.* Vol. 13, Is. 5. May 2006. P. 308-311. DOI:10.1109/LSP.2006.870086.

151. Carew T. J, Magsamen S. H. Neuroscience and education: an ideal partnership for producing evidence-based solutions to Guide 21(st) Century Learning. *Neuron*. 2010. Sep 9, 67(5):685-8. DOI:10.1016/j.neuron.2010.08.028.

152. Chandrani S., Washef A., Soma M., Debasis M. Facial Expressions: A Cross-Cultural Study. *Emotion Recognition: A Pattern Analysis Approach.* Wiley Publ., 2015. P. 69-87. DOI:10.1002/9781118910566.

153. Chang K., Bowyer K.W., Sarkar S., Victor B. Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2003. Vol. 25, No. 9. P. 1160-1165.

154. Chen Q., Zhang X., Xiong S., Chen X. Short-term power load forecasting with least squares support vector machines and wavelet transform. *International*

Conference on Machine Learning and Cybernetics. 2008. Vol. 3. P. 1425-1429.

155. Chirchi V., Waghmare L. Iris Biometric Authentication used for Security Systems. *International Journal of Image, Graphics and Signal Processing*, 2014. Vol. 6, No. 9. P. 54-60.

156. Cho K. et al. On the properties of neural machine translation: Encoder-decoder approaches. 2014. (preprint arXiv:1409.1259v2).

157. Chung J. et al. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014. (preprint arXiv:1412.3555).

158. Connaughton R., Bowyer K. W., Flynn P. J. Fusion of Face and Iris Biometrics. *Handbook of Iris Recognition*. Springer, 2013. P. 219-237.

159. Danchuk V., Lemeshko Y., Lemeshko T. Concept of system-synergistic approach in project management. *Bulletin of the National Transport University*, 2012. Vol 26(2). P. 128-133.

160. Dawson M., Schell A. Handbook of Psychophysiology. Cambridge University Press. New York, 2000.

161. Dedov I. et al. Prevalence of diabetic retinopathy and cataract in adult patients with type 1 and type 2 diabetes in Russia. *Rev. Diabet Stud*, 2009. Vol. 6, No 2. P. 124-129.

162. Deng Y., Zhong Y. Keystroke Dynamics Advances for Mobile Devices Using Deep Neural Network. *Recent Advances in User Authentication Using Keystroke Dynamics Biometrics*. Publisher: Science Gate PublishingEditors. 2015. Vol. 2, Ch. 4. P. 59-70. DOI: 10.15579/gcsr.vol2.ch4.

163. Dey S., Madikeri S., Ferras M. Deep neural network based posteriors for text-dependent speaker verification. *IEEE International Conference on Acoustics, Speech and Signal Processing*. 2016. P. 5050-5054.

164. Drugman T., Dutoit T. On the Potential of Glottal Signatures for Speaker Recognition. *Interspeech*. 2010. P. 2106-2109.

165. Dua M., Gupta R., Khari M. et al. Biometric iris recognition using radial basis function neural network. *Soft Comput.* 2019. No 23. P. 11801-11815.

166. Ekman P. Emotions Revealed, Second Edition: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Holt Paperbacks. 2007. 320 p.

167. Ekman P., Friesen W.V. Nonverbal leakage and clues to deception. *Psychiatry MMC*, 1969. Vol. 32, Is. 1. P. 88-106.

168. El-Abed M, Charrier C., Rosenberger C. Evaluation of Biometric Systems. *New Trends and Developments in Biometrics*. 2012. Vol. 7. P. 159-169. DOI: 10.5772/52084.

169. Epp C., Lippold M., Mandryk R. Identifying Emotional States Using Keystroke Dynamics. *Conference on Human Factors in Computing Systems*. May 2011. Vancouver, Canada: ACM. P. 715-724.

170. Ertugrul O., Jeni L., Cohn J. FACSCaps: Pose-Independent Facial Action Coding with Capsules. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Salt Lake City, 2018. P. 2211-221109.

171. Fedotov D. V., Popov V. A. Optimisation of convolutional neural network structure with self-configuring evolutionary algorithm in one identification problem. *Vestnik SibGAU*. 2013. Vol. 16, No. 4. P. 857-862.

172. Ganchev T., Fakotakis N., Kokkinakis G. Comparative evaluation of various MFCC implementations on the speaker verification task. *10th International Conference on Speech and Computer*. Patras, Greece. 2005.

173. Geeta N., Soni M. A New Design Approach for Speaker Recognition Using MFCC and VAD. *International Journal of Image, Graphics and Signal Processing*. 2013. Vol.5, No.9. P.43-49.

174. Ghosh S., Laksana E., Scherer S., Morency L. A multi-label convolutional neural network approach to cross-domain action unit detection. *IEEE*

International Conference on Affective Computing and Intelligent Interaction. 21-24 September 2015. P. 609-615.

175. Ghosh S., Goenka S., Ganguly N., Mitra B., De P. Representation Learning for Emotion Recognition from Smartphone Keyboard Interactions. *8th International Conference on Affective Computing and Intelligent Interaction*. 3-6 September 2019. P. 704-710. DOI: 10.1109/ACII.2019.8925518.

176. Goodfellow J., Erhan D., Carrier L. Challenges in representation learning: a report on three machine learning contests. *Neural Networks*. 2015. Vol. 64. P. 59-63.

177. Gref K., Srivastava R., Koutník J., Steunebrink B., Schmidhuber J. LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*. October 2017. Vol. 28, No 10. P. 2222-2232.

178. Gudi A., Tasli H., Den U., Maroulis A. Deep learning based FACS Action Unit occurrence and intensity estimation. *IEEE 11th International Conference and Workshops on Automatic Face and Gesture Recognition*. Vol. 6. P. 1-5.

179. Hajaria K., Ujwalla Gawandeb, Golharc Y. Neural Network Approach to Iris Recognition in Noisy Environment. *International Conference on Information Security & Privacy.* 11-12 December 2015. Nagpur, India. Procedia Computer Science, 78. P. 675-682.

180. Hamsa A. Fingerprint Identification System Using Neural Networks. *Nahrain University, College of Engineering Journal*. 2012. Vol.15, No.2. P. 234-244.

181. Han J. et al. Sentiment Pen: Recognizing Emotional Context Based on Handwriting Features. *10th Augmented Human International Conferences*. 11-12 March 2019. Reims, France. P. 1-9. DOI: 10.1145/3311823.3311868.

182. Hayreddin C., Shambhu U. Sensitivity analysis in keystroke dynamics using convolutional neural networks. *IEEE Workshop on Information Forensics and*

Security. 4-7 December 2017. P. 1-6.

183. He J., Li D., Yang B., Cao S., Sun B., Yu L. Multi view facial action unit detection based on CNN and BLSTM RNN. *12th IEEE International Conference on Automatic Face & Gesture Recognition*. 2017. P. 848-853.

184. He L. Stress and Emotion Recognition in Natural Speech in the Work and Family Environments: Ph.D. Dissertation. RMIT University, 2010. 197 p.

185. Hinton G., Sabour S., Frosst N. Matrix capsules with EM routing. *ICLR Conference*. Vancouver, BC, Canada. 2018.

186. Hu Z., Tereikovskyi I., Chernyshev D., Tereikovska L., Tereikovskyi O., Wang D. Procedure for Processing Biometric Parameters Based on Wavelet Transformations. *International Journal of Modern Education and Computer Science*. 2021. Vol. 13, No 2. P. 11-22.

187. Hu Z., Tereikovskyi I., Korystin O., Mihaylenko V., Tereikovska L. Two-Layer Perceptron for Voice Recognition of Speaker's Identity. *Advances in Intelligent Systems and Computing*. 2021. Springer, Cham. Vol. 1247. P. 508-517.

188. Hu Z., Tereikovskyi I., Tereikovska L., Tsiutsiura M., Radchenko K. Applying Wavelet Transforms for Web Server Load Forecasting. *Advances in Intelligent Systems and Computing*. 2020. Springer, Cham. Vol 938. P. 13-22.

189. Hu Z., Tereykovskiy I., Zorin Y., Tereykovska L., Zhibek A. Optimization of convolutional neural network structure for biometric authentication by face geometry. *Advances in Intelligent Systems and Computing*. 2019. Vol. 754. P. 567-577.

190. Hwang, K., Sung W. Single stream parallelization of generalized LSTM-like RNNs on a GPU. 2015 (preprint arXiv:1503.02852).

191. Iandola F.N., Han S. W., Moskewicz M.W. SqueezeNet: AlexNet level accuracy with 50x fewer parameters and <0.5MB model size. 2016. 13 p. (preprint arXiv:1602.07360v4).

192. Ilbeygi M., Shah-Hosseini H. A novel fuzzy facial expression recognition system based on facial feature extraction from color face images. *Engineering Applications of Artificial Intelligence*. 2012. Vol. 25. P. 130-146.

193. Ingale A. B., Chaudhari D. S. Speech emotion recognition. *International Journal of Soft Computing and Engineering*. 2012. Vol. 2, No 1. P. 235-238.

194. Ding I., Yen C., Hsu Y. Developments of Machine Learning Schemes for Dynamic Time-Wrapping-Based Speech Recognition. *Mathematical Problems in Engineering*. 2013. Vol. 2013. P. 56-68.

195. Irum A., Salman A. Speaker Verification Using Deep Neural Networks: A Review. *International Journal of Machine Learning and Computing*. 2019. Vol. 9, No. 1. P. 20-25.

196. ISO/IEC 19794-4:2005, Information technology. Biometric data interchange formats. Part 4: Finger image data (IDT).

197. ISO/IEC 19794-6:2005, Information technology. Biometric data interchange formats. Part 6: Iris image data.

198. ISO/IEC 19794-5:2011/AMD 2:2015, Information technology. Biometric data interchange formats. Part 5: Face image.

199. ISO 3382-2:2008, Acoustics - Measurement of room acoustic parameters. Part 2: Reverberation time in ordinary rooms.

200. Jung N, Wranke C, Hamburger K, Knauff M. How emotions affect logical reasoning: evidence from experiments with mood-manipulated participants, spider phobics, and people with exam anxiety. *Frontiers in Psychology*. 2014. Vol. 5. P. 1-12. DOI:10.3389/fpsyg.2014.00570.

201. Juslin P., Laukaa P. Communication of emotions in vocal expression and music performance: different channels, same code. *Psychological bulletin.* 2003. Vol. 129, No 5. P. 770-814.

202. Jyoti M., Dhiraj G., Ratna D., Sainarayanan G. Reference Threshold

Calculation for Biometric Authentication. *Image, Graphics and Signal Processing*. 2014. Vol. 6, No. 2. P. 46-53.

203. Kanimozhi M., Puvirajasingam K., Avitha M. S. Survey on keystroke dynamics for a better biometric authentication system. *International Journal of Emerging Technologies and Engineering.* 2014. Vol. 1, No 9. P. 116-139.

204. Kanishcheva O., Sharonova N. Image and video tag aggregation. *CEUR Workshop Proceedings*. 2018. No 2268. P. 161-172.

205. Karam Z., Campbell W. A new kernel for SVM MLLR based speaker recognition. *8th Annual Conference of the International Speech Communication Association*. August 27-31, 2007. Antwerp, Belgium. P. 290-293.

206. Kayaoglu M., Topcu B., Uludag U. Standard Fingerprint DatabasesStandard. Manual Minutiae Labeling and Matcher Performance Analyses. 2013. 13 p. (preprint arXiv:1305.1443).

207. Kenny P., Gupta V., Stafylakis T., Ouellet P., Alam J. Deep neural networks for extracting baum-welch statistics for speaker recognition. *Proc. The Speaker and Language Recognition Workshop*. Odyssey, 2014. P. 293-298. DOI:10.21437/Odyssey.2014-44.

208. Kevin C. Using keystroke analytics to improve pass-fail classifiers. *Journal of Learning Analytics*. 2017. Vol. 4, No. 2. P. 189-211.

209. Khadkevich M., Omologo M. Time-frequency reassigned features for automatic chord recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*. May 22-27, 2011. Prague, Czech Republic. P. 181-184.

210. Kobojek P., Saeed K. Application of Recurrent Neural Networks for User Verification based on Keystroke Dynamics. *Journal of Telecommunications and Information Technology*. 2016. No 3. P. 80-90.

211. Kołakowska A., Landowska A. Keystroke Dynamics Patterns While Writing Positive and Negative Opinions. *Sensors*. 2021. Vol. 21(5963). P. 1-21.

DOI: 10.3390/s21175963.

212. Konar A., Chakraborty A. Emotion recognition: a pattern analysis approach. Wiley, 2015. 584 p.

213. Krak I., Barmak O., Mazurets O. The practice implementation of the information technology for automated definition of semantic terms sets in the content of educational materials. *CEUR Workshop Proceedings*. 2018. Vol. 2139. P. 245-254.

214. Kumar V.K., Srinivasan B. Ear Biometrics in Human Identification System. *International Journal of Information Technology and Computer Science.* 2012. Vol.4, No.2. P.41-47. DOI: 10.5815/ijitcs.2012.02.06.

215. Kun L., Jiafen L., Ping T., Qing L. Keystroke Biometrics for Freely Typed Text Based on CNN model. *Twenty-Second Pacific Asia Conference on Information Systems*. Japan, 2018. P. 2278-2324.

216. LeCun Y., Bottou L, Bengio Y., Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998. Vol. 86, No11. P. 2278-2324.

217. LeCun Y., Huang F.J., Bottou L. Learning methods for generic object recognition with invariance to pose and lighting. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2004. Vol. 2. P. 97-104.

218. Learned-Miller E., Huang G., RoyChowdhury A., Li H., Gang H. Labeled faces in the wild: a survey. *In Advances in Face Detection and Facial Image Analysis. Springer*, 2016. P. 189-248.

219. Lee P., Tsui W., Hsiao T. The Influence of Emotion on Keyboard Typing: An Experimental Study Using Auditory Stimuli. *PLoS ONE*. 2015. Vol. 10(6). P. 1-16. DOI: 10.1371/journal.pone.0129056.

220. Li L., Wang D., Rozi A., Fang T. Cross-lingual speaker verification with

deep feature learning. 2017 (preprint arXiv:1706.07861).

221. Li C., Ma X., Jiang B., Li X., Zhang X., Liu X., Cao Y., Kannan A., Zhu Z. Deep Speaker: an End-to-End Neural Speaker Embedding System. 2017 (preprint arXiv:1705.02304).

222. Lin C., Liu J., Lee K. On neural networks for biometric authentication based on keystroke dynamics. *Sensors and Materials*. 2018. Vol. 30, No 3. P. 385-396.

223. Littlewort, G., Whitehill J., Wu T., Fasel I., Frank M., Movellan J., Bartlett M. The Computer Expression Recognition Toolbox. *IEEE International Conference on Automatic Face & Gesture Recognition*. 2011. P. 298-305.

224. Liu M., Guan J. User keystroke authentication based on convolutional neural network. *Communications in Computer and Information Science*. 2019. Vol. 971. P. 157-168.

225. Lyon R.F. Machine Hearing: An emerging field. *IEEE signal processing magazine*. 2010. Vol. 27, No. 5. P. 131-139.

226. Lyubchyk L., Grinberg G. Inverse Dynamic Models in Chaotic Systems Identification and Control Problems. *International Conference on Advanced Computer Information Technologies*. June 1-3, 2018. Ceske Budejovice, Czech Republic. P. 34-37.

227. Mallouh A., Qawaqneh Z., Barkana B. New transformed features generated by deep bottleneck extractor and a GMM–UBM classifier for speaker age and gender classification. *Neural Computing and Applications*. 2017. P. 1- 13.

228. Makarova V., Petrushin V. Ruslana: a database of russian emotional utterances. *7th International Conference on Spoken Language Processing*. Denver, Colorado, USA. September 16-20, 2002. P. 2041-2044.

229. Marak P., Hambalik A. Fingerprint recognition system using artificial neural network as feature extractor: design and performance evaluation. *Tatra Mountains Mathematical Publications*. 2016. 67(1). P. 117-134.

DOI: 10.1515/tmmp-2016-0035.

230. Marchi E., Ferroni G., Eyben F., Gabrielli L., Squartini S., Schuller B. Multi-resolution linear prediction based features for audio onset detection with bidirectional LSTM neural networks. *IEEE International Conference on Acoustics, Speech and Signal Processing*. 2014. P. 2164-2168.

231. Martens J. The Hermite transform-theory. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1990. Vol. 38. No. 9. P. 1595-1606.

232. Matsumoto D., Hwang H., Ekman P. Encyclopedia of Behavioral Neuroscience: Second Edition. 2021. P. 671-677.

233. Mauch M., Dixon S. Approximate Note Transcription for the Improved Identification of Difficult Chords. *11th International Society for Music Information Retrieval Conference*. Utrecht, Netherlands. August 9-13, 2010. P. 135-140.

234. Michelsanti D., Guichi Y., Ene A., Stef R., Nasrollahi K., Moeslund T. Fast Fingerprint Classification with Deep Neural Network. *International Conference on Computer Vision Theory and Applications.* Porto, Portugal. 2017.

235. Milborrow S., Nicolls F. Active Shape Models with SIFT Descriptors and MARS. *International Conference on Computer Vision Theory and Applications*. Lisbon, Portugal. 05-08 January 2014. P. 380-387.

236. McLaren M., Lei Y., Scheffer N., Ferrer L. Application of convolutional neural networks to speaker recognition in noisy conditions. *15th Annual Conference of the International Speech Communication Association*. Singapore. 14-18 September, 2014. P. 686-690.

237. Mikhaylenko V., Tereikovska L. Encoding the output signal of the convolutional neural network intended to recognize handwritten symbol. *International scientific-practical conference of young scientists «Build-Master-Class-2017»*. Kyiv. 16-18 of November, 2017. P. 349-350.

238. Min B., Hyung G., Kang R. Noisy Ocular Recognition Based on Three

Convolutional Neural Networks. *Sensors*. 2017. Vol. 17(12). P. 1-26. DOI:10.3390/s17122933.

239. Miyazawa K., Ito K., Aoki T., Kobayashi K., Nakajima H. An effective approach for iris recognition using phase-based image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2008. Vol. 30, № 10. P. 1741-1756.

240. Mohammadreza H., Mehrbaksh N., Othman I., Ali Z. F., Sarminah S. Authentication systems: A Literature Review and Classification. *Telematics and Informatics.* 2018. Vol. 35, No 5. P. 1491–1511.

241. Muhammad A., Rizwan A., Dong S., Phong H., Muhammad O. IrisDenseNet: Robust Iris Segmentation Using Densely Connected Fully Convolutional Networks in the Images by Visible Light and Near-Infrared Light Camera Sensors. *Sensors*. 2018. Vol. 18(5). P. 1-30. DOI:10.3390/s18051501.

242. Nazarkevych M., Buriachok V., Lotoshynska N., Dmytryk S. Research of Ateb-Gabor filter in biometric protection systems. *IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies*. 2018. Vol.1. P. 310-313. DOI: 10.1109/STC-CSIT.2018.8526607.

243. Nikulchenko A. A., Dorofieiev Yu. I., Lyubchyk L. M. Robust stabilizing inventory control in supply networks under uncertainty of external demand and supply time-delays. *Journal of Computer and Systems Sciences International*. 2014. Vol. 53, No. 5. P. 761-775.

244. Oksiiuk O., Tereikovska L., Tereikovskiy I. Determination of expected output signals of the neural network model intended for image recognition. *4th International Scientific-Practical Conference Problems of Infocommunications, Science and Technology.* 2017. P. 596-599.

245. Oksiiuk O., Tereikovska L., Tereikovskiy I. Adaptation of the neural network model to the identification of the cyberattacks type "denial of service". *14th International Conference on Advanced Trends in Radioelecrtronics,*

Telecommunications and Computer Engineering. 2018. P. 502-505.

246. Oyebade K. Pattern Recognition: Invariance Learning in Convolutional Auto Encoder Network. *International Journal of Image, Graphics and Signal Processing*. 2016. Vol. 8, No.3. P.19-27.

247. Partila P., Voznak M., Tovarek J. Pattern Recognition Methods and Features Selection for Speech Emotion Recognition System. *Scientific World Journal*. 2015. Vol. 2015. P. 1-7. DOI:10.1155/2015/573068.

248. Peralta D., Triguero I., Garcıa S., Saeys Y., Benitez J., Herrera F. On the use of convolutional neural networks for robust classification of multiple fingerprint captures. *International Journal of Intelligent Systems*. 2018. Vol. 33(1). P. 213-230.

249. Penagarikano M., Bordel G. Layered Markov models: a new architectural approach to automatic speech recognition. Proceedings of the IEEE Signal Processing Society Workshop Machine Learning for Signal Processing. 2004. P. 305-314.

250. Pham V., Bluche T., Kermorvant C., Louradour J. Dropout improves recurrent neural networks for handwriting recognition. *IEEE 14th International Conference on Frontiers in Handwriting Recognition*. 2014. P. 285-290.

251. Po-Ming L., Wei-Hsuan T., Tzu-Chien H. The influence of emotion on keyboard typing: an experimental study using visual stimuli. *BioMedical Engineering OnLine.* 2014. Vol. 13. P. 1-12. DOI:10.1186/1475-925X-13-81.

252. Preeti K., Sasikumar M. Recognising Emotions from Keyboard Stroke Pattern. *International Journal of Computer Applications*. 2010. Vol. 11, No 9. P. 1-5.

253. Proenca H. Iris recognition: on the segmentation of degraded images acquired in the visible wavelength. *IEEE Transaction on Pattern Analysis and Machine Intelligence.* 2010. Vol. 32, No. 8. P. 1502-1516.

254. Ranganathan H., Chakraborty S., Panchanathan S. Multimodal emotion

recognition using deep learning architectures. *IEEE Winter Conference on Applications of Computer Vision*. 2016. P. 1-9.

255. Richardson F., Reynolds D., Dehak N. A unified deep neural network for speaker and language recognition. 2015. (preprint arXiv:1504.00923).

256. Roizenblatt R. Iris recognition as a biometric method after cataract surgery. *BioMedical Engineering OnLine*. 2004. Vol. 3. No 1. P. 1-7. DOI:10.1186/1475-925X-3-2.

257. Russell J., Bachorowski J., Fernandez-Dols J. Facial and vocal expressions of emotion. *Annual Review of Psychology*. 2002. Vol. 54. P. 329-349. DOI: 10.1146/annurev.psych.54.101601.145102.

258. Sabour S., Frosst N., Hinton G. Dynamic Routing Between Capsules. *Advances in Neural Information Processing Systems*. 2017. P. 3857-3867.

259. Sadikoglua F., Uzelaltinbulatb S. Biometric retina identification based on neural network. *12th International Conference on Application of Fuzzy Systems and Soft Computing*. 29-30 August 2016. Vienna, Austria. P. 26-33.

260. Sahoo T., Patra S. Silence Removal and Endpoint Detection of Speech Signal for Text Independent Speaker Identification. *International Journal of Image, Graphics and Signal Processing*. 2014. Vol. 6 (6). P. 27-35.

261. Saket M., Soumyajit G., Vikram P. Deep Secure: A Fast and Simple Neural Network Based Approach for User Authentication and Identification via Keystroke Dynamics. *International Joint Conference on Artificial Intelligence*. Melbourne, Australia. 2017. P. 34-40.

262. Sanjay R., Mirza S. Iris Recognition System and Analysis Using Neural Networks. *International Journal of Engineering Research & Technology*. 2013. Vol. 2, No 7. P. 1051-1054.

263. Sassi A., Ouarda W., Amar C., Miguet S. Sky-CNN: A CNN-based Learning Approach for Skyline Scene Understanding. *International Journal of*

Intelligent Systems and Applications. 2019. Vol. 11, No 4, P. 14-25.

264. Satyanand S., Abhay K., Kolluri D. Efficient Modelling Technique based Speaker Recognition under Limited Speech Data. *International Journal of Image, Graphics and Signal Processing* 2016. Vol.8, No.11. P. 41-48.

265. Satyanand S., Ajeet S. Accuracy Comparison using Different Modeling Techniques under Limited Speech Data of Speaker Recognition Systems. *Global Journal of Science Frontier Research: F Mathematics and Decision Sciences*. 2016. Vol 16, No 2. P. 1-17.

266. Savchenko L. V., Savchenko A. V. Fuzzy Phonetic Encoding of Speech Signals in Voice Processing Systems. *Journal of Communications Technology and Electronics*. 2019. Vol. 64. No. 3. P. 238-244.

267. Savchenko V. V., Savchenko A. V. Information Theoretic Analysis of Efficiency of the Phonetic Encoding-Decoding Method in Automatic Speech Recognition. *Journal of Communications Technology and Electronics*. 2016. Vol. 4, No 61. P. 430-435.

268. Savchenko V. V. The Principle of the Information-Divergence Minimum in the Problem of Spectral Analysis of the Random Time Series Under the Condition of Small Observation Samples. *Radiophysics and Quantum Electronics*. 2015. Vol. 5, No 58. P. 373-379.

269. Savchenko V. V. Enhancement of the Noise Immunity of a Voice-Activated Robotics Control System Based on Phonetic Word Decoding Method. *Journal of Communications Technology and Electronics*. 2016. Vol. 12, No 61. P. 1374-1379.

270. Savvides M., Kumar B., Khosla P. Eigenphases vs Eigenfaces. *17th International Conference on Pattern Recognition*. 2004. Vol. 3. P. 810-813.

271. Schmidt S. J. What Does Emotion Have to Do with Learning? *Journal of Food Science Education*. 2017. Vol. 16. P. 64-66.

272. Seyed M., Marjan A. Face emotion recognition system based on fuzzy logic using algorithm improved Particle Swarm. *International Journal of Computer Science and Network Security*. 2016. Vol.16, No.7. P. 157-166.

273. Shankar S., Robertson D., Ioannou Y., Criminisi A., Cipolla R. Refining Architectures of Deep Convolutional Neural Networks. *IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV. 2016. P. 2212-2220.

274. Shanshan L. Facial expression recognition algorithm based on local Gabor wavelet automatic segmentation. *Journal of Computer Applications*. 2009. Vol. 11. P. 29-37.

275. Sharada V., Mahesh S. Robust spectral features for automatic speaker recognition in mismatch condition. *Second International Symposium on Computer Vision and the Internet.* Procedia Computer Science, 2015. Vol. 58. P. 272-279.

276. Sharonova N., Doroshenko A., Cherednichenko O. Issues of fact-based information analysis. *CEUR Workshop Proceedings*. 2018. Vol. 2136. P. 11-19.

277. Sharonova N., Kyrychenko I., Tereshchenko G. Application of big data methods in E-learning systems. *CEUR Workshop Proceedings*. 2021. Vol. 2870. P. 1302-1311.

278. Schupp H., Stockburger J., Codispoti M., Junghofer M., Weike A., Hamm A. Selective visual attention to emotion. *Journal of Neuroscience*. 2007. Vol. 27 (5). P. 1082-1089. DOI: 10.1523/JNEUROSCI.3223-06.2007.

279. Sitnikov V. S., Bilenko A.A. Classification of wavelet functions. *Труды Одесского политехнического университета*. Одесса. 2008. Вып. 1(29). С. 168-171.

280. Snyder D., Ghahremani P., Povey D., Garcia-Romero D., Carmiel Y., Khudanpur S. Neural network-based speaker embeddings for end-to-end speaker verification. *IEEE Spoken Language Technology Workshop*. 2016.

281. SorokinV. N., Tsyplikhin A. I. Speaker verification using the spectral

and time parameters of voice signal. *Journal of Communications Technology and Electronics*. 2010. Vol. 55, No 12. P. 1561-1574.

282. Starasotnikau N. A., Feodortsau R. V. Accuracy comparison of algorithms for determination of image center coordinates in optoelectronic devices. *Science & Technique.* 2018. Vol. 17, No 1. P. 79-86.

283. Steinbuch M. Wavelet theory and applications. A literature study. Eindhoven University of Technology. 2005. 50 p.

284. Stephane K., Claude T. Fingerprint Recognition with Artificial Neural Networks: Application to E-Learning. *Journal of Intelligent Learning Systems and Applications*. 2016. Vol. 8. P. 39-49.

285. Sumi K., Itoyama K., Yoshii K. et al. Automatic Chord Recognition Based on Probabilistic Integration of Chord Transition and Bass Pitch Estimation. *9th International Conference on Music Information Retrieval*. Philadelphia, USA. September 14-18, 2008. P. 39-44.

286. Tai S. et al. Improved semantic representations from tree-structured long short-term memory networks. 2015 (preprint arXiv:1503.00075).

287. Tariq U., Lin K. et al. Emotion Recognition from an Ensemble of Features. *IEEE International Conference on Automatic Face & Gesture Recognition*. 2011. Vol. 42 (4). P. 872-977.

288. Tereikovska L.A., Radchenko K.O., Kharlamov K.O. The application of discrete wavelet transform for forecasting the operational load of the Web server of the distance education system. *Eighth World Congress «Aviation in the XXI-st Century» - «Safety in Aviation and Space Technologies»*. Kyiv.2018. P. 3.2.13-3.2.17.

289. Tereykovska L., Tereykovskiy I., Aytkhozhaeva E., Tynymbayev S., Imanbayev A. Encoding of neural network model exit signal, that is devoted for distinction of graphical images in biometric authenticate systems. *News of the*

national academy of sciences of the republic of Kazakhstan series of geology and technical sciences. 2017. Vol. 6, No 426. P. 217-224.

290. Tereikovska L., Tereikovskyi I., Beketova A., Karaman G., Makovetska N. Recognition of Speaker's Emotion by Squeezenet Convolutional Neural Network. *Journal of Theoretical and Applied Information Technology*. March 2021. Vol. 99, No 5. P. 1139-1148.

291. Tereykovskaya L., Tereykovskiy I., Aytkhozhaeva E., Tynymbayev S., Imanbayev A. Improvement of learning efficiency of the neural networks, intended for recognition of graphic images in systems of biometric authentication. *Computer Modelling & New Technologies*. 2017. Vol. 21, Is. 2. P. 54-57.

292. Tereikovska L., Tereikovskyi I., Mussiraliyeva S., Akhmed G. Recognition of emotions by facial geometry using a capsule neural network. *International Journal of Civil Engineering and Technology*. April 2019.Vol. 10, Is. 03. P. 1424-1434.

293. Tereikovska L., Tereikovskyi I. Optimization of Convolutional Neural Network Structure for Biometric Authentication by Face Geometry. *Управління розвитком технологій* : матеріали V Міжнар. наук.-практ. конф. м. Київ. 19-20 травня 2018 р. С 42-43.

294. Terejkowski I., Aitchanov B., Bapiev I., Terejkowska L., Pogorelov V. Calculation of expected output signal of neural network model for detecting of cyber-attack on network resources. *15th International Scientific Conference Information Technologies and Management*. Riga, Latvia. April 27-28, 2017. P. 59-62.

295. Tereikovskyi I., Chernyshev D., Tereikovska L., Mussiraliyeva S., Akhmed G. The procedure for the determination of structural parameters of a convolutional neural network to fingerprint recognition. *Journal of Theoretical and Applied Information Technology*. April 2019. Vol. 97, No 8. P. 2381-2392.

296. Tereikovskyi I., Mussiraliyeva S., Kosyuk Y., Bolatbek M., Tereikovska

L. An experimental investigation of infrasound influence hard drives of a computer system. *International Journal of Civil Engineering and Technology*. June 2018. Vol. 9, Iss. 6. P. 1558-1566.

297. Tereikovskyi I., Mussiraliyeva S., Tereikovska L., Chernyshev D., Nyussupov A., Abaiuly Y. The Procedure of Adapting the Design Parameters of the Convolutional Neural Network During the Speaker's Emotions Recognition. *International Conference on Smart Information Systems and Technologies.* Nur-Sultan. 28-30 April 2022. P. 68-73.

298. Tereikovskyi I., Tereikovska L., Korystin O., Mussiraliyeva S., Sambetbayeva A. User Keystroke Authentication and Recognition of Emotions Based on Convolutional Neural Network. *Advances in Intelligent Systems and Computing*. 2020. Springer, Cham. Vol. 1126. P. 283-292.

299. Tereikovskyi I., Tereykovska L., Mussiraliyeva S., Tsiutsiura M., Achkoski J. Markov Model of Unsteady Profile of Normal Behavior of Network Objects of Computer Systems. *Conflict Management in Global Information Networks*. 2019. *CEUR Workshop Proceedings*. Vol. 2588. P. 140-152.

300. Toliupa S., Kulakov Y., Tereikovskyi I., Tereikovskyi O., Tereikovska L., Nakonechnyi V. Keyboard Dynamic Analysis by Alexnet Type Neural Network. *IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering*. 2020. P. 416-420.

301. Toliupa S., Nakonechnyi V., Tereikovskyi I., Tereikovska L., Korystin O. One-Periodic Template Marks Model of Normal Behavior of the Safety Parameters of Information Systems Networking Resources. *IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology*. 2019. P. 764-768.

302. Toliupa S., Tereikovska L., Korystin O., Chernyshev D., Tereikovskyi I. Low-Resource Convolution Neural Network for Keyboard Recognition of the User.

IEEE International Conference on Advanced Trends in Information Theory. 2019. P. 222-226.

303. Toliupa S., Tereikovska L., Tereikovskyi I., Doszhanova A., Alimseitova Z. Procedure for Adapting a Neural Network to Eye Iris Recognition. *IEEE International Conference on Problems of Infocommunications, Science and Technology*. 2020. P. 167-171.

304. Toliupa S., Tereikovskiy I., Dychka I., Tereikovska L., Trush A. The Method of Using Production Rules in Neural Network Recognition of Emotions by Facial Geometry. *3rd International Conference on Advanced Information and Communications Technologies*. 2019. P. 323-327.

305. Toser Z., Jeni L., Lorincz A., Cohn J. Deep learning for facial action unit detection under large head poses. *European Conference on Computer Vision*. Springer, 2016. P. 359-371.

306. Trojahn M., Arndt F., Weinmann M., Ortmeier F. Emotion Recognition Through Keystroke Dynamics on Touchscreen Keyboards. *15th International Conference on Enterprise Information Systems.* 4-7 July 2013. P. 31-37. DOI: 10.5220/0004415500310037.

307. Turaga S., Murray J., Jain V., Roth F., Helmstaedter M., Briggman K., Denk W., Seung H. Convolutional networks can learn to generate affinity graphs for image segmentation. *Neural Comput*. 2010. Vol. 22(2). P. 511-538.

308. Um E., Plass J., Hayward E., Homer B. Emotional design in multimedia learning. *Journal of Educational Psychology*. 2012. Vol. 104(2). P. 485-498.

309. Usham D., Vinita F., Sandeep P. Amanda F. A Neural Network Based Iris Recognition System for Personal Identification. *Journal on Soft Computing*. 2010. No 2. P. 78-83.

310. Variani E., Lei X., McDermott E., Lopez Moreno I., Gonzalez-Dominguez J. Deep neural networks for small footprint text-dependent speaker

verification. *IEEE International Conference on Acoustic, Speech and Signal Processing*. 2014. P. 4052-4056.

311. Vaziri G., Almasganj F., Behroozmand R. Pathological assessment of patients' speech signals using nonlinear dynamical analysis. *Computers in Biology and Medicine*. 2010. Vol. 40, No 1. P. 54-63.

312. Viola P. Robust realtime face detection. *International Journal of Computer Vision*. 2004. Vol. 57, No 2. P. 137-154.

313. Vysotska O., Davydenko A. Keystroke Pattern Authentication of Computer Systems Users as One of the Steps of Multifactor Authentication. *Advances in Computer Science for Engineering and Education II*. 2019. Springer, Cham. Vol 938. P. 356-368.

314. Wang R., Han C., Wu Y., Guo T. Fingerprint Classification Based on Depth Neural Network. 2014 (preprint 1-14arXiv:1409.5188v1).

315. Wang Z., Li Y., Wang S., Ji Q. Capturing global semantic relationships for facial action unit recognition. *IEEE International Conference on Computer Vision*. 2013. P. 3304-3311.

316. Wu T., Butko N., Ruvolo P., Whitehill J., Bartlett M., Movellan J. Multi-Layer Architectures for Facial Action Unit Recognition. *IEEE Transactions on Systems, Man and Cybernetics*. 2012. Vol. 42, No 4. P. 1027-1038.

317. Xiaofeng, L., Shengfei, Z., Shengwei, Y. Continuous authentication by free-text keystroke based on CNN plus RNN. *Procedia Computer Science*. 2019. Vol. 147. P. 314-318.

318. Yao S., Hu C., Sun M. Prediction of Web Traffic Based on Wavelet and Neural Network. *6th World Congress on Intelligent Control and Automation.* 2006. Vol. 1. P. 4026-4028.

319. Yao S., Hu C., Peng W. Server Load Prediction Based on Wavelet Packet and Support Vector Regression. *International Conference on Computational*

Intelligence and Security. 2006. Vol. 2. P. 1016-1019.

320. Yilong Y., Xianjing M., Gongping Y., Xiaoming X. Retinal Identification Based on an Improved Circular Gabor Filter and Scale Invariant Feature Transform. *Sensors*. 2013. Vol. 13(7). P. 9248-9266. DOI: 10.3390/s130709248.

321. Yu Y., Acton S. Speckle Reducing Anisotropic Diffusion. *IEEE Transactions on Image Processing*. 2002. Vol. 11, No 11. P. 1260-1270.

322. Yudin O., Toliupa S., Korchenko O., Tereikovska L., Tereikovskyi I., Tereikovskyi O. Determination of Signs of Information and Psychological Influence in the Tone of Sound Sequences. *IEEE 2nd International Conference on Advanced Trends in Information Theory.* 2020. P. 276-280.

323. Zhang W., Deng Y., He L., Liu J. Variant Time-Frequency Cepstral Features for Speaker Recognition. *Interspeech*. 2010. P. 2122-2125. DOI: 10.21437/Interspeech.2010-160.

324. Zhao K., Chu W., Zhang H. Deep region and multilabel learning for facial action unit detection. *IEEE Conference on Computer Vision and Pattern Recognition*. 2016. P. 3391-3399.

325. Zhengbing H., Tereykovskiy I., Tereykovska L., Pogorelov V. Determination of structural parameters of multilayer perceptron designed to estimate parameters of technical systems. *International Journal of Intelligent Systems and Applications*. 2017. No 10. P. 57-62.

326. Zijiang Y. Research on Server Load Prediction Based on Wavelet Packet Theory. *First IEEE International Symposium on Information Technologies and Applications in Education.* 2007. P. 610-613.

327. Zoubida L., Adjoudj R. Integrating Face and the Both Irises for Personal Authentication. *International Journal of Intelligent Systems and Applications*. 2017. Vol.9, No.3. P. 8-17.

ДОДАТОК А
АКТИ ВПРОВАДЖЕНЬ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОГО ДОСЛІДЖЕННЯ

"ЗАТВЕРДЖУЮ"
Проректор з наукової роботи
та інноваційного розвитку
Київського національного
університету
будівництва і архітектури

О.Ю. Ковальчук

«__» _____ 2022 року

АКТ ВПРОВАДЖЕННЯ

результатів дисертаційних досліджень здобувача

Терейковської Людмили Олексіївни

Цим актом підтверджується, що метод формування вхідного поля нейромережевої моделі аналізу клавіатурного почерку слухача системи дистанційного навчання, розроблений в рамках дисертаційної роботи «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання», використано при виконанні НДР «Нейромережеві методи, моделі та засоби інформаційної технології управління емоційним станом слухачів системи дистанційного навчання» (№ДР 0122U001768). НДР виконувалась на кафедрі інформаційних технологій проєктування та прикладної математики Київського національного університету будівництва і архітектури МОН України під керівництвом Терейковської Л.О. В результаті використання даного методу ефективність засобів автоматизованого розпізнавання емоційного стану за клавіатурним почерком підвищилась приблизно в 1,1 рази.

Завідувач кафедрою інформаційних
технологій проєктування
та прикладної математики,
д.т.н., професор

Олександр ТЕРЕНТЬЄВ

"ЗАТВЕРДЖУЮ"
Проректор з наукової роботи та
інноваційного розвитку
Київського національного університету
будівництва і архітектури

О.Ю. Ковальчук

«__» _____ 2022 року

АКТ

про впровадження результатів дисертаційної роботи
Терейковської Людмили Олексіївни
«Методологія автоматизованого розпізнавання емоційного стану слухачів
системи дистанційного навчання»
на здобуття наукового ступеня доктора технічних наук

Цей акт складено у тому, що під час роботи над НДР №ДР 0121U110859 «Методи та засоби розпізнавання емоцій диктора», яка виконувалась на кафедрі інформаційних технологій проєктування та прикладної математики Київського національного університету будівництва і архітектури МОН України під науковим керівництвом Терейковської Л.О., реалізовано запропонований у дисертації метод формування вхідного поля нейромережевої моделі аналізу голосового сигналу.

Запропоноване у результаті виконання наукових досліджень рішення забезпечило приблизно на 10% підвищення ефективності засобів автоматизованого розпізнавання емоційного стану за голосом.

Завідувач кафедрою інформаційних
технологій проєктування
та прикладної математики,
д.т.н., професор

Олександр ТЕРЕНТЬЄВ

"ЗАТВЕРДЖУЮ"
Проректор з наукової роботи та
інноваційного розвитку
Київського національного університету
будівництва і архітектури

О.Ю. Ковальчук

«10» листопада 2022 року

АКТ

впровадження результатів дисертаційної роботи Терейковської Людмили Олексіївни на тему «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання»

Цим актом підтверджується, що результати наукових досліджень у рамках дисертаційної роботи «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання» використано на кафедрі інформаційних технологій проєктування та прикладної математики при виконанні НДР №ДР 0121U110668 «Теоретичні основи інформаційної технології розпізнавання психоемоційного стану слухачів системи дистанційного навчання» (науковий керівник Терейковська Л.О.). Впроваджено запропоновані підходи до визначення конструктивних параметрів, навчання та мінімізації ресурсоємності нейромережевих моделей, що забезпечило можливість розробки ефективних методів розпізнавання емоційного стану.

Завідувач кафедрою інформаційних
технологій проєктування
та прикладної математики,
д.т.н., професор

Олександр ТЕРЕНТЬЄВ

"ЗАТВЕРДЖУЮ"
Декан факультету прикладної математики
Національного технічного університету
України "Київський політехнічний
інститут імені Ігоря Сікорського"

_____ І.А. Дичка

«18» травня 2022 року

АКТ ВПРОВАДЖЕННЯ

результатів дисертаційної роботи Терейковської Людмили Олексіївни «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання» на здобуття наукового ступеня доктора технічних наук

Цим актом підтверджується, що метод нейромережевого розпізнавання емоцій, розроблений в рамках дисертаційної роботи Терейковської Л.О. «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання», використано на кафедрі системного програмування і спеціалізованих комп'ютерних систем факультету прикладної математики Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського" при виконанні НДР «Методи, моделі та комп'ютерні засоби виявлення деструктивного впливу в медіапросторі» (№ДР 0121U110662). Використання даного методу забезпечило можливість виявлення деструктивного впливу в медіапросторі за рахунок оцінки емоцій аудиторії.

Завідувач кафедрою системного
програмування і спеціалізованих
комп'ютерних систем,
д.т.н., професор Віталій РОМАНКЕВИЧ

"ЗАТВЕРДЖУЮ"
Проректор з наукової роботи
та інноваційного розвитку
Київського національного
університету
будівництва і архітектури

_____ О.Ю. Ковальчук

«___» _____ 2022 року

А К Т

впровадження результатів дисертаційної роботи Терейковської Людмили Олексіївни на тему «МЕТОДОЛОГІЯ АВТОМАТИЗОВАНОГО РОЗПІЗНАВАННЯ ЕМОЦІЙНОГО СТАНУ СЛУХАЧІВ СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ» на здобуття наукового ступеня доктора технічних наук

Комісія у складі: голова – завідувач кафедри управління проєктами, д.т.н., проф. Бушуєв С.Д.; члени комісії – професор кафедри управління проєктами, д.т.н., проф. Веренич О.В., професор кафедри управління проєктами, д.е.н., проф. Бондар О. А. цим Актом засвідчує, що результати дисертаційного дослідження Терейковської Людмили Олексіївни використані співробітниками кафедри управління проєктами при підготовці і викладанні курсу лекцій і лабораторних робіт навчальної дисципліни «Нейронні мережі та глибинне навчання».

Зокрема впроваджено: новий метод розробки архітектури згорткової нейронної мережі, призначеної для аналізу біометричних параметрів, в якому за рахунок використання запропонованих критеріїв ефективності та принципів адаптації конструктивних параметрів нейронної мережі до умов функціонування, забезпечується можливість зменшення обсягу експериментальних досліджень, пов'язаних з розробкою архітектури згорткової нейронної мережі; новий метод формування навчальних прикладів для нейромережевої моделі аналізу біометричних параметрів, в якому за рахунок обробки навчальних прикладів ймовірнісною нейронною мережею забезпечується можливість врахування в очікуваному вихідному сигналі близькості еталонів визначених класів.

Голова комісії
д.т.н., проф. Сергій БУШУЄВ

Члени комісії
д.т.н., проф. Олена ВЕРЕНИЧ

д.е.н., проф. Олена БОНДАР

"ЗАТВЕРДЖУЮ"
Декан факультету прикладної математики
Національного технічного університету
України "Київський політехнічний
інститут імені Ігоря Сікорського"

_____ І.А. Дичка

«25» травня 2022 року

АКТ ВПРОВАДЖЕННЯ

в навчальний процес кафедри прикладної математики факультету прикладної математики Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського" результатів дисертаційної роботи Терейковської Людмили Олексіївни на тему «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання» на здобуття наукового ступеня доктора технічних наук

Комісія у складі: голова – завідувач кафедри, д.т.н., проф. Чертов О.Р.; члени комісії – доцент кафедри прикладної математики, к.т.н., доц. Олефір О. С., доцент кафедри прикладної математики, к.ф.-м.н., доц. Третиник В.В. склали цей акт про те, що результати дисертаційної роботи Терейковської Л.О. «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання» впроваджені у навчальний процес на кафедрі прикладної математики у 2021-2022 навчальному році при викладанні дисципліни «Методи штучного інтелекту». Зокрема в лекційні матеріали впроваджено методи формування вхідного поля нейромережевих моделей, призначених для аналізу клавіатурного почерку та для аналізу голосового сигналу слухача системи дистанційного навчання.

Голова комісії
д.т.н., проф. Олег ЧЕРТОВ

Члени комісії
к.т.н., доц. Олександр ОЛЕФІР

к.ф.-м.н., доц. Віолета ТРЕТИНИК

"ЗАТВЕРДЖУЮ"
Декан факультету кібербезпеки,
комп'ютерної та програмної інженерії
Національного авіаційного
університету
_____ К.С. Нестеренко
«26» березня 2022 року

АКТ

про впровадження в навчальний процес кафедри безпеки інформаційних технологій факультету кібербезпеки, комп'ютерної та програмної інженерії Національного авіаційного університету результатів дисертаційної роботи Терейковської Людмили Олексіївни на тему «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання»

Комісія у складі: голова – завідувач кафедри безпеки інформаційних технологій, д.т.н., проф. Корченко О.Г.; члени комісії – професор кафедри безпеки інформаційних технологій, к.т.н., проф. Іванченко Є.В., доцент кафедри безпеки інформаційних технологій, к.т.н., доц. Хохлачова Ю. Є. склали цей акт про те, що результати дисертаційної роботи Терейковської Людмили Олексіївни «Методологія автоматизованого розпізнавання емоційного стану слухачів системи дистанційного навчання» впроваджені у навчальний процес на кафедрі безпеки інформаційних технологій у 2021-2022 навчальному році при викладанні дисципліни «Технології штучного інтелекту».

№ з/п	Назва роботи, що впроваджується	Форма впровадження	Ефективність від впровадження
1.	Підходи до визначення конструктивних параметрів нейромережевих моделей	Лекція	Ознайомлення студентів з підходами до побудови архітектури нейромережевої моделі
2.	Метод розробки архітектури згорткової нейронної мережі	Лекція	Ознайомлення студентів з базовими положеннями в області згорткових нейронних мереж
3.	Застосування вейвлет-перетворень для фільтрації вхідного поля нейромережевої моделі	Лекція	Ознайомлення студентів з методами застосування вейвлет-перетворень при формуванні навчальних прикладів

Голова комісії
д.т.н., проф. Олександр КОРЧЕНКО
Члени комісії
к.т.н., проф. Євгенія ІВАНЧЕНКО
к.т.н., доц. Юлія ХОХЛАЧОВА

ДОДАТОК Б
ПРОГРАМНИЙ КОД ЗАСТОСУНКУ ДЛЯ РОЗПІЗНАВАННЯ ЕМОЦІЙ ТА ОСОБИ ЛЮДИНИ ЗА ГОЛОСОМ

```
import soundfile as sf
from python_speech_features import mfcc
import os
import tensorflow
import numpy as np
from python_speech_features import mfcc
from tensorflow.keras.layers import Dense, Activation, Flatten, Dropout, BatchNormalization
from tensorflow.keras.utils import to_categorical
from tensorflow.keras.models import Sequential, Model, load_model
from tensorflow.keras import regularizers, optimizers, backend
import matplotlib.pyplot as plt
import time
mel_count = 24
def emotio_train():
    audio_files = []
    directory="Soundfiles\\res\\dataset1"
    count =0
    y=[]
    for dir in os.listdir(directory):
        for file in os.listdir(os.path.join(directory,dir)):
            if file.endswith(".wav"):
                audio_files.append(os.path.join(directory, dir, file))
                y.append(int(dir))
                count+=1
    y=np.array(y)
    x = []
    for flName in audio_files :
        print(flName)
```

```
            data, samplerate = sf.read(flName)
            data = np.resize(data,[samplerate*2,])
            temp = mfcc(data,samplerate,winlen=0.016, winstep=0.008,
numcep=mel_count, nfft=8000, appendEnergy=True)
            tempArr2=[]
            for j in range(0, 240):
                val = temp[j]
                tempArr1 = []
                for i in range(3,10):
                    tempArr1.append(val[i])
                tempArr2.append(tempArr1)
            x.append(tempArr2)
            count += 1
        x=np.array(x)
        np.random.seed(int(time.clock()))
        test_part = 0.25
        x=np.reshape(x,(x.shape[0],x.shape[1]*x.shape[2]))
        index=np.arange(x.shape[0])
        np.random.shuffle(index)
        x=x[index]
        y=y[index]
        x_test=x[:int(x.shape[0]*test_part)]
        y_test=y[:int(x.shape[0]*test_part)]
        x_train=x[int(x.shape[0]*test_part):]
        y_train=y[int(x.shape[0]*test_part):]
        backend.clear_session()
        model = Sequential()
        model.add(Dense(256, activation='relu', input_shape=(x.shape[1],)))
        model.add(Dense(128, activation='relu'))
        model.add(Dropout(0.25))
```

```
        model.add(Dense(4, activation='softmax'))
        sgd = optimizers.SGD(learning_rate=0.0005, momentum=0.0)
        model.compile(optimizer=sgd, loss='categorical_crossentropy',
metrics=['accuracy'])
        model.summary()
        history = model.fit(x_train, to_categorical(y_train), epochs=50,
batch_size=16, validation_split=0.1)
        y_pred = model.predict(x_test)
        score = model.evaluate(x_test, to_categorical(y_test), verbose=0)
        print(score)
        model.save("emotions_model.h5")
        plt.plot(history.history['accuracy'])
        plt.plot(history.history['val_accuracy'])
        plt.title('Model accuracy')
        plt.ylabel('Accuracy')
        plt.xlabel('Epoch')
        plt.legend(['Train', 'Test'], loc='lower right')
        plt.show()
        plt.plot(history.history['loss'])
        plt.plot(history.history['val_loss'])
        plt.title('Model loss')
        plt.ylabel('Loss')
        plt.xlabel('Epoch')
        plt.legend(['Train', 'Test'], loc='upper right')
        plt.show()
    def emotion_train_2():
        audio_files = []
        directory="Soundfiles\\res\\dataset2"
```

```
count =0
y=[]
for dir in os.listdir(directory):
    for file in os.listdir(os.path.join(directory,dir)):
        if file.endswith(".wav"):
            audio_files.append(os.path.join(directory,dir,file))
            s = file.split('-')
            y.append(int(s[2])-1)
            count+=1
y=np.array(y)
x = []
for flName in audio_files :
    print(flName)
    data, samplerate = sf.read(flName)
    data = np.resize(data,[samplerate*2,])
    temp = mfcc(data,samplerate,winlen=0.016, winstep=0.008,
numcep=mel_count, nfft=8000, appendEnergy=True)
    tempArr2=[]
    for val in temp:
        tempArr1 = []
        for i in range(3,20):
            tempArr1.append(val[i])
        tempArr2.append(tempArr1)
    x.append(tempArr2)
    count += 1
x=np.array(x)
np.random.seed(int(time.clock()))
test_part = 0.1
x=np.reshape(x,(x.shape[0],x.shape[1]*x.shape[2]))
index=np.arange(x.shape[0])
```

```
        np.random.shuffle(index)
        x=x[index]
        y=y[index]
        x_test=x[:int(x.shape[0]*test_part)]
        y_test=y[:int(x.shape[0]*test_part)]
        x_train=x[int(x.shape[0]*test_part):]
        y_train=y[int(x.shape[0]*test_part):]
        backend.clear_session()
        model = Sequential()
        model.add(Dense(256, activation='relu', input_shape=(x.shape[1],)))
        model.add(Dense(128, activation='relu'))
        model.add(Dropout(0.25))
        model.add(Dense(8, activation='softmax'))
        sgd = optimizers.SGD(learning_rate=0.0005, momentum=0.0)
        model.compile(optimizer=sgd, loss='categorical_crossentropy',
metrics=['accuracy'])
        model.summary()
        history = model.fit(x_train, to_categorical(y_train), epochs=150,
batch_size=32, validation_split=0.1)
        y_pred = model.predict(x_test)
        score = model.evaluate(x_test, to_categorical(y_test), verbose=0)
        print(score)
        model.save("emotions2_model.h5")
        plt.plot(history.history['accuracy'])
        plt.plot(history.history['val_accuracy'])
        plt.title('Model accuracy')
        plt.ylabel('Accuracy')
        plt.xlabel('Epoch')
        plt.legend(['Train', 'Test'], loc='lower right')
        plt.show()
```

```
        plt.plot(history.history['loss'])
        plt.plot(history.history['val_loss'])
        plt.title('Model loss')
        plt.ylabel('Loss')
        plt.xlabel('Epoch')
        plt.legend(['Train', 'Test'], loc='upper right')
        plt.show()
    def person_train():
        audio_files = []
        directory="Soundfiles\\res\\dataset2"
        count =0
        y=[]
        for dir in os.listdir(directory):
            s = dir.split('_')
            for file in os.listdir(os.path.join(directory,dir)):
                if file.endswith(".wav"):
                    audio_files.append(os.path.join(directory,dir,file))
                    y.append(int(s[1])-1)
                    count+=1
        y=np.array(y)
        x = []
        for flName in audio_files :
            print(flName)
            data, samplerate = sf.read(flName)
            data = np.resize(data,[2*samplerate,])
            temp = mfcc(data,samplerate,winlen=0.016, winstep=0.008,
numcep=mel_count, nfft=8000, appendEnergy=True)
            tempArr2=[]
            for val in temp:
                tempArr1 = []
```

```
                for i in range(0,20):
                    tempArr1.append(val[i])
                tempArr2.append(tempArr1)
            x.append(tempArr2)
            count += 1
        x=np.array(x)
        np.random.seed(int(time.clock()))
        test_part = 0.1
        x=np.reshape(x,(x.shape[0],x.shape[1]*x.shape[2]))
        index=np.arange(x.shape[0])
        np.random.shuffle(index)
        x=x[index]
        y=y[index]
        x_test=x[:int(x.shape[0]*test_part)]
        y_test=y[:int(x.shape[0]*test_part)]
        x_train=x[int(x.shape[0]*test_part):]
        y_train=y[int(x.shape[0]*test_part):]
        backend.clear_session()
        model = Sequential()
        model.add(Dense(256, activation='relu', input_shape=(x.shape[1],)))
        model.add(Dense(256, activation='relu'))
        model.add(Dropout(0.25))
        model.add(Dense(24, activation='softmax'))
        sgd = optimizers.SGD(learning_rate=0.001, momentum=0.0)
        model.compile(optimizer=sgd, loss='categorical_crossentropy',
metrics=['accuracy'])
        model.summary()
        history = model.fit(x_train, to_categorical(y_train), epochs=150,
batch_size=32, validation_split=0.1)
        y_pred = model.predict(x_test)
```

```
        score = model.evaluate(x_test, to_categorical(y_test), verbose=0)
        print(score)
        model.save("person_model.h5")
        plt.plot(history.history['accuracy'])
        plt.plot(history.history['val_accuracy'])
        plt.title('Model accuracy')
        plt.ylabel('Accuracy')
        plt.xlabel('Epoch')
        plt.legend(['Train', 'Test'], loc='lower right')
        plt.show()
        plt.plot(history.history['loss'])
        plt.plot(history.history['val_loss'])
        plt.title('Model loss')
        plt.ylabel('Loss')
        plt.xlabel('Epoch')
        plt.legend(['Train', 'Test'], loc='upper right')
        plt.show()
    def get_emotion():
        model= load_model("emotions_model.h5")
        x = []
        file = input("Enter file name: ")
        data, samplerate = sf.read('Soundfiles\\res\\'+ file +".wav")
        data = np.resize(data, [samplerate*2, ])
        temp = mfcc(data, samplerate, winlen=0.016, winstep=0.008,
numcep=mel_count, nfft=8000, appendEnergy=True)
        tempArr2 = []
        for j in range(0, 240):
            val = temp[j]
            tempArr1 = []
            for i in range(3, 10):
```

```
                tempArr1.append(val[i])
            tempArr2.append(tempArr1)
        x.append(tempArr2)
        x = np.array(x)
        x = np.reshape(x, (x.shape[0], x.shape[1] * x.shape[2]))
        y = model.predict(x)
        percentage=int(y[0][0]*100)
        print("Sad: ",percentage,"%")
        percentage = int(y[0][1] * 100)
        print("Neutral: ",percentage,"%")
        percentage = int(y[0][2] * 100)
        print("Happy: ",percentage,"%")
        percentage = int(y[0][3] * 100)
        print("Angry: ",percentage,"%")
    def get_person():
        model= load_model("person_model.h5")
        file = input("Enter file name: ")
        data, samplerate = sf.read('Soundfiles\\res\\'+ file +".wav")
        data = np.resize(data, [2 * samplerate, ])
        temp = mfcc(data, samplerate, winlen=0.016, winstep=0.008,
numcep=mel_count, nfft=8000, appendEnergy=True)
        tempArr2 = []
        x = []
        for val in temp:
            tempArr1 = []
            for i in range(0, 20):
                tempArr1.append(val[i])
            tempArr2.append(tempArr1)
        x.append(tempArr2)
        x = np.array(x)
```

```
        x = np.reshape(x, (x.shape[0], x.shape[1] * x.shape[2]))
        y = model.predict(x)
        print("Actor ", np.argmax(y)+1)
    def emDb2():
        audio_files = []
        directory = "Soundfiles\\res\\dataset2"
        count = 0
        y = []
        for dir in os.listdir(directory):
            for file in os.listdir(os.path.join(directory, dir)):
                if file.endswith(".wav"):
                    audio_files.append(os.path.join(directory, dir, file))
                    s = file.split('-')
                    y.append(int(s[2]) - 1)
                    count += 1
        y = np.array(y)
        x = []
        for flName in audio_files:
            print(flName)
            data, samplerate = sf.read(flName)
            data = np.resize(data, [samplerate * 2, ])
            temp = mfcc(data, samplerate, winlen=0.016, winstep=0.008,
numcep=mel_count, nfft=8000, appendEnergy=True)
            tempArr2 = []
            for val in temp:
                tempArr1 = []
                for i in range(3, 20):
                    tempArr1.append(val[i])
                tempArr2.append(tempArr1)
            x.append(tempArr2)
```

```
        count += 1
    x = np.array(x)
    x = np.reshape(x, (x.shape[0], x.shape[1] * x.shape[2]))
    model = load_model("emotions2_model.h5")
    score = model.evaluate(x, to_categorical(y), verbose=0)
    print(score)
while(True):
    print("1 - Train for emotion recognition on dataset1")
    print("2 - Train for emotion recognition on dataset2")
    print("3 - Train for person recognition on dataset2")
    print("4 - Get emotion recognition for wav file (dataset1)")
    print("5 - Get person recognition for wav file (dataset2)")
    print("6 - Exit")
    s=input("Enter your choice: ")
    if (s=="1"):
        emotio_train()
    elif (s=="2"):
        emotion_train_2()
    elif (s=="3"):
        person_train()
    elif (s=="4"):
        get_emotion()
    elif (s=="5"):
        get_person()
    elif (s == "6"):
        break;
    else:
        print("Wrong input")
```