

РЕЦЕНЗІЯ

доцента кафедри інформаційних систем і технологій, кандидата технічних наук, доцента Комісаренко Олени Сергіївни
на дисертаційну роботу Луца Владислава Євгеновича
на тему «Методи і засоби трансформації звукової інформації в жестову мову з використанням нейронної мережі»,
яка представлена на здобуття ступеня доктора філософії у галузі 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки»

Актуальність теми дисертаційної роботи та її зв'язок з науковими програмами, планами, темами

Дисертація Луца Владислава Євгеновича присвячена розробці методів і засобів автоматичного перетворення усного мовлення в текстове представлення, придатне для подальшого перекладу на Українську жестову мову, на основі нейромережевих технологій.

Питання інформаційної доступності для людей з порушеннями слуху має суттєве соціальне значення та досі не отримало належного розв'язання ні в Україні, ні за кордоном. Існуюча мережа перекладачів УЖМ не покриває реальних потреб глухих і слабочуючих громадян, що зумовлює потребу в автоматизованих технологічних рішеннях для подолання цього дефіциту.

Автор довів необхідність комплексного підходу, який включає адаптивну фільтрацію аудіосигналу, автоматичне розпізнавання мовлення, каскадну автокорекцію помилок розпізнавання та концептуально-семантичну лематизацію тексту відповідно до структури жестового словника. Виявлена автором ключова наукова проблема полягає в тому, що класична граматична лематизація не узгоджується з вимогами жестового словника УЖМ — наявні інструменти не враховують концептуально-семантичної організації жестових лексем, де одна лема може охоплювати цілий клас семантично пов'язаних словоформ.

Дисертацію виконано у межах науково-дослідних тем кафедри інформаційних систем і технологій Національного транспортного університету відповідно до планів наукової діяльності, зокрема в рамках теми «Математичні методи, моделі і технології дослідження складних систем», державний реєстраційний номер 0124U003679



Наукова новизна і практичне значення отриманих результатів

Наукова новизна дисертаційної роботи полягає у розробці методів і засобів трансформації звукової інформації на жестову мову, що забезпечують підвищення якості транскрипції та коректність концептуально-семантичної лематизації для УЖМ.

У рамках дослідження *вперше*:

— запропоновано та реалізовано каскадний підхід до постобробки результатів розпізнавання мовлення у задачі перекладу на УЖМ, що поєднує словниковий пошук та неймережевий fallback на основі character-level Seq2Seq;

— запропоновано технологію автоматизованого формування тренувального датасету реальних помилок розпізнавання через TTS-Whisper pipeline, що дозволяє отримувати предметно-специфічні дані для донавчання моделі без ручної обробки.

Удосконалено:

— технологію фільтрації аудіосигналів для ASR-систем шляхом розробки адаптивного багатокомпонентного каскаду з параметричною конфігурацією, оптимізованого під голосовий діапазон;

— підхід до лематизації флективних мов шляхом впровадження character-level Seq2Seq архітектури з двонаправленим GRU-кодувальником та механізмом уваги, адаптованої до концептуально-семантичного відображення для УЖМ.

Набуло подальшого розвитку:

— технологія конфігурованих ланцюжків обробки для низькоресурсних мов із флективною морфологією у задачах перекладу на жестову мову;

— підхід до контекстної корекції помилок розпізнавання на рівні речення шляхом адаптації байтової Transformer-моделі до специфіки українських ASR-помилки.

Практична значущість отриманих результатів полягає у розробці програмного комплексу для автоматичного перекладу українського мовлення на УЖМ у реальному часі. Каскад аудіофільтрації забезпечує ефективне зниження рівня шуму при збереженні акустичних характеристик голосу. Підсистема автокорекції суттєво перевищує наявні аналоги за точністю виправлення помилок розпізнавання при прийнятному часі обробки. Підсистема концептуально-семантичної лематизації демонструє суттєво вищу точність порівняно з граматичними лематизаторами

загального призначення. Комплексне тестування підтвердило практичну придатність системи для застосування в режимі реального часу.

Обґрунтованість наукових положень, висновків і рекомендацій, та достовірність отриманих у роботі результатів підтверджується коректним вибором математичних моделей і методів, детальним порівняльним аналізом з відомими аналогами та верифікацією на реальних мовленнєвих даних із застосуванням стандартних метрик оцінки якості.

Повнота опублікування основних положень дисертаційної роботи

За результатами дисертаційного дослідження підготовлено 23 наукових праць, в тому числі 5 статей у фахових наукових виданнях.

Наукові положення та новизна дисертації були презентовані та обговорені на 18 наукових заходах всеукраїнського й міжнародного рівня, що засвідчено відповідними публікаціями тез.

Відсутність порушення академічної доброчесності

Дисертаційна робота Луца В.Є. була перевірена за допомогою програмного забезпечення з метою виявлення текстових запозичень. За результатами перевірки встановлено відсутність плагіату, а також підтверджено коректність оформлення посилань на використані джерела.

Структура та зміст дисертаційного дослідження, рівень виконання поставленого наукового завдання

Перший розділ присвячено аналізу соціального контексту та актуальності проблеми, огляду вітчизняних та міжнародних досліджень у сфері автоматичного розпізнавання мовлення і лематизації. Виявлено принципову прогалину в існуючих підходах і сформульовано мету, задачі та наукову новизну дослідження.

Другий розділ містить теоретичні основи та методи трансформації мовлення на жестову мову. Розглянуто теоретичні засади автоматичного розпізнавання мовлення, методи корекції помилок розпізнавання, методи концептуально-семантичної лематизації, методи цифрової фільтрації аудіосигналів та синтезу жестового відображення. Обґрунтовано вибір компонентів системи та описано програмні засоби реалізації.

Третій розділ присвячено практичній реалізації та верифікації підсистем: адаптивної аудіофільтрації, автокорекції помилок розпізнавання та концептуально-

семантичної лематизації. Для кожної підсистеми наведено результати порівняльного тестування.

Четвертий розділ описує інтеграцію підсистем у єдиний ланцюжок обробки реального часу з графічним інтерфейсом. Наведено результати комплексного тестування, аналіз розподілу обчислювального часу та оцінку досягнутих показників якості. Узагальнені результати підтверджують ефективність запропонованих методів і засобів для вирішення задачі трансформації звукової інформації в жестову мову.

Узагальнені результати підтверджують ефективність запропонованих методів і засобів для вирішення задачі трансформації звукової інформації в жестову мову.

Зауваження та побажання до дисертаційної роботи

У цілому дисертаційна робота Луца Владислава Євгеновича заслуговує позитивної оцінки, проте необхідно зазначити деякі зауваження.

1. У роботі поняття «концептуально-семантична лематизація» використовується як центральний науковий термін, однак його чітке формальне визначення у тексті відсутнє. Рекомендується у Розділі 2 навести авторське визначення цього поняття та обґрунтувати його відмінність від суміжних термінів («граматична лематизація», «словникова нормалізація», «концептуальне відображення»).
2. Датасет концептуально-семантичних лем формувався з використанням великої мовної моделі (Ollama) з подальшою валідацією через ruyorphy3. Однак процедура верифікації якості самого датасету — зокрема відсоток помилково розмічених пар — у роботі не описана. Рекомендується навести оцінку якості датасету, наприклад через ручну перевірку репрезентативної вибірки.
3. У розділі 4 аналіз якості системи проводився лише для одного диктора. Рекомендується розширити тестову вибірку за рахунок записів щонайменше трьох-п'яти дикторів різної статі, вікової категорії та регіону, оскільки якість розпізнавання мовлення суттєво залежить від індивідуальних акустичних особливостей. Наведені зауваження мають рекомендаційний характер і не зменшують значущості отриманих наукових результатів.

Загальні висновки

Аналіз дисертаційної роботи свідчить про те, що вона є цілісним науковим дослідженням, виконаним на високому теоретичному та прикладному рівні. Всі результати, що мають елементи наукової новизни, обґрунтовані автором і підтверджені експериментальними дослідженнями, що дало змогу сформулювати практичні рекомендації щодо розробки систем автоматичного перекладу мовлення на жестову мову.

Тематика дослідження повністю відповідає спеціальності 122 «Комп'ютерні науки» галузі знань 12 «Інформаційні технології» та сприяє подальшому розвитку як прикладних, так і теоретичних положень у сфері нейромережевої обробки мовлення та інклюзивних інформаційних технологій.

Беручи до уваги рівень наукової новизни, достовірність отриманих результатів, використання сучасного методичного інструментарію, а також наявність належної публікаційної активності, можна зробити висновок, що дисертаційна робота й оприлюднені за її результатами наукові праці повністю відповідають вимогам, визначеним пунктами 6-9 Постанови Кабінету Міністрів України від 12 січня 2022 року №44 «Про затвердження Порядку присудження ступеня доктора філософії». Автор дисертації, Луц Владислав Євгенович, продемонстрував достатній рівень володіння методологією наукових досліджень і заслуговує на присудження наукового ступеня доктора філософії в галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Доцент кафедри інформаційних
систем і технологій
Національного транспортного
університету,
кандидат технічних наук, доцент



Олена КОМІСАРЕНКО

