

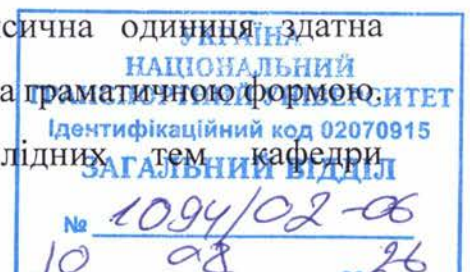
ВІДГУК ОФІЦІЙНОГО ОПОНЕНТА

декана факультету комп'ютерних наук
Харківського національного університету радіоелектроніки,
кандидата технічних наук, доцента
Золотухіна Олега Вікторовича
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. Дисертаційна робота Луца В.Є. пов'язана з розв'язанням задачі перетворення усного мовлення на текст із подальшою можливістю його перекладу на українську жестову мову (УЖМ). Для цього автор застосовує сучасні нейромережеві методи обробки сигналів і природної мови. Обмежена доступність інформації для людей з порушеннями слуху й дотепер залишається невирішеною соціальною проблемою, причому не лише в Україні, а й у світі загалом.

Кількість перекладачів УЖМ значно поступається реальному попиту з боку глухих і слабочуючих людей, тож створення автоматизованих технологічних інструментів, здатних хоча б частково зняти цю напругу, є нагальним завданням сьогодення. Дисертант аргументовано доводить, що ефективно розв'язання цієї задачі можливе лише за умови комплексного підходу, що поєднує в собі адаптивну обробку звукового сигналу, розпізнавання мовлення, багаторівневе виправлення помилок розпізнавання та зведення слів до концептуальних лексичних одиниць жестового словника. Окремої уваги заслуговує сформульована автором наукова проблема: усталені методи граматичної лематизації не узгоджуються з логікою побудови жестового словника УЖМ, адже не враховують того, що в жестовій мові одна лексична одиниця здатна об'єднувати цілу групу слів, споріднених за змістом, а не за граматичною формою.

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.

Структура та зміст дисертаційного дослідження, рівень виконання поставленого наукового завдання. Дисертаційна робота складається зі вступу, чотирьох основних розділів та висновків.

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.

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 нейромережевий 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 флексивних 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 флексивною 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-помилوک.

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.

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.

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.

Зауваження та побажання до дисертаційної роботи

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

1. У роботі описано двонаправне зчитування слова кодувальником (зліва направо і справа наліво одночасно) для задачі лематизації. Не пояснено, наскільки суттєвим є внесок саме двонапрямого читання порівняно з однонапрямним, тобто чи перевірялося окремо, наскільки погіршиться результат без цієї особливості архітектури.

2. У дисертації не проаналізовано, як змінюється якість роботи нейромережових моделей залежно від довжини вхідного речення, чи однаково добре система обробляє короткі репліки та розгорнуті багатослівні висловлювання.

3. У роботі використано символний, а не словниковий підхід до подання тексту всередині нейромережових моделей. Не обговорено потенційні недоліки такого вибору, зокрема, чи не втрачається рівень смислового зв'язку між словами через те, що модель "бачить" лише окремі літери, а не слова цілком.

Наведені зауваження мають дискусійний характер і не позначаються на позитивній оцінці дисертації та практичній цінності її 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.

Декан факультету комп'ютерних наук
Харківського національного університету
радіоелектроніки, к.т.н., доцент

Олег ЗОЛОТУХІН

ПІДПИС ЗАСВІДЧУЮ:
Факівець І.каг. ВК
Начальник відділу кадрів
"10" 08 2026



Золіне З.І.