

## **РЕЦЕНЗІЯ**

на дисертаційну роботу

Борисова Гліба Олександровича

на тему «Адаптивні системи оброблення акустичної інформації для створення персоналізованого медіаконтенту»,

представлену на здобуття ступеня доктора філософії

в галузі знань 17 «Електроніка та телекомунікації»

за спеціальністю 171 «Електроніка»

### **Актуальність теми дисертації.**

Створення якісного персоналізованого медіаконтенту, насамперед мовної записаної інформації, може бути доцільним у випадках, коли потрібно, наприклад, розробити електронні пристрої для запису аудіокниг з текстового джерела (системи озвучування розмовного тексту), при налаштуванні та створенні голосових асистентів чи помічників. І тут, окрім збереження фонетичних особливостей мови, а саме акценту, дикції, інтонації, артикуляції необхідно забезпечити підтримку записів української мовою з точним передаванням місцевого діалекту. Існуючі розробки від компаній Google чи Amazon ці нюанси вимови та наголосів українських слів у словосполученнях найчастіше просто не враховують.

У сучасних технологіях обробки акустичної інформації нейронні мережі відіграють ключову роль, завдяки своїй здатності адаптуватися до різних умов і вирішувати складні завдання. Ці технології дозволяють здійснювати високоточний аналіз мовленнєвих сигналів, ефективне шумозаглушення, розпізнавання мовлення, синтез мовлення та навіть клонування голосу. Нейронні мережі здатні виявляти специфічні особливості акустичних сигналів і адаптуватися до їх складних структур. Тобто, застосування нейронних мереж являє собою актуальну задачу, адже такий підхід дозволяє реалізувати системи розпізнавання мови, синтезу мови з урахуванням цих особливостей.

Також нейронні мережі можна використати при дослідженні аномалій у випадках функціонування різних машин та механізмів у промисловості. Їх виявлення дозволить уникнути відмов при роботі відповідного обладнання. Іншим аспектом застосування мереж може стати їх впровадження в системах відеонагляду, причому не лише в якості основи голосових асистентів з підтримкою української мови, а й коли необхідно, задля підвищення рівня захисту території, швидко отримати аналіз і реакцію на незвичні звукові сигнали (вторгнення у приміщення, розбиття скла, тощо).



## **Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.**

Наукова новизна результатів дисертаційного дослідження полягає в наступному:

- вперше розроблено систему ідентифікації за голосом, яка є стійкою до штучно підробленого голосу і показує високу точність схожості з еталонним записом відразу за 4 критеріями.

- вперше побудовано акустичну модель розпізнавання мовних сигналів з підтримкою нейронної мережі, яка дозволяє в якості вхідної інформації використовувати українські словосполучення. Для її реалізації розроблено змінену рекурентну нейронну мережу, яка вирізняється тим, що за рахунок вбудованої пам'яті в структурі етап навчання та тестування нейронної мережі моделі можна проводити одночасно.

- удосконалено програмний алгоритм дереверберації записаних аудіо сигналів з адитивним додаванням шуму, де використано згорткову нейронну мережу за архітектурою U-Net і яка адаптована до запису не тільки тестових сигналів типу "сплеск" або "постріл", але й словосполучень українською мовою.

- набуло подальшого розвитку створення систем клонування голосу за рахунок введення послідовно трьох попередньо навчених нейронних мереж. Такий підхід дозволив зберегти акцент, інтонаційні та інші фонетичні особливості у синтезованих фразах як англійської, так і української мов.

Наукові дослідження були виконані здобувачем на кафедрі Акустичних та мультимедійних електричних систем КПІ ім. Ігоря Сікорського в рамках ініціативної теми "Автоматизована система суб'єктивного та об'єктивного оцінювання якості мовлення" (№ держреєстрації 0116U008959).

Отже, в дисертаційній роботі поставлене наукове завдання щодо розроблення різних за призначенням адаптивних електронних систем оброблення акустичної інформації з підтримкою нейронних мереж виконано повністю, здобувач повною мірою оволодів методологією наукової діяльності.

## **Оцінка змісту дисертації, її завершеність та дотримання принципів академічної доброчесності.**

За своїм змістом дисертаційна робота здобувача Борисова Г.О. повністю відповідає Стандарту вищої освіти зі спеціальності 171 «Електроніка» та напрямкам досліджень відповідно до освітньої програми «Електроніка».

Дисертаційна робота є завершеною науковою працею і свідчить про наявність особистого внеску здобувача у науковий напрям розроблення електронних засобів оброблення аудіо інформації з використанням нейронних мереж.



Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадіння, можна зробити висновок, що дисертаційна робота Борисова Гліба Олександровича є результатом самостійних досліджень здобувача і не містить елементів фальсифікації, компіляції, фабрикації, плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають належні посилання на відповідне джерело.

### **Мова та стиль викладення результатів.**

Дисертаційна робота написана українською мовою.

Робота вирізняється доступністю викладення основних положень, стиль мовлення є науковим з використанням стандартної термінології, яка прийнята в галузі електроніки та прикладної акустики.

Дисертація складається з вступу, трьох розділів, висновків, списку літератури та додатків. Робота містить 29 рисунків та 21 таблицю. Загальний обсяг дисертації складає 137 сторінок.

У вступі містяться дані про актуальність дослідження, її мету та завдання. Додатково зазначено методи дослідження та наукова новизна одержаних результатів.

В першому розділі роботи зазначено теоретичні дані щодо принципів та підходів оброблення аудіо інформації. Зокрема, визначено базові співвідношення окремих методів аналізу мовлення, а саме, формули які пояснюють аналіз лінійного передбачення, частотний аналіз, та кепстральний аналіз сигналів. Для розбірливості мовлення, зазначено, що її можна описати через емпіричну функцію і ця характеристика є важливою для різних галузей досліджень, інженерії та діагностики, і дозволяє кількісно оцінити різноманітні явища, а точніше якість записів, пристроїв для комунікації та відтворення, реверберацію у залах, характеристики порушення слуху, користь від використання слухових апаратів або поєднання цих факторів. Як один з підходів для оброблення аудіо сигналів запропоновано використання нейронних мереж, завдяки чому для користувача стає доступним простий і ефективний інструментарій з налаштування та адаптації цих мереж у випадку, коли результат оброблення сигналу потребує забезпечення певної якості відтворення, як от у випадку роботи з мовними сигналами.

Другий розділ роботи присвячений розгляду засад адаптивних систем оброблення акустичної інформації, і розглянуто основні принципи побудови таких систем. Описано ключові риси таких систем, включаючи модулі вилучення ознак, класифікації та синтезу мовлення. Проведено аналіз сучасних архітектур нейронних мереж, які є основою для створення адаптивних рішень. Представлено приклади використання систем із застосуванням часово-частотного представлення сигналів, що дозволяє досягти високої точності



вилучення інформативних ознак. Розглянуто перспективи використання адаптивних систем для вирішення прикладних задач, таких як синтез персоналізованого голосу, автоматичне розпізнавання мовлення та аудіообробка в реальному часі. Особливий акцент зроблено на значенні цих систем для інноваційних сфер, таких як голосові асистенти, медичні пристрої, системи безпеки та мультимедійні додатки.

Третій розділ роботи містить результати проведений експериментальних випробувань з перевірки якості функціонування розроблених алгоритмів для завдань вилучення ознак, синтезу мовлення та адаптації аудіосигналів у різних акустичних умовах. Зокрема, розглянуто методи зменшення впливу шумів та реверберації, а також забезпечення персоналізації голосу. Описано проведення експериментів на різних наборах аудіоданих, що у свою чергу дозволило оцінити стабільність і точність запропонованих підходів. Висвітлено практичну цінність отриманих результатів у реальних сценаріях, таких як створення персоналізованого медіаконтенту, ідентифікація за голосом та обробка аудіо у складних умовах.

Дисертаційна робота оформлена відповідно до вимог наказу МОН України від 12 січня 2017 р. № 40 «Про затвердження вимог до оформлення дисертації».

#### **Оприлюднення результатів дисертаційної роботи.**

Наукові результати дисертації висвітлені у 4 наукових публікаціях здобувача, серед яких: 4 статті у наукових виданнях, включених на дату опублікування до переліку наукових фахових видань України. Також результати дисертації були апробовані на 1 міжнародній науково-практичній конференції.

Таким чином, наукові результати описані в дисертаційній роботі повністю висвітлені у наукових публікаціях здобувача.

#### **Недоліки та зауваження до дисертаційної роботи.**

1. В другому розділі роботи доцільно було б окремо зазначити за яким принципом використовується MEL-шкала для врахування особливостей сприйняття слуху людиною і як розраховуються MEL-кепстральні коефіцієнти при цьому.

2. В практичній частині роботи (стор.60) при навчанні системи ідентифікації за голосом краще було б зазначити окремо за якими ознаками формувалась вибірка аудіо сигналів.

3. Не зовсім зрозуміло зі змісту роботи, виходячи з наведених результатів у п.3.3.3-3.3.6 за яких умов сформована експертна група оцінювала якість сигналів.

Вважаю, що висловлені зауваження не є визначальними і не зменшують загальну наукову новизну та практичну значимість результатів та не впливають на позитивну оцінку дисертаційної роботи.

### **Висновок про дисертаційну роботу.**

Вважаю, що дисертаційна робота здобувача ступеня доктора філософії Борисова Гліба Олександровича на тему «Адаптивні системи оброблення акустичної інформації для створення персоналізованого медіаконтенту» виконана на високому науковому рівні, не порушує принципів академічної доброчесності та є закінченим науковим дослідженням, сукупність теоретичних та практичних результатів якого розв'язує наукове завдання, що має істотне значення для галузі знань «Електроніка та телекомунікації». Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в п.6 – 9 «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44.

Здобувач Борисов Гліб Олександрович заслуговує на присудження ступеня доктора філософії в галузі знань 17 «Електроніка та телекомунікації» за спеціальністю 171 «Електроніка».

### **Рецензент:**

Доцент кафедри акустичних та  
мультимедійних електронних систем  
КГП ім. Ігоря Сікорського



«    »      20     року

