

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.

Оброблення акустичної інформації ставить за мету виконання різних завдань – від безпосереднього покращення звукового потоку знешумлення або дереверберації яке широко використовується в слухових апаратах, до виявлення аномалій в аудіоданих, для випадків функціонування різних машин та механізмів у промисловості. Для вирішення цих або подібних завдань, які пов'язані з обробленням мовних сигналів, останнім часом все частіше використовують нейромереві алгоритми. Такий підхід вирішує відразу декілька актуальних проблем, пов'язаних з ідентифікацією мовлення та синтезу голосу. Так, на сьогодні технології розпізнавання мовлення інтегровані у смарт-пристрої, є базовими в системах "розумний будинок" та використовуються для керування транспортними засобами. Окрім цього, технології ідентифікації голосу широко застосовуються в системах безпеки, що займаються аутентифікацією та верифікацією користувачів при доступі до захищених персональних даних. Більшість систем голосового доступу базуються на унікальних біометричних характеристиках користувача, зокрема голосу. В даному випадку якраз застосування нейронних мереж дозволяє забезпечити адаптивність, гнучкість та підтримку діалектів української мови та словосполучень при функціонуванні цих систем.

Оцінка обґрунтованості наукових результатів дисертації, їх достовірності та новизни.

Достовірність побудови акустичної моделі розпізнавання мовлення підтверджується реалізацією модифікованої рекурентної нейронної мережі, яка здатна працювати з українськими словосполученнями як вхідними даними. Обґрунтованість наукового підходу забезпечено архітектурним рішенням із вбудованою пам'яттю, що дозволяє поєднувати процеси навчання і тестування, підвищуючи ефективність роботи моделі. Запропонована модель є новою з точки зору її адаптації до особливостей української мови та одночасної оптимізації навчального процесу.

Результати, що стосуються синтезу мовлення (клонування голосу), є новими завдяки впровадженню концепції трьох послідовно з'єднаних попередньо навчених нейронних мереж. Даний підхід забезпечив високу достовірність відтворення індивідуальних мовних характеристик (інтонацій, акцентів, фонетичних нюансів) як для англійської, так і для української мов, що свідчить про його універсальність і наукову обґрунтованість. Результати

підтверджують подальший розвиток систем синтезу мовлення з адаптацією до конкретного мовця.

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 виконано поставлене наукове завдання щодо розроблення адаптивних систем оброблення аудіо інформації з підтримкою українського лінгвістичного апарату.

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 третього рівня вищої освіти «Електроніка».

Дисертація складається з вступу, трьох розділів, висновків, списку літератури та трьох додатків загальним об'ємом 137 сторінок.

У вступі міститься актуальність дослідження, її мета та основні завдання. Зазначено предмет та об'єкт дослідження, сформульовано наукову новизну та практичне значення отриманих результатів.

У першому розділі наведено приклади використання часово-частотного представлення сигналів (спектрограм, мел-спектрограм) для вилучення інформативних ознак з аудіоданих. Також визначено актуальні підходи до адаптації моделей до оброблення сигналів у складних акустичних умовах. Розглянуто методи оцінювання ефективності нейромережових моделей, а також перспективи їхнього використання для персоналізації голосу в різних прикладних задачах. Визначено основну мету клонування голосу, яка полягає у тому, щоб максимально зберегти природність і унікальні особливості оригінального голосу.

Другий розділ роботи містить інформацію про способи реалізації систем ідентифікації особи за голосовими біометричними даними, розпізнавання мови, підвищення розбірливості та клонування голосу з включенням нейронних мереж різних видів. При цьому нейронна мережа, може бути як складова частина, як от для системи ідентифікації за голосом. А може бути реалізація нейронної мережі повністю для вирішення задачі підвищення розбірливості мови. Комбінацію декількох нейронних мереж запропоновано для створення системи клонування голосу.

Третій розділ роботи містить результати перевірки за якістю розроблених автором систем, що ототожнюють різні алгоритми оброблення акустичної інформації. В даному випадку це система ідентифікації за голосом, система розпізнавання мови, система підвищення розбірливості мови з функцією дереверберації та система клонування голосу. Відмічено, що усі розроблені системи мають працювати з аудіо контентом, який створено або записано не лише англійською мовою, але й українською мовою, враховуючи місцевий

діалект. Крім цього, визначено завдання при розробленні цих систем щодо досягнення переваг їх у порівнянні з існуючими аналогами, насамперед з точки зору досягнення швидкодії навчання та роботи, точності та головне, розроблені системи мають вирізнятись властивістю адаптації роботи до зміни умов функціонування.

Дисертаційна робота є завершеною науковою працею та свідчить про особистий внесок здобувача у науковий напрям розроблення електронних засобів оброблення мовної інформації з використанням інструментарію нейронних мереж.

Розглянувши звіт подібності за результатами перевірки дисертаційної роботи на текстові співпадиння, 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 електронних засобів прикладної акустики.

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. Останній абзац п.1.4 роботи (стор.30), де визначено доцільність застосування нейронних мереж для роботи з мовними сигналами логічніше було б розмістити у вступній частині дисертації. Це буде краще, адже саме нейронні мережі в дисертаційній роботі є ключовим інструментом для вирішення поставлених завдань з оброблення акустичних даних.

2. Варто було б пояснити на основі чого при формуванні вихідних даних до експерименту при аналізі системи покращення розбірливості мови відстань від джерела сигналу до мікрофону прийнята рівною 150 см.

3. Після рис. 2.5 на стор. 54 автор використовує посилання на технологію SV2TTS, проте про особливості цієї технології для вирішення задач дослідження нічого не наводить.

4. В роботі за текстом зустрічаються орфографічні помилки, наприклад на стор. 81, 86, 89 та стор. 111 неправильно написано слово “група”.

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:

В.о. завідувача кафедри акустичних та
мультимедійних електронних систем
КПІ ім. Ігоря Сікорського
К.Т.Н., доцент



Олександр ДРОЗДЕНКО
ЗАСВІДЧУЮ
Відділ кадрів та вченої справи

Підпис _____



« 19 » 06 2025 року