

## **ВІДГУК**

офіційного опонента 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 може переслідувати декілька причин. Наприклад, коли необхідно розробити систему клонування голосу для запису аудіовізуальних творів, які потрібно щоб декларували відомі легендарні постаті в житті країни. Крім цього, інформаційні українські навчальні системи, їх аудіо частина найчастіше пов'язана з обробленням мовної інформації. Ще однією причиною може бути нагальна необхідність створення електронних додатків з підтримкою голосових помічників. Зручним і одним з можливих засобів для створення систем оброблення акустичної інформації можуть слугувати нейронні мережі різної архітектури. Враховуючи їх особливості та принцип відкритості архітектури, з'являється можливість створити актуальні нові електронні системи, які підвищують розбірливість української мови, забезпечують ідентифікацію за голосом або спрямовані на клонування голосових зразків.

### **Оцінка обгрунтованості 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



У вступі наведено актуальність роботи, мета та завдання дослідження, сформульовано об'єкт і предмет дослідження. Також визначено методи дослідження, наукову новизну та особистий внесок автора дисертації.

У першому розділі проаналізовано існуючі інноваційні підходи до оброблення аудіо інформації. І наголошено, що одним з перспективних підходів для оброблення аудіоданих може стати використання нейронних мереж, завдяки чому для користувача стає доступним простий і ефективний інструментарій з налаштування і адаптації цих мереж у випадку, коли результат оброблення сигналу потребує забезпечення певної якості відтворення, наприклад при обробленні мовних сигналів зі змістовним наповненням. Додатково наведено теоретичні основи кепстрального аналізу сигналів та визначено положення щодо сутності проведення дереверберації сигналів. Окремо, проаналізовано різні типи нейронних мереж з позиції їх використання саме для оброблення мовної інформації.

Другий розділ дисертації розкриває особливості побудови різних систем оброблення акустичної інформації з застосуванням можливостей нейронних мереж. Так, при створенні системи ідентифікації за голосом автор пропонує використовувати нейронні мережі рекурентного типу. Система підвищення розбірливості мови характеризується наявністю згорткової нейронної мережі. Що стосується системи клонування голосу, то задля збереження лінгвістичних та фонетичних характеристик мовлення запропоновано використовувати нейронні мережі глибокого навчання.

Третій розділ дослідження містить результати проведеного практичного експерименту. Зокрема, для системи ідентифікації за голосом автором використано рекурентна нейронна мережа, а не згорткова. Такий підхід дозволяє враховувати часові залежності, послідовність, глобальний контекст і дозволяє працювати зі змінною довжиною сигналів. Для цієї мережі проведено її навчання та перевірка на основі тестової вибірки з 23 людей. Нейронна мережа для системи покращення розбірливості мовлення, а точніше її перевірка відбувається на основі залучення експертної різновікової групи з 10 людей. При цьому запис сигналів проведено як в житловій кімнаті, так і в різних навчальних аудиторіях 12 корпусу університету. Додатково проведено експерименти з мовними записами українською мовою з додаванням стороннього шуму. Для розробленої системи клонування голосу автором проаналізовано MEL-спектрограми сигналів, які записано як англійською так і українською мовами.



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. На сторінці 26 першого розділу роботи при опису функції розбірливості мовлення (*SI*) краще уникати використання жаргонних термінів, таких як “підігнати”. В даному випадку наведене речення можна побудувати таким чином: “параметри функції *SI* можна визначити такими чином, щоб її значення в окремих крапках були б наближені до відповідних значень, отриманих емпіричним шляхом”.
2. Другий абзац дисертації на сторінці 30, де визначено ідеї застосування нейронних мереж для оброблення мовних сигналів, доцільно розмістити у вступі до дисертації.
3. У п.2.2.2 роботи бажано було б додатково навести, яким чином застосовується на практиці функція активації Leaky ReLU (стор.52).
4. На сторінці 61 роботи було б краще відмітити переваги використання оптимізатора Адама перед іншими видами оптимізаторів для досягнення цілей навчання розробленої нейронної мережі.

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



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.

Отже, в дисертаційній роботі поставлене наукове завдання 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 «Електроніка».

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 викладена на високому науково-технічному рівні з використанням стандартної термінології, що прийнята в галузі електроніки та прикладної акустики.

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

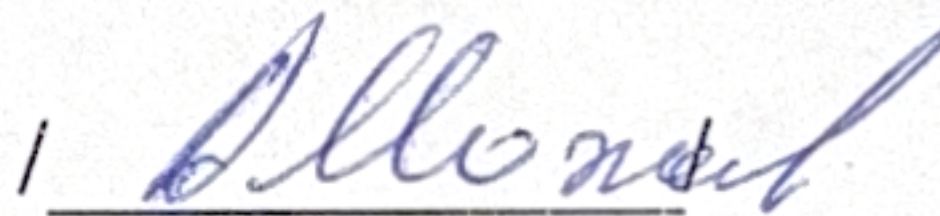


значення для галузі знань «Електроніка та телекомунікації». Дисертаційна робота за актуальністю, практичною цінністю та науковою новизною повністю відповідає вимогам чинного законодавства України, що передбачені в п.6 – 9 «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої 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».

**Офіційний опонент:**

Професор кафедри кібербезпеки  
та DATA-технологій  
Навчально –наукового інституту №5  
Харківський національний університет  
внутрішніх справ  
д.т.н., професор

/ 

Олександр МОЖАЄВ

«\_\_\_» \_\_\_\_\_ 20\_\_ року

*Згідно з доповіддю  
Директора інституту*



Оксана БРУСАКОВА