

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДОНЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ВАСИЛЯ СТУСА

ЗВІЗДА МИКОЛА РУСЛАНОВИЧ

Допускається до захисту:

В.о. завідувача кафедри біофізики і фізіології
кандидат хім. наук, доцент _____ О. І. Доценко
« ____ » _____ 2025 р.

**РОЗРОБКА ТА ВИПРОБУВАННЯ НЕ ДОРОГОГО ПРИСТРОЮ ДЛЯ
ПАСИВНОГО АКУСТИЧНОГО МОНІТОРИНГУ БІОРІЗНОМАНІТТЯ
ТА ДОВКІЛЛЯ**

Спеціальність 091 Біологія та біохімія

Кваліфікаційна (магістерська) робота

Керівник: Міщенко А.М., к.б.н.,
старший викладач кафедри біофізики та фізіології

Оцінка: ____ / ____ / ____

(бал/за шкалою ЄКТС/за національною шкалою)

Голова ЕК: _____

Вінниця 2025

Анотація

Звізда М.Р. Розробка та випробування не дорогого пристрою для пасивного акустичного моніторингу біорізноманіття та довкілля. Кваліфікаційна робота магістра. Спеціальність 091 Біологія та біохімія. Донецький національний університет імені Василя Стуса, Вінниця, 2025.

У роботі представлено розробку та випробування недорогого пристрою для пасивного акустичного моніторингу (РАМ) біорізноманіття з акцентом на виявлення та ідентифікацію видів птахів. Дослідження базується на створенні автономного аудіозаписувального модуля на основі мікропроцесора Raspberry Pi або ESP32, який забезпечує довготривале збирання екологічних акустичних даних у польових умовах. Актуальність роботи зумовлена зростаючою потребою у доступних, масштабованих і технічно гнучких інструментах моніторингу біорізноманіття, оскільки традиційні методи, засновані на ручних обстеженнях, мають низку суттєвих обмежень — від суб'єктивності до високих трудових витрат.

У межах дослідження реалізовано апаратну та програмну конфігурацію пристрою, включно з автономним записом аудіо та можливістю локального аналізу даних за допомогою нейронної мережі BirdNET (для варіанту з Raspberry Pi). Проведено оцінку точності автоматичного розпізнавання видів птахів, визначення оптимальних порогових значень довіри для різних умов запису та аналіз впливу цих порогів на обсяг і якість виявлень. Результати дослідження демонструють потенціал недорогих модульних систем РАМ для підвищення ефективності польових екологічних досліджень, розширення можливостей громадської науки та створення відкритих технологічних рішень у сфері екоакустики.

Ключові слова: пасивний акустичний моніторинг, BirdNET, Raspberry Pi, ESP32, біорізноманіття, екоакустика, птахи, автоматичне розпізнавання.

54 с., 3 табл., 23 рис., 19 джерел.

ABSTRACT

Zvizda M.R. The development and evaluation of a low-cost device for passive acoustic monitoring biodiversity and the environment. Master's qualification work. Specialty 091 Biology and Biochemistry. Vasyl Stus Donetsk National University, Vinnytsia, 2025.

This study presents the development and testing of a low-cost device for passive acoustic monitoring (PAM) of biodiversity, with a particular focus on detecting and identifying bird species. The research is based on the creation of an autonomous audio-recording module using a Raspberry Pi or ESP32 microcontroller, enabling long-term collection of ecological acoustic data in field conditions. The relevance of the work is driven by the growing need for accessible, scalable, and technically flexible biodiversity monitoring tools, as traditional survey-based methods remain limited by subjectivity, high labor requirements, and restricted temporal coverage.

The project includes the hardware and software implementation of the device, featuring autonomous audio recording and the possibility of on-board data analysis using the BirdNET neural network (for the Raspberry Pi configuration). The study evaluates the accuracy of automatic bird species identification, determines optimal confidence thresholds for various recording environments, and analyzes the influence of these thresholds on the volume and quality of detections. The results demonstrate the potential of low-cost modular PAM systems to enhance the effectiveness of field ecological research, support citizen science initiatives, and contribute to the development of open technological solutions in ecoacoustics.

Keywords: passive acoustic monitoring, BirdNET, Raspberry Pi, ESP32, biodiversity, ecoacoustics, birds, automatic identification.

ЗМІСТ

ВСТУП	5
РОЗДІЛ 1. ЛІТЕРАТУРНИЙ ОГЛЯД	8
1.1 Характеристики мікрофонів.....	8
1.2 Вибір відповідних моделей мікрофонів.....	9
1.3 Літературний огляд	10
РОЗДІЛ 2. МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕННЯ	24
2.1 Методологія тестування частотної характеристики мікрофонів	24
2.2 пасивна акустична реєстрація звуків птахів та кажанів в польових умовах	24
2.3 Скрипт для розрізання аудіо запису на фрагменти на основі файлу анотації	25
2.4 Метрики для оцінки точності нейромережі BirdNet	27
2.5 Довірчий бал BirdNet, оцінка мінімального значення довірчого балу	28
РОЗДІЛ 3. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТУ ТА ЇХ ОБГОВОРЕННЯ	33
3.1 Блок схема автономного записувального пристрою (АЗП).....	33
3.1.1 Процесорний блок.....	34
3.1.2 Версія на основі Raspberry Pi.....	34
3.1.3 Версія на основі ESP32	35
3.2 Вибір моделі мікрофона	36
3.2.1 Висновок і рекомендації.....	41
3.3 Тестування придбаних зразків мікрофонів	42
3.3.1 Некалібрована частотна характеристика	42
3.3.2 Калібровані частотні характеристики	43
3.3.3 Тест на чутливість до ультразвуку (40 кГц).....	45
3.3.4 Висновок	48
3.4 Розпізнавання видів птахів за допомогою BirdNet.....	48
3.5 Оцінка точності нейромережі BirdNet	52
3.6 Вибір граничних значень довірчої оцінки BirdNet.....	56
ВИСНОВКИ.....	60
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	62

ВСТУП

Біорізноманіття є фундаментальним компонентом екологічної стабільності та функціонування природних екосистем. Її зміни слугують чутливими індикаторами антропогенного впливу, деградації природних середовищ і глобальних екологічних загроз. У контексті зростаючої економічної активності, фрагментації ландшафтів та прискорення зміни клімату, систематичний моніторинг біорізноманіття стає критично важливим для своєчасної оцінки стану екосистем і формування ефективних політик захисту довкілля. Птахи є особливо важливою групою індикаторних організмів, які оперативно реагують на зміни навколишнього середовища і широко використовуються в наукових програмах моніторингу. Традиційні методи моніторингу птахової фауни, засновані на ручних візуальних та акустичних обстеженнях, мають низку суттєвих обмежень. Вони залежать від кваліфікації спостерігачів, охоплюють обмежений період часу, потребують значних людських ресурсів і часто характеризуються суб'єктивними оцінками. Крім того, такі методи важко масштабувати на великі території або забезпечити одночасне покриття багатьох населених пунктів. Це вимагає переходу до технологічно розвинених автоматизованих систем запису та аналізу даних. Однією з найпрогресивніших технологій є пасивний акустичний моніторинг (PAM) — метод, заснований на довготривалому автономному записі екологічної звукової активності. PAM дозволяє отримувати великі набори даних про вокальну активність птахів у різний час доби, незалежно від присутності спостерігача, погодних умов чи доступності місцевості. Ці технології отримали значний розвиток завдяки методам автоматичного аналізу акустичних даних, зокрема машинному та глибокому навчанню. Одним із найефективніших інструментів сьогодні є нейронна мережа BirdNET, здатна автоматично розпізнавати види птахів за короткими аудіофрагментами. Це відкриває можливість швидкого та масштабного аналізу великих аудіоколекцій, суттєво знижуючи трудомісткість традиційної експертної обробки.

Незважаючи на наявність комерційних акустичних рекордерів, їхня висока вартість, закриті обладнання, обмежені можливості модифікації та складність інтеграції з додатковими програмними інструментами суттєво знижують їхню придатність для наукових досліджень і освітніх проєктів. У цьому контексті розробка власного недорогого пристрою для РАМ на базі мікропроцесора Raspberry Pi або ESP32 є актуальною та перспективною. Такі рішення поєднують низьку вартість, гнучкість конфігурації, відкритість програмного забезпечення та можливість інтегрувати алгоритми аналізу даних безпосередньо «на борту» пристрою. У випадку Raspberry Pi стає можливим не лише автономний аудіозапис, а й локальний аналіз даних із використанням нейронної мережі BirdNET, що значно підвищує швидкість обробки та ефективність польових досліджень. Одним із важливих завдань сучасної екоакустики є оцінка точності автоматичного виявлення та визначення оптимальних порогових значень оцінки довіри для конкретних дослідницьких умов. Оскільки характеристики шумового середовища, видова структура птахової фауни та акустичний ландшафт можуть суттєво відрізнятися між місцями, універсальні порогові значення не існують. Тому для правильної інтерпретації результатів необхідно переоцінити точність виявлення BirdNET на отриманих польових даних, визначити оптимальні порогові значення та проаналізувати їхній вплив на об'єм і якість виявлень. Отже, розробка та тестування доступного акустичного пристрою для РАМ, а також аналіз ефективності автоматичної ідентифікації видів птахів за допомогою BirdNET є важливим кроком до вдосконалення інструментарію моніторингу біорізноманіття. Це сприяє підвищенню точності оцінок, розширенню можливостей польових досліджень і створенню відкритих технологічних рішень, доступних для науковців, природоохоронців і громадських ініціатив.

Мета: Метою даної роботи є розробка, створення та експериментальне випробування недорогого пристрою для пасивного акустичного моніторингу біорізноманіття, побудованого на базі Raspberry Pi або мікропроцесора ESP32, а також оцінка ефективності автоматичної детекції видів птахів за записаними

аудіоданими з використанням нейромережі BirdNET, включно з визначенням точності класифікації та оптимальних порогових значень довірчої оцінки (confidence score) для конкретних умов дослідження.

Задачі. В рамках окресленої мети виконати наступні задачі:

1. Виконати літературний огляд що стосується рекомендацій по вибору мікрофона для аудіомоніторингу дикої природи. Здійснити огляд технічної документації моделей мікрофонів доступних на ринку.
2. Виміряти амплітудно-частотні характеристики мікрофонів доступних на ринку, оцінити можливість їх використання для запису птахів та кажанів.
3. Спланувати та провести реєстрацію звуків птахів та кажанів в польових умовах.
4. Для отриманих аудіоданих провести детекцію видів птахів нейромережею BirdNet.
5. Для обраних видів птахів оцінити точність детекції та граничне значення довірчої оцінки (confidence score).

РОЗДІЛ 1. ЛІТЕРАТУРНИЙ ОГЛЯД

Мікрофон є критично важливим компонентом автономного блоку запису (ARU), відповідальним за захоплення акустичних сигналів від дикої природи, включно з вокалізаціями птахів (100–12 кГц) та ехолокацією кажанів (до 150 кГц). Вибір відповідного мікрофона передбачає оцінку ключових характеристик, які відповідають критеріям проєкту: енергоефективність, пропускна здатність, низький рівень шуму, низька вартість і компактний розмір. Додаткові аспекти включають частотну характеристику, динамічний діапазон, чутливість та екологічну стійкість, які безпосередньо впливають на ефективність ARU у пасивному акустичному моніторингу (PAM) для різноманітних екологічних застосувань

1.1 Характеристики мікрофонів

Наступні характеристики мікрофона є критично важливими для конструкції ARU:

Співвідношення сигнал/шум (SNR): Високий SNR є необхідним для виявлення слабких звуків дикої природи, таких як далекі пташині крики або ехолокація кажанів, на фоновому шумі. Як показано у (K. F. A. Darras et al., 2020), вищий SNR значно збільшує дальність виявлення (наприклад, від 207 до 10 477 м² для чутних звуків) і покращує автоматизовану функцію виявлення як для пташиних, так і для кажанів. Для високоякісних записів переважають $SNR \geq 70$ дБ при 1 кГц [3].

Смуга пропускання та частотна характеристика: Мікрофон повинен охоплювати широкий діапазон частот (від 20 Гц до >100 кГц), щоб захоплювати як низькочастотні співи птахів, так і ультразвукові крики кажанів. Плоска частотна характеристика забезпечує точне відображення звукового ландшафту, тоді як розширена ультразвукова реакція критично важлива для моніторингу кажанів.

Динамічний діапазон: Широкий динамічний діапазон дозволяє мікрофону сприймати як тихі, так і гучні звуки (наприклад, тихі пташині крики до гучних навколишніх шумів) без спотворень. Бажаний динамічний діапазон ≥ 60 дБ, як показано у (Mydlarz et al., 2014), щоб враховувати різноманітні рівні звуку в природних умовах [9].

Чутливість: Вища чутливість (наприклад, -30 дБ або вище) покращує захоплення сигналів низької амплітуди — це критично важливо для віддалених або слабких вокалізації дикої природи. Однак надчутливі мікрофони можуть перевантажуватися гучними звуками, що вимагає ретельного вибору залежно від випадку.

Екологічна стійкість: мікрофон повинен витримувати зовнішні умови, включно з коливаннями температури (від -20°C до 50°C), вологістю та дощем. MEMS-мікрофони часто мають кращу стійкість порівняно з електретами, як зазначено у (Mydlarz et al., 2014), завдяки ширшому робочому температурному діапазону та імунності до RFI/EMI.

Споживання енергії: Низьке споживання енергії критично важливе для акумуляторних ARU. MEMS-мікрофони зазвичай споживають менше енергії (наприклад, 120 мкА для SPU0410LR5H-QB), ніж електрети, що відповідає критеріям енергоефективності.

Розмір і вартість: компактні мікрофони полегшують інтеграцію в портативні ARU, тоді як недорогі варіанти ($\leq \$10$) підтримують масштабовані розгортання для масштабних РАМ-досліджень.

1.2 Вибір відповідних моделей мікрофонів

Вибір мікрофонів для ARU базувався на попередньому досвіді моніторингу звуку дикої природи, задокументованому в літературі про некомерційні конструкції ARU, індивідуальні мікрофонні системи (наприклад, Sonitor (K. Darras et al., 2021)) та дослідження оцінки продуктивності мікрофона в РАМ, особливо для птахів і кажанів (K. F. A. Darras et al., 2020) [3, 4]. Технічні листи мікрофонів та рекомендації польових експертів (Влад,

2022; Рекомендації для блокфлейти Зака – Зак Пофф, без дати) були перевірені для забезпечення сумісності з критеріями проекту.

Вибір мікрофона здійснювався:

1. Література про некомерційні ARU та проектування мікрофонів для екологічного запису
2. Рецензовані дослідження, що оцінюють SNR, дальність виявлення та довговічність птахів і кажанів
3. Технічні листи виробника
4. Практичний досвід і спільноти мікрофонів для самостійного використання
5. Доступність на українському ринку

1.3 Літературний огляд

Як співвідношення сигнал/шум (SNR) мікрофонів впливає на якість акустичного моніторингу дикої природи

У статті (K. F. A. Darras et al., 2020) досліджується, як співвідношення сигнал/шум (SNR) мікрофонів впливає на якість акустичного моніторингу дикої природи, зосереджуючись на птахів і кажанах. Автори виявили, що мікрофони з вищим SNR значно покращують якість акустичного дискретизації, що є критично важливим для екологів, які вивчають сонантне біорізноманіття [3]. Дослідження включало тестування 12 різних моделей мікрофонів від шести виробників. Серед них були як традиційні циліндричні капсули, такі як WM-61A (Panasonic), AOM-5024L-HD-R (PUI Audio) та EM258 (Primo), так і мікроелектромеханічні системи (MEMS) — SPU0410LR5H-QB (Knowles) та ICS-40720 (Invensense).

Ключові кількісні результати виявили такі впливи вищого мікрофона SNR:

Простори для виявлення звуку: Мікрофони з вищим SNR призвели до значного збільшення простору для виявлення звуку. Для чутного звуку (1 кГц)

ці площі значно збільшилися з 207 до 10 477 м², що означає збільшення у 1,7 рази від найнижчого до найвищого мікрофона SNR. Для ультразвуку (40 кГц) площа виявлення збільшилася з 353 до 3632 м², що свідчило про ще більше збільшення у 10 разів. Ці зростання були статистично значущими ().

Активність вокалізації: Зафіксована вокалізація дикої природи також значно зросла при підвищенні SNR. У птахів активність вокалізації зросла у 1,6 раза (від 950 до 1539 секунд). Для бит це зростання було ще більш помітним — у 9,7 раза (від 33 до 320 секунд). Варто зазначити, що модель мікрофона з найнижчим SNR (POM-1345P-C3310-R) не зафіксувала жодних пташиних криків.

Різноманіття видів: Пул видів як птахів, так і кажанів не міг бути повністю виявлений мікрофонами з нижчим SNR. Вищий мікрофонний SNR зазвичай призводив до вищого максимального рівня багатства видів кажанів і швидше. У птахів спостерігалися вищі максимальні рівні насиченості вибірки при збільшенні SNR в один із двох днів відбору.

Автоматизоване виявлення: Продуктивність автоматичного виявлення як пташиного, так і кажанського крику позитивно впливала через вищий мікрофонний SNR. Діапазон порогових значень балів, які мали прийнятну пригадку та точність (визначені як мінімум 0,5), значно збільшився на 0,9 і 0,4 на кожну одиницю SNR для птахів і кажанів відповідно. Записи з мікрофонів з низьким SNR виявилися неприйнятними з точки зору точності та запам'яті.

Дослідження підкреслює, що мікрофонний SNR — це ключова характеристика, яка суттєво впливає на акустичну дискретизацію птахів і кажанів, і її слід максимізувати шляхом вибору відповідних мікрофонів. Також було відзначено, що вимірювані значення SNR показують значні розбіжності між виробниками мікрофонних елементів; наприклад, PUI Audio SNR у середньому були на 26 дБ нижчими, ніж для мікрофонних елементів Knowles, що свідчить про те, що характеристики виробника не завжди є послідовними або порівнянними. Це підкреслює критичну важливість вибору мікрофонів з

високим SNR для досягнення максимальних діапазонів виявлення та чіткіших сигналів, що полегшує як людський, так і автоматизований аналіз записів.

AOM-5024L-HD-R є взаємозамінним із елементом EM172, який використовується в системі запису з відкритим кодом Solo (Whytock & Christie, 2017)

SPU0410LR5H-QB у проекті міського звукового сенсорування

У статті (Mydlarz et al., 2014) представлено початкове дослідження, яке порівнює акустичну продуктивність спеціального MEMS-мікрофона, зокрема Knowles SPU0410LR5H-QB, з трьома не-MEMS споживчими мікрофонами: електретним Panasonic WM-64PC, електретним мікрофоном Polson OLM-10 на лацкані та конденсаторним USB-мікрофоном Blue Snowflake [9]. Це порівняння було проведено в контексті вибору найбільш відповідної мікрофонної технології для масштабного, довгострокового та безперервного міського проекту звукового сенсування в Нью-Йорку. Проект має на меті впровадження численних пристроїв дистанційного зондування (RSD) для точного моніторингу складного та динамічного звукового середовища міста, що вимагає надійного, стійкого, високоякісного, енергоспоживання та недорогого рішення, яке можна масштабувати. Результати показали, що мікрофон MEMS значно перевершив конкурентів за ключовими акустичними характеристиками, актуальними для міського звукового сенсування.

- Частотна характеристика: Індивідуальний мікрофон MEMS мав найближчу до плоскої частотної характеристики в широкому діапазоні, що є критично важливим для точної трансдукції та представлення різноманітного міського звукового ландшафту. Для порівняння, електрети Panasonic і Polson мали плоскі відгуки лише на нижчих частотах, з чіткими високочастотними піками та падіннями, тоді як Blue Snowflake демонстрував нерегулярну частотну характеристику з численними великими жолобами, що робило його менш придатним.

- Динамічний діапазон: Мікрофон MEMS демонстрував ширший динамічний діапазон у 62,4 дБ, порівняно з 60,2 дБ у Panasonic, 54,2 дБ у Polson

і 49,6 дБ у Blue, що насамперед пояснюється нижчим рівнем шуму (44 дБ). Хоча більшість мікрофонів, включно з MEMS (106,9 дБ), демонстрували порівнянно високий максимальний звуковий тиск (SPL), ширший динамічний діапазон MEMS дозволяє точно захоплювати як низькорівневі сигнали, так і високорівневі імпульсивні джерела, поширені в міських умовах.

- Екологічна стійкість: MEMS-мікрофони також користуються перевагою завдяки своїй винятковій стійкості до змінних умов навколишнього середовища, включаючи ширший робочий температурний діапазон (від -40 до 100 °C) порівняно з електретними мікрофонами, а також кращу стійкість до радіочастотних (RFI) та електромагнітних перешкод (EMI).

- Мініатюризація та послідовність: Їхній надзвичайно малий розмір (зазвичай у 10 разів менший за електрети) дозволяє інтегрувати додаткові схеми, такі як підсилювачі та аналогово-цифрові перетворювачі, в автономний блок, зменшуючи загальну площу та енергоспоживання.

Крім того, MEMS-мікрофони забезпечують надзвичайно високий рівень послідовності акустичних характеристик між деталями, що робить їх дуже придатними для багатокапсульних і багатосенсорних масивів. Крім того, прототип MEMS-масиву, що складається з мікрофонів Knowles SPU0410LR5H-QB, був протестований на спрямованість. Цей масив мав більш спрямовані характеристики на вищих частотах (понад 2 кГц) і збільшені відстані, що свідчить про її потенційну корисність для локалізації джерел звуку (наприклад, сирени, спів птахів) та застосування на далеких полях у міському моніторингу. На основі цих комплексних оцінок мікрофон MEMS пропонується для використання у майбутньому розробленні сенсорів проекту.

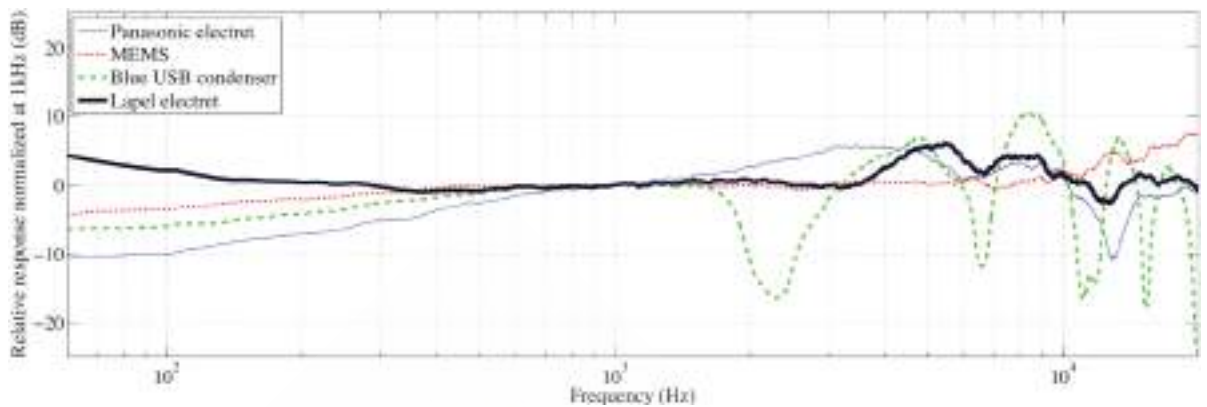


Рис.1.1- Порівняння частотної характеристики мікрофона [9]

Збирання індивідуальних мікрофонів для запису наземної дикої природи: система Sonitor

У статті (К. Даррас та ін., 2021) представлено Sonitor — відкриту, високопродуктивну та модульну мікрофонну систему, розроблену для задоволення потреб екологів у акустичному моніторингу наземної дикої природи. Головна мета — надати економічно вигідні альтернативи дорогим комерційним мікрофонам, які часто мають неповні технічні характеристики і швидко псуються під впливом зовнішніх умов. Система Sonitor має на меті охопити весь спектр звукових частот сонантної наземної дикої природи за частку вартості комерційних рішень [3, 4].

Система переважно використовує два основні типи мікрофонних елементів: електроконденсаторні мікрофони (ECM) та мікроелектромеханічні системи (MEMS). MEMS-мікрофони здебільшого обирали за високу продуктивність у малих розмірах, потенціал більшого приросту продуктивності порівняно з традиційними капсулами ECM, а також за меншу варіативність від частини до частини та чутливість до змін температури. Однак дослідження також визнає переваги ECM, зокрема їхню вищу чутливість і співвідношення сигнал/шум при 1 кГц. Некомерційні мікрофонні елементи, такі як Knowles SPU0410LR5H-QB, успішно використовуються компаніями, такими як Biotope. Їх як заміників комерційних мікрофонів і у власних корпусах авторів з 2017 року для запису птахів і кажанів за допомогою

рекордерів SM2Bat+. Рекордер AudioMoth, пристрій з відкритим кодом, має вбудований мікрофонний елемент MEMS, здатний відбирати звук до 192 кГц.

Процес проєктування включав оцінку кількох вроджених компромісів у конструкції мікрофона для використання на відкритому повітрі. До них належать:

- Захист від погодних умов проти заглушення звуку: Захист мікрофонів від проникнення води за допомогою акустичних вентиляційних отворів може суттєво зменшити звук. Наприклад, вентиляція GAW325 при тестуванні на мікрофоні Knowles SPU0410LR5H-QB приглушувала звуки 1 кГц на -7,6 дБ спереду і -9,4 дБ збоку, а 40 кГц — на -2,5 дБ спереду. Ще важливіше, що три нещодавно протестовані моделі вентиляційних отворів показали неприйнятні рівні ультразвукового ослаблення, що перевищують 7 дБ, що може скоротити діапазон виявлення вдвічі.

- Вітроізоляція проти сушіння після дощу: хоча лобові скла зменшують шум вітру, вони можуть затримувати воду після дощу, що заважає передачі звуку. Вентиляційний отвір GAW112 у поєднанні з лобовим склом потребував значно більше часу для висихання, ніж 6 мм клаксонів. У вологому стані (1–3 години після замочування) високі чутні частоти (10 кГц) ослаблювалися приблизно на 20 дБ, а ультразвук — приблизно на 30 дБ більше, ніж у трубі. Мікрофону знадобилося до 18 годин, щоб повністю висохнути і повернутися до нормального рівня запису.

- Довжина кабелю проти втрат сигналу: Довгі кабелі можуть спричинити погіршення сигналу, особливо на високих частотах через ємність кабелю. Тести показали, що кабелі діаметром 52,5 метра знижують рівень звуку тестових сигналів на 40 кГц на 1,2–1,3 дБ порівняно з 5-метровими кабелями. Однак ця втрата не була визнана «великою», що свідчить про те, що попередні підсилювачі можуть бути непотрібними для мікрофонів із помірною чутливістю.

- Спрямованість проти підсилення: Акустичні роги підсилюють звук, концентруючи його на мікрофонному елементі, але при цьому підвищують

спрямованість, звужуючи патерн звукознімача. Ультразвуковий рупор тестував підсилений звук спереду на 9,1 дБ, з 45° на 2,5 дБ, а також ослаблений звук з 90° на 5,2 дБ. Великі акустичні рупори для чутного звуку підсилювали звуки 1 кГц на 3,4 дБ і звуки 10 кГц на 7,2 дБ спереду.

У статті пропонується п'ять конфігурацій мікрофонів, включаючи Sonitor Bufo та Sonitor Pavo, оптимізованих для конкретних випадків використання.

- Sonitor Bufo (Дешевий аудібл):

- Використовується мікрофон: використовує капсули ECM, при цьому для високого співвідношення сигнал/шум рекомендують Primo EM172 або PUI AOM-5024L-HD-R.

- Особливості конструкції: Це найпростіший і найдешевший дизайн, що складається лише з аудіороз'єму, проводів, ECM з приклеєним вентиляційним отвором GAW325 та епоксидного клею. Він не потребує друкованих плат чи металевих корпусів.

- Цільовий звук і характеристики: призначений лише для моніторингу звуків у чутному діапазоні, таких як амфібії, птахи та примати. Він має теоретичне співвідношення сигнал/шум 80 дБ при 1 кГц і вартість 8 євро за одиницю з часом складання 6 хвилин. Він несумісний з акустичними рогами.

- Sonitor Pavo (Дешевий універсал):

- Використовується мікрофон: Використовує мікрофон Vesper VM1000 MEMS.

- Конструктивні особливості: Виготовлено з мікрофоном Vesper VM1000 на друкованій платі (PCB), з приклеєним вентиляційним отвором GAW112 та епоксидним клеєм. Він не модульний, тобто різки не можна прикріпити.

- Цільовий звук і характеристики: Розроблений для повноспектральних записів, охоплюючи все — від низькочастотних викликів амфібії до високочастотних викликів кажана. Він має співвідношення сигнал/шум 62 дБ на 1 кГц і 36–54 дБ при 40 кГц. Вартість становить 12 євро

за одиницю з часом складання 6 хвилин. Vesper VM1000 стійкий до вищих екологічних навантажень завдяки своїй п'єзоелектричній конструкції.

Бінауральні мікрофони з низьким рівнем шуму (Primo EM272 та PUI 5024)

(Низькошумні бінауральні мікрофони (Primo EM272 та PUI 5024) – Зак Пофф, без дати)Протягом років я експериментував із капсулами Panasonic WM-61a, щоб виготовляти дешеві бінауральні мікрофони. Вони чудові для міських звукових ландшафтів, але надто шиплячі для запису тихих звуків. Мод Linkwitz трохи допомагає, але здебільшого впливає на частотну характеристику та обробку SPL, а не на власний шум.... Primo EM272 вже має вбудований FET, тож поводить себе так само, як старий Panasonic WM-61a (тільки не шумний). Майк веде чудовий блог із масштабними експериментами зі створення мікрофонів із звуковими кліпами, поясненнями та вимірюваннями.

- Primo EM172 був замінений на Primo 272. Той самий звук.
- PUI AOM-5024L-HD-R — це дуже схожа капсула, яку набагато легше купити (і вона дешевша).

Рекомендації для блокфлейти Зака

(Рекомендації блокфлейти Зака – Зак Пофф, без дати)Мої студенти іноді просять у мене поради щодо купівлі власних блокфлейтів і мікрофонів. Ось мої рекомендації щодо гнучкого обладнання для запису, яке підходить для серйозного звукозапису. Внизу сторінки є посилання на поради інших людей. (Будь-які посилання на торговців — лише для довідки. Я не отримую відкатів за ваші кліки.)

- Виробники мікрофонів нижче використовують ті ж кілька мікрофонних капсул. Мікрофони з однаковою капсулою звучатимуть однаково, але відрізняються різними дизайнами корпусів, стратегіями використання кабелів, парами, аксесуарами тощо.

- 10-мм всенаправлена капсула Primo EM272 має дуже низький рівень самошуму (14 дБ). Це значно менше, ніж у традиційних лав-мікрофонах, тому вони чудово підходять для тихих джерел. Він має більш «м'яку»

високочастотну характеристику, тому може знадобитися підсилювач еквалайзера вище 10 кГц (що частково нівелює перевагу шуму).

- 6-мм всенаправлена капсула Primo EM258 має вищий власний шум (22 дБ), але яскравіший характер, який не потребує еквалайзера. Звучить дуже схоже на дорогі лавантри на кшталт DPA 4060. Він також має корисну реакцію на УЗД.

- Кардіоїдні мікрофонні капсули, такі як Primo EM204 / EM423, мають низький власний шум і збалансовану частотну характеристику. Вони добре порівнюються з багатьма студійними мікрофонами.

- Більшість цих мікрофонів дуже компактні. Їх можна прикріпити до окулярів або пов'язки для створення бінауральних записів, або прикріпити до гілок дерев, штатив тощо.

- Продавець із Великої Британії (але доставка до Північної Америки та інших країн легка)

- Окремі мікрофони, відповідні стереопари та саморобні деталі (включаючи капсули для сирого мікрофона для виготовлення власних мікрофонів)

- У них є мікрофони на базі Primo EM272, Primo EM258 та інших капсул (включаючи мікрофони з ультразвуковим підходом).

Найкращі мікрофони для запису звуків природи та дикої природи

(Vlad, 2022) Lav мікрофони, такі як Clippy EM272 або Lom mikroUsi, дають відмінні записи за свій розмір і ціну, хоча й не такі якісні, як у згаданих мною мікрофонах Sennheiser [17]. Це також пов'язано з преампами, і Sony A10 ніколи не буде таким хорошим, як рекордер Sound Devices або Zoom F серії.

Динамічний діапазон частково пов'язаний із власним шумом, хоча все трохи складніше. Мікрофони МКН мають відмінний динамічний діапазон — я записував усе — від м'якого снігопаду до двигунів літака на своїх 8040. Я пробував записувати двигуни автомобілів мікрофонами Clippy та mikroUsi, і

капсули швидко перевантажувалися. Вони чудово підходять для тонких або середніх гучних звуків, але не призначені для надто гучних джерел.

Мікрофони, такі як DPA 4062 (або його новіший аналог), набагато краще підходять для запису гучних двигунів або вибухів. Ще одна річ, що стосується якості звуку — це частотна характеристика.

Мікрофони MKN80xx вловлюють контент до 40-50 кГц, хоча шум, про який я згадував раніше, стає помітним при 30 кГц.

Мікрофони MikroUsi захоплюють пристойний контент вище 30 кГц, тоді як мікрофони Clippy зазвичай не захоплюють багато контенту вище 20 кГц. Це не найважливіший аспект, але варто пам'ятати, якщо ви шукаєте HF-контент, наприклад, з високими обертонами чи кажанами.

Clippy EM172 або 272Pro:

- Гарний шум про себе
- Працює з живленням від розетки
- Низька ціна
- Малий розмір і площа
- Легко захищається від м'якого вітру
- Легше купити, ніж mikroUsi

Мінуси:

- Легко перевантажується гучними звуками
- Не переживає більше кількох ночей у тропічному лісі з Windbubbles
- Гризуни люблять їсти кабелі
- Погана частотна характеристика понад 20 кГц

Поради щодо USB-мікрофонних систем

*(Поради щодо USB-мікрофонних систем · Mcguirepr89/BirdNET-Pi · Обговорення #39, без дати)*Здається, існує хибне уявлення про різницю між всенаправленими мікрофонами (більшість Lavalier мікрофонів, електретних капсул, згаданих у цій темі тощо) і напрямними мікрофонами (кардіоїдні, суперкардіоїдні, гіперкардіоїдні, шотган-мікрофони). Дозвольте пояснити,

чому, на мою думку, немає переваги використання спрямованих мікрофонів саме для виявлення птахів за допомогою системи BirdNET на базі Raspberry-Pi: Уявіть два мікрофони, розміщені в полі: один DIRECTIONAL MIC (можливо, shotgun) і один OMNIDIRECTIONAL MIC (як згаданий вище лавальєрний мікрофон):

- У фронтальному напрямку всенаправлений мікрофон і спрямований мікрофон мають однакову ймовірність виявлення птаха (обмежені самошумом, а не чутливістю

- Зверніть увагу, що напрямний мікрофон НЕ має жодної переваги у чуванні птахів у фронтальному напрямку, оскільки не має додаткового безшумного підсилення в цьому напрямку

- Направний мікрофон має недолік у всіх інших напрямках, оскільки він спрямований, тому має АТЕНУАЦІЮ в інших напрямках, ніж фронтальний

- Всенаправлений мікрофон має однакову чутливість у всіх напрямках

- Отже, ЯКЩО радіус у фронтальному напрямку для спрямованого мікрофона становить 25 м, то для всенаправленого мікрофона він також 25 м (за рівних самошумів і чутливості), але для напрямного мікрофона він значною мірою зменшується у всіх інших напрямках

- Тепер ми також можемо розглянути співвідношення сигнал-шум (SNR)

- SNR буде кращим у напрямку мікрофона, бо «шум» (цей «шум» часто стосується інших птахів, які можуть вас зацікавити!) з боків і ззаду певною мірою відхиляється

Отже, підсумовуючи, напрямлений мікрофон НЕ більш чутливий до фронтального напрямку, ніж всенаправлений.

До речі, ми ще не говорили про шум вітру в мікрофонах-shotgun. Він значно вищий, ніж у інших мікрофонах, тому такий мікрофон потрібно накрити дирижаблем і/або мертвим котом, що робить систему набагато громіздкішою та дорожчою

Якщо запитати орнітолога (я, до речі, орнітолог), він/вона порекомендує напрямний мікрофон для польових досліджень у ситуаціях, коли ви

контролюєте напрямок мікрофона і можете активно спрямовувати мікрофон на птаха, і ваш головний інтерес — якісний і чистий запис ОДНОГО виду птахів. АЛЕ: у системі BirdNET це не так:

- У системі BirdNET мене цікавить система, яка слухає ВСІХ птахів у ВСІХ напрямках і ідентифікує їх
- Для цього потрібен всенаправлений мікрофон
- До речі: наприклад, BirdNET здатний одночасно ідентифікувати співаючих птахів у ранковому приспіві

МОЯ РЕКОМЕНДАЦІЯ щодо мікрофона в системі BirdNET:

- Використовуйте всенаправлений мікрофон з мінімальним самошумом, який можете собі дозволити
- Цього можна досягти, використовуючи електретну капсулу EM272 з дуже низьким самошумом, 14dBA (SNR 80dBA) — >дивіться вище посилання на Lavalier мікрофон на основі цієї капсули — > ви можете помітити, що згаданий вище мікрофон Rode має навіть трохи ВИЩІЙ рівень самошуму, як 15dBA (79dBA SNR).
- Мій польовий досвід із мікрофонами на базі EM272 показує, що вони чудові і досить дешеві!
- Якщо гроші для вас не важливі, купіть супер-дупер, ультранизкий автошум, перевірений і надійний Sennheiser MKH8020 та USB-інтерфейс з фантомним живленням 48V (наприклад, Focusrite Scarlett 2i2 — > я ще не тестував це з RPi), щоб отримати максимальну продуктивність.

Електретні та MEMS-мікрофони

Як мікрофони Electret Condenser (ECM), так і мікроелектромеханічні системи (MEMS) широко використовуються в акустичній екології. Кожен має свої переваги та компроміси:

Таблиця 1.1 Порівняльна характеристика Електретних та MEMS-мікрофонів

Ознака	Електретні мікрофони (ЕСМ)	Мікрофони MEMS
Чутливість	Висока чутливість (наприклад, -28 дБ для EM272Z1)	Може знадобитися додаткове підсилення через нижчу чутливість (наприклад, -38 дБ для SPU0410LR5H-QB), що підвищує складність схеми.
SNR	Часто вище на 1 кГц (наприклад, 80 дБ для Primo EM272)	Нижчий SNR у деяких моделях (наприклад, 63 дБ для SPU0410LR5H-QB) порівняно з висококласними ЕСМ, такими як EM272Z1, потенційно обмежує дальність виявлення.
Частотна характеристика	Добре працює у чутному та низькому ультразвуковому діапазоні	У плоскому стані деякі можуть захоплювати частоту до 80 кГц
Довговічність	Чутливий до вологи та навколишнього середовища	Стійкість до пилу, води, RFI/EMI
Послідовність	Помірна варіативність між одиницями	Висока послідовність завдяки напівпровідниковому процесу, ідеально підходить для багатомікрофонних масивів у великомасштабних РАМ

Продовження таблиці 1.1

Ознака	Електретні мікрофони (ЕСМ)	Мікрофони MEMS
Розмір/Форм-фактор	Громіздких; Крихіткі кабелі	Дуже маленькі; може інтегрувати АЦП, підсилювач на чіпі
Міць	Зазвичай живлення від розетки 2–5 В	Працює при 1,8–3,6 В, низькому струмі (<200 мкА)
Ціна та доступ	Економічно вигідно, моделі на кшталт WM-61A коштують \$2.65, що відповідає критеріям низької вартості.	Легко доступний у світових постачальників

Підсумовуючи, електретні мікрофони досі часто користуються перевагою для запису з чутним діапазоном через вищу чутливість і SNR, тоді як MEMS-мікрофони відзначаються міцністю, мініатюризацією та високочастотною характеристикою, що робить їх придатними для ультразвукового запису та масштабного застосування.

РОЗДІЛ 2. МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕННЯ

2.1 Методологія тестування частотної характеристики мікрофонів

Щоб охарактеризувати частотну характеристику цих мікрофонів, ми використали програмне забезпечення Room EQ Wizard (REW). Синусоїдальний сигнал свіпу генерувався і відтворювався через пару динаміків Microlab Solo7C, керовані звуковою картою ноутбука. Акустична відповідь записувалася за допомогою пристрою ARU, де кожен тестовий мікрофон підключався до MAX4466 попереднього підсилювача з максимальним підсиленням (~30 дБ), після чого відбувалася оцифрування через PCM1808 АЦП.

Для калібрування системної відповіді (без урахування впливу мікрофона) вихід звукової карти ноутбука безпосередньо підключався до входу ARU (обходячи мікрофон), а отримана відповідь зберігалася як файл калібрування звукової карти. Цей файл застосовувався в REW для корекції нелінійності звукової системи.

Оскільки жоден референсний мікрофон не використовувався для визначення абсолютної частотної характеристики самих динаміків, отримані частотні криві відображають комбіновану частотну характеристику тестових мікрофонів і динаміків. Ці криві вказують на відносну чутливість, а не на абсолютну частотну характеристику.

2.2 Пасивна акустична реєстрація звуків птахів та кажанів в польових умовах

Протокол акустичного моніторингу

Здійснювався пасивний акустичний моніторинг з метою реєстрації вокалізації птахів, а також інших звукових подій характерних для обраної місцевості.

А. Кінець літа початок осені (всього місяць), Україна вінниця, районі з приватною забудовою.

Б. З вечора 21:00 по 11:00 ранку, Запис по 10 хвилин з паузами 20 хвилин.

В. Використовувався захист приладу від дощу, вітрозахист мікрофона.

Г. Вибір місця розташування приладу (висота над ґрунтом тощо) - запропонуй свій варіант виходячи зі розповсюджених методик.

Д. Параметри запису: моно, 96кГц, 24-біт. Перепиши тезисно описане мною як розділ методики. Обґрунтуй вибір пунктів 2-5

2.3 Скрипт для розрізання аудіо запису на фрагменти на основі файлу анотації

У мене є папка з аудіофайлами та їхніми анотаціями. Аудіофайли можуть бути у форматах wav mp3 та flac (наприклад, night_20250901_053003.flac). Аудіодані були анотовані написом «Raven Pro». Файли анотації (наприклад, night_20250815_051913.Table.1.selections.txt) мають формат csv, роздільник — це tab.

Отримайте список аудіофайлів (['*.wav', '*.mp3', '*.flac']) з папки введення, які містять файл анотації

А. На основі інформації з анотованих файлів вирізайте фрагменти з аудіофайлів і зберігайте їх у підпапках розділеної папки.

- о інформація про фрагменти, які будуть вирізані та збережені, доступна у колонках 'Start Time (s)', 'End Time (s)' та 'Species' у файлах анотацій.

- о фрагмент слід зберігати в підпапці розділеної папки з іменем, що відповідає текстовій мітці в стовпці 'Вид'.

- о Залежно від значення clips_naming_convention змінної, фрагменти слід зберігати відповідно до різних іменних умов. наприклад, фрагменти 'night_20250901_053003.flac' можна зберегти як

- А) 'night_20250901_053003_10_13.flac',
'night_20250901_053003_21_14.flac' ... (де 10_13, 21_14 ... — це час початку і кінця фрагментів).
- В) 'night_20250901_053003_1.flac', 'night_20250901_053003_2.flac', ... (де 1, 2 ... є номерами фрагментів у відповідному файлі ('night_20250901_053003.flac')).
- С) '1.flac', '2.flac', ... (де 1, 2, ... — це кількість фрагментів у поточній підпапці)

Б. Під час завантаження файлу анотації скрипт повинен перевірити, чи містить він стовпець Confidence. Якщо ні, то цей стовпець слід додати до Dataframe з стандартним vale 100 у всіх чанках. Якщо цей стовпець існує у файлі анотацій, можливо, що деякі рядки в цьому стовпці мають відсутні значення. У такому випадку відсутні значення слід заповнити числом 100 (у датафреймі).

В. Існує кілька процедур вилучення фрагментів.

- 'центр'. Центри фрагментів розташовані в центрах інтервалів ['Час початку (s)' ; «Кінець часу (терміни)»]. Усі фрагменти мають бути довжиною 3 секунди (+- 3/2 сек від центру; якщо центр фрагмента розташований на відстані менше 3/2 сек від початку або кінця аудіозапису, його межі слід відповідно обирати.

- 'перекриття'. Ця процедура має розділяти вміст обмежувальної рамки (між «Час початку» та «Кінець часу») на накладаються фрагменти довжини «FRAGMENT_DURATION». Лівий край першого фрагмента дорівнює «Початку часу» коробки, а лівий край останнього фрагмента менший або дорівнює «Кінцевому часу» коробки. Якщо поточна коробка менша за 'FRAGMENT_DURATION', її слід пропускати під час обробки.

- 'випадковий'. Ця процедура видалить фрагмент, центр якого не збігається з центром коробки. Відстань між центром коробки та фрагментом має бути випадковим значенням. Ця відстань має бути такою, щоб виконувати такі умови: $(t1_box - t1_fragm) < \alpha * L$ та $(t1_box - t1_fragm) < \alpha * L$, де

$t1_fragm$ і $t2_fragm$ — час початку і кінця фрагмента; $t1_box$ і $t2_box$ — час початку і завершення коробки; альфа — це параметр, дробове число між 0 і 1; $L = t2_box - t1_box$, довжина коробки. Іншими словами, лівий/правий край коробки не може простягатися за лівий/правий край фрагмента більше, ніж $\alpha * L$. Довжина фрагмента має бути `FRAGMENT_DURATION`. Процедура обробляє крайні випадки (фрагменти біля початку/кінця аудіо). Метод створення фрагмента на основі певного поля залежатиме від значень 'Species' та 'Confidence' у стовпцях таблиці анотації для цього поля. Перед застосуванням процедур вилучення фрагментів можливі значення зі стовпця видів мають бути зіставлені з:

- тип процедури екстракції фрагмента;
- Порогове значення для оцінки довіри (`confidence_min`). Якщо інформація про збіг не вказана користувачем для певного виду, тоді будуть застосовані значення за замовчуванням (процедура вилучення фрагментів за замовчуванням — «центрування фрагментів у середині між Часом початку та Кінцевим часом»; значення за замовчуванням `confidence_min` дорівнює 0. Якщо значення 'Confidence' з таблиці анотацій для поточного боксу менше за `confidence_min`, Тоді цю коробку пропускають, і процедура екстракції фрагментів не застосовується до цієї коробки. Також скрипт має вказувати як параметр кількість разів, коли слід використовувати процедуру вилучення фрагментів, описану в пункті 2. Цю процедуру можна застосовувати до певної коробки кілька разів, у результаті чого для однієї коробки буде створено кілька різних фрагментів.

2.4 Метрики для оцінки точності нейромережі BirdNet

Для оцінки точності нейромережі BirdNet ми використовували такі метрики: Влучність, Повнота F1-міра (Precision, Recall, F1-Score) Порівняно з експертною ідентифікацією, результати BirdNET можуть бути такими: справжні позитивні результати (правильно ідентифіковані види), хибнопозитивні результати (неправильно ідентифіковані види) або

хибнонегативні (види, які не були ідентифіковані). На основі цього ми розраховували точність, відклик і F1, щоб перевірити продуктивність BirdNET. Точність (екв. 1) вимірює частку видів, ідентифікованих BirdNET, які є правильними. Recall (екв. 2) вимірює, скільки видів, ідентифікованих експертом, правильно ідентифіковані BirdNET. Через компроміс між пригадуванням і точністю, тобто максимізація одного зменшує іншу, ми використовуємо F1-бали як основний показник ефективності, які забезпечують збалансовану інтеграцію точності та відтворення (рівняння 3).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (1)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

2.5 Довірчий бал BirdNet, оцінка мінімального значення довірчого балу

Рекомендації щодо правильного використання балів BirdNET та інших результатів детекторів. Оцінка впевненості Рекомендації щодо роботи з балами BirdNET.

Оцінка порогу довіри

Щоб оптимізувати компроміс між точністю (мінімізація хибнопозитивних результатів) і відкликанням (мінімізація хибнонегативних) для видоспецифічних аналізів, ми впровадили підхід на основі даних для визначення відповідних порогів довіри BirdNET за допомогою логістичної регресії. Цей метод виходить за межі довільних фіксованих порогів, використавши ручну валідацію підмножини виявлень для статистичного

моделювання взаємозв'язку між автоматизованим балом і фактичною точністю виявлення [1].

Процедура застосовувалася індивідуально для кожного цільового виду:

А. Ручна валідація: Репрезентативна підмножина виявлень BirdNET була випадково вибіркована на сайтах дослідження та вручну перевірена експертом-людиною (ім'я автора/рецензента/ініціали) [1]. Кожному виявленню присвоюється бінарний результат:

- Справжній позитив (TP): Вид був правильно ідентифікований на момент позначення.
- Хибнопозитивний результат (FP): Вид не був присутній, або виявлення було неправильним.

Б. Модель логістичної регресії: Бінарна модель логістичної регресії була встановлена за допомогою статистичних моделей бібліотеки Python. Для кожного виду як залежну змінну використовували бінарний результат, підтверджений експертами (TP=1, FP=0). Сирий, необмежений логіт-бал, який надає BirdNET для кожного виявлення, використовувався як незалежна змінна. Було обрано використання логіт-балів замість обмежених 0-1 довірчих балів, оскільки вони відображають сирий результат фінального шару моделі і більше підходять для статистичного аналізу.

Logit score = $\ln(1/(1 - \text{оцінка довіри}))$

В. Поріговий розрахунок: Отримане логістичне рівняння регресії використовувалося для обчислення порогу логіт-балу, специфічного для виду, що відповідає бажаному рівню точності (наприклад, 95% точності, $p=0,95$). Цей поріг відображає бал, вище якого виявлення має 95% ймовірність бути справді позитивним. Поріг розраховувався за такою формулою:

$$\text{BirdNET score threshold} = \frac{\ln\left(\frac{p}{1-p}\right) - \beta_0}{\beta_{\text{BirdNET_score}}}.$$

де p — бажана ймовірність істинно позитивного (наприклад, 0,95), β_0 — перехоплення, а β_1 — коефіцієнт логіт-балу з логістичної регресійної моделі.

Усі виявлення з показником логіту BirdNET нижче відповідного порогу, специфічного для виду, були відфільтровані та виключені з подальших аналізів. Цей суворий, орієнтований на дані підхід до фільтрації забезпечує високу якість даних і підвищує достовірність результатів для всіх цільових видів [1].

Обґрунтування роцдури. Підвищення порогу неминує перетворить деякі справжні позитивні результати на хибнонегативні, але описана процедура логістичної регресії навмисно приймає цей компроміс з певної причини: основна мета цього методу порогового визначення — максимізувати точність (надійність виявлень), а не відкликати (мінімізувати хибнонегативні результати).

Ось чому процедура побудована саме так:

А. Фокус — на точності, а не на пригадуванні

Підхід логістичної регресії, який часто використовується в дослідницьких рекомендаціях (наприклад, Wood & Kahl, 2024), явно розроблений для забезпечення високої позитивної прогностичної цінності (PPV), тобто точності [20]. Дослідники часто фільтрують результати BirdNET, щоб відповісти на питання на кшталт: «Чи дійсно цей вид присутній у цьому регіоні?» або «Чи можна довіряти кожній точці на цій карті?».

У таких випадках хибнопозитивні результати значно більше шкодять аналізу, ніж хибнонегативні. Декілька пропущених виявлень (хибнонегативних результатів) можуть трохи недооцінити їхню поширеність, але високий рівень хибнопозитивних результатів може призвести до абсолютно хибних висновків про присутність, поширення або фенологію виду.

Б. Природа даних BirdNET

BirdNET за своєю суттю створює дуже велику кількість виявлень, часто з низьким порогом за замовчуванням (наприклад, 0,1), що призводить до

багатьох хибнопозитивних результатів. Виклик для дослідників полягає в тому, щоб керувати цим обсягом потенційно ненадійних даних.

- Підвищуючи поріг через логістичну регресію, ви робите правило класифікації більш жорстким.
- Це зменшує кількість хибнопозитивних (покращує точність), але збільшує кількість хибнонегативних (зменшує відклик).
- Компроміс між точністю та відтворенням є притаманним будь-якому порогу класифікації.

В. Визначення набору даних валідації

Конкретна модель логістичної регресії вище використовує лише *позитивні виявлення* з результатів BirdNET, які були надіслані на аналіз на людях.

- Ви не перевіряєте вручну кожну секунду аудіо, де BirdNET нічого не повідомляє.
- Отже, дані, що використовуються для логістичного регресійного аналізу, за своєю суттю не містять інформації про справжні відсутності, що необхідно для прямого розрахунку рівня хибнонегативного показника ($FN / [TP + FN] = \text{Відкликання}$) по всьому набору даних.
- Модель вивчає ті точки, які вона передбачила як позитивні, визначаючи, наскільки надійними були ці прогнози.

Зведення

Описаний метод є практичним, орієнтованим на дані підходом, щоб забезпечити високу надійність (високу точність) фінального фільтрованого набору даних для досліджень, де хибнопозитивні результати є серйозною проблемою. Вона приймає неминучий наслідок нижчого рівня запам'ятовування (більше хибнонегативних результатів) як необхідну ціну для точності даних.

Якщо ваша мета дослідження — мінімізувати хибнонегативні результати (наприклад, у контексті збереження, де пропустити крик рідкісного виду критично важливо), ви б зробили:

А. Ручну перевірку великих фрагментів аудіо, включаючи ті, де BirdNET нічого не повідомляв.

Б. Використовуйте іншу метрику, наприклад F1-бал, або чітко пріоритетно відтворюйте відкликання під час вибору порогу (що передбачає зниження порогу, збільшення хибнопозитивних результатів).

$$\text{Confidence score} = \frac{1}{1 + e^{\text{logit score} \cdot \text{sensitivity}}}$$



РОЗДІЛ 3. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТУ ТА ЇХ ОБГОВОРЕННЯ

3.1 Блок схема автономного записувального пристрою (АЗП)

Загалом, аудіорекордер складається з мікрофона, передпідсилювача, аналогово-цифрового перетворювача (АЦП) та процесорного блока. Мікрофон захоплює звукові хвилі та перетворює їх на аналоговий електричний сигнал. Попередній підсилювач підсилює слабкий сигнал від мікрофона до придатного для використання рівня. АЦП оцифрує цей сигнал. Нарешті, процесор керує та обробляє цифрові аудіодані, виконуючи такі завдання, як фільтрація, пониження дискретизації та збереження даних на носії. Цей пристрій може бути реалізований як за допомогою мікроконтролера, так і одноплатного комп'ютера (SBC).

Вибір конкретних електронних компонентів для кожного блоку визначався наступними критеріями:

- Енергоефективність – Реєстратор буде жити автономно (наприклад, від батареї або сонячної панелі), тому мінімізація споживання енергії є необхідною для подовження часу роботи.
- Смуга пропускання – система повинна підтримувати частоти дискретизації, придатні для запису як вокалізації птахів (зазвичай до 10–12 кГц), так і ехолокації кажанів (які можуть перевищувати 100 кГц).
- Низький рівень шуму – Щоб забезпечити якісні аудіозаписи, усі компоненти — особливо мікрофон, попередній підсилювач і АЦП — повинні створювати мінімальний шум.
- Низька вартість – пасивний акустичний моніторинг (РАМ) у дослідженнях дикої природи часто вимагає розміщення багатьох автономних блоків реєстрації (ARU) на великій території. Щоб зробити масштабне впровадження можливим, виробничі витрати кожного пристрою мають бути максимально низькими.

- Масштабованість і гнучкість – бажано, щоб дизайн підтримував модульні оновлення або модифікації, такі як додавання GPS, бездротового зв'язку або додаткових датчиків.

3.1.1 Процесорний блок

Було розроблено дві версії аудіореєстратора, які відрізнялися процесорними блоками. Кожен підхід має свої компроміси щодо продуктивності, енергоспоживання, вартості та розширення.

3.1.2 Версія на основі Raspberry Pi

Перша версія використовує одноплатний комп'ютер із сімейства Raspberry Pi — зокрема Raspberry Pi Zero 2 W — як процесорний блок (@fig:f1). Цей SBC працює на повній операційній системі Linux і підтримує широкий спектр програмних бібліотек і периферійних пристроїв, що робить його надзвичайно гнучким для розробки та майбутніх оновлень.

Ця модель має такі переваги:

- Низька вартість (приблизно \$15–20 USD);
- Помірне споживання енергії (приблизно 1,5 Вт у піковій фазі, 0,8 Вт у нормальній роботі та <0,2 Вт у режимах простою/сну з управлінням живленням);
- 512 МБ оперативної пам'яті, достатньо для буферизації аудіоданих та виконання легких аналітичних завдань;
- Повна підтримка SD-карт великої ємності та USB-пристроїв для зберігання даних;
- Простота розробки програмного забезпечення, включаючи підтримку Python, C/C++, ALS та різних інструментів Linux;
- Вбудоване бездротове з'єднання (Wi-Fi, Bluetooth);
- Підтримка віддаленої передачі даних через зовнішні модулі, такі як SIMCom A7670E, що дозволяє зв'язок на основі GSM/4G під час польових розгортань.

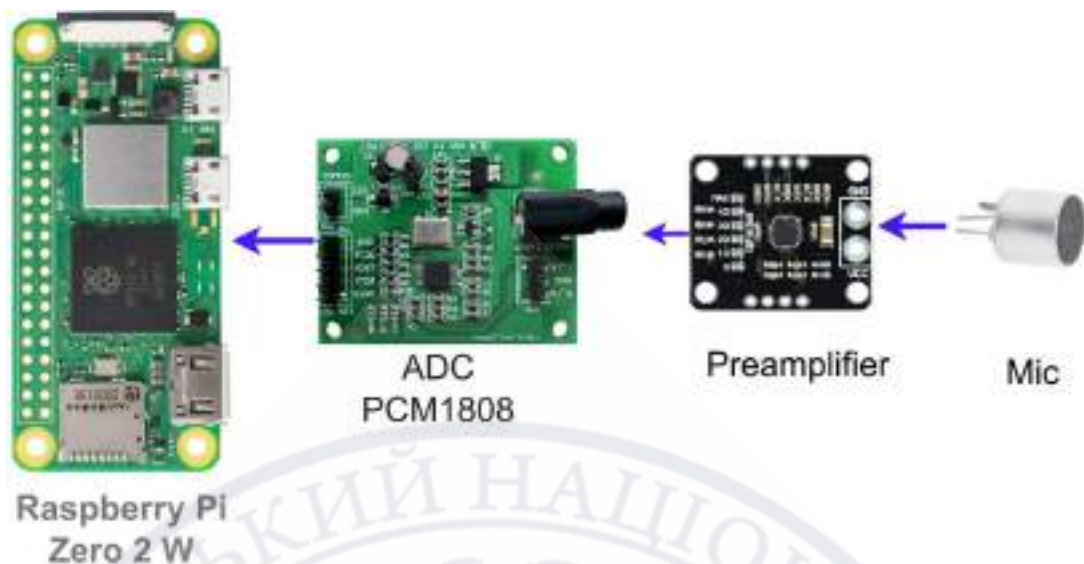


Рис 3.1 - Пристрій створений на платформі Raspberry PI Zero 2 W

3.1.3 Версія на основі ESP32

Друга версія базується на ESP32 — недорогому мікроконтролері з інтегрованим Wi-Fi та Bluetooth (@fig:f2). ESP32 широко використовується в IoT-застосуваннях завдяки компактному форм-фактору, низькому енергоспоживанню та достатнім обчислювальним ресурсам для легких аудіозавдань.

Ключові характеристики:

- двоядерний процесор Tensilica Xtensa (до 240 МГц);
- 520 КБ SRAM, з додатковою PSRAM, доступною для деяких варіантів;
- Режими наднизького енергоспоживання, що робить його дуже придатним для розгортання на батарейках;
- Підтримка I²S, що дозволяє прямий інтерфейс із зовнішніми аудіо-АЦП, такими як PCM1808;
- Ультра-низьке енергоспоживання: Працює приблизно на 0,6 Вт під час активної обробки та при 0,01 Вт у режимі глибокого сну, ідеально для тривалих розгортань з обмеженими енергоресурсами.

Переваги у порівнянні з Raspberry Pi:

- Нижче енергоспоживання;
- Швидші переходи між пробудженням/сну;

- Дешевша вартість і простіше обладнання.

Мінуси у порівнянні з Raspberry Pi:

- Обмежена обчислювальна потужність: Двоядерний процесор ESP32 з частотою 240 МГц і 520 КБ SRAM менш здатні виконувати складні аудіозавдання або великі набори даних порівняно з чотирядерним процесором і середовищем Linux у Raspberry Pi.

- Немає нативної операційної системи: Не має повноцінної операційної системи, що вимагає розробки прошивки в середовищах, таких як FreeRTOS або Arduino, що може ускладнювати розробку.

- Обмеження зберігання: Використовує меншу флеш-пам'ять або зовнішні SD-карти, що обмежує місткість вбудованого зберігання порівняно з підтримкою Raspberry Pi для більших рішень для зберігання.

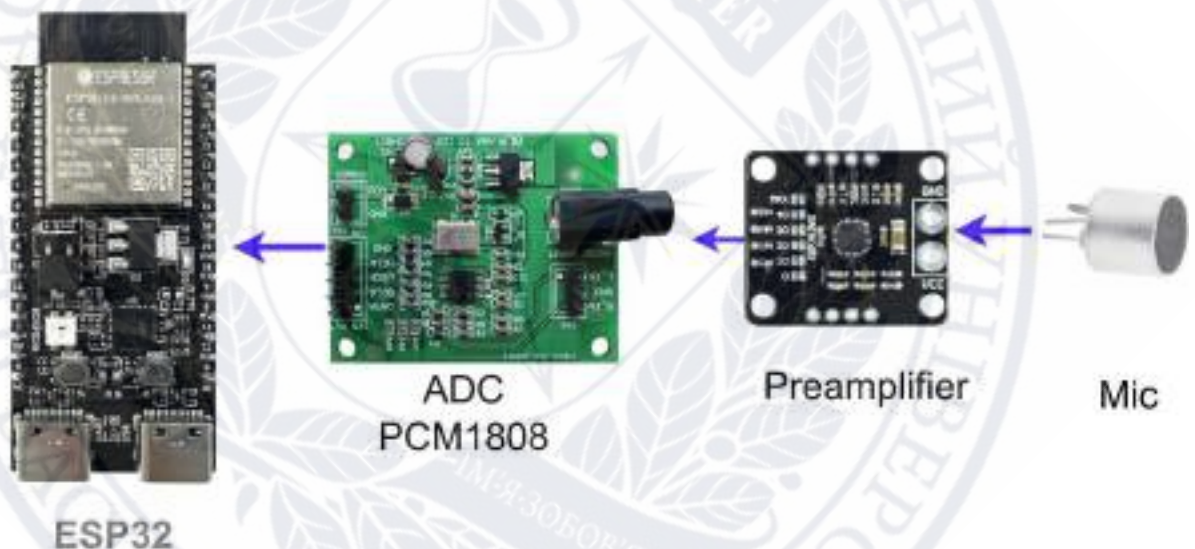


Рис 3.2 - Пристрій створений на платформі ESP32

3.2 Вибір моделі мікрофона

Технічні характеристики вибраних мікрофонів

Використовуючи наведені вище літературні дані, ми обрали найбільш відповідні моделі мікрофонів, доступні для придбання в Україні. Далі ми надамо технічні характеристики вибраних мікрофонів, отриманих із технічних шитів. Таблиця 1 надає порівняльний огляд ключових характеристик,

використання та ціноутворення. Фігури відображають типову криву частотної характеристики вибраних мікрофонів.

Таблиця 3.1 Порівняння мікрофонів для пасивного акустичного моніторингу

Мікрофон	Формат	Виробник	SNR	Чутливість	Макс. Вхід (dB)	Напруга (V)	Ціна
EM172	Капсули	Прімо	80	-28 ±3	119	3	24
EM258	Капсули	Прімо	74	-32 ±3	115	3	6.3
EM272Z 1	Капсули	Прімо	80	-28 ±3	119	3	24
EM189T	Капсули	Прімо	68	-43 ±3	128	2	7.3
EM289	Капсули	Прімо	68	-44 ±3	128	2	9
EM288Z 1	Капсули	Прімо	70	-43 ±3	127	2	13
EM294Z 1	Капсули	Прімо	68	-43,5 ±3	141	2	31
WM-61A	Капсули	Panasonic	>62	-35 ±3	-	2	2.65
AOM-5024L-HD-R	Капсули	PUI Audio	80	-24	119	3	5
ICS-40180	MEMS	Долучайтесь			-	3.3	2.5
ICS-40720	MEMS	Долучайтесь	70	-38	124	3.3	6

Продовження таблиці 3.1

Мікрофон	Формат	Виробник	SNR	Чутливість	Макс. Вхід (dB)	Напруга (V)	Ціна
ICS-40730	MEMS	Долучайтесь	74	-38	124	3.3	5
SPU0410LR5H-QB	MEMS	Ноулз	63	-38	106.9	3.6	1
PMM-3738-VM1000-R	Piezo MEMS	PUI Audio	62	-38	-	1.8	3
FG-23629-P16	Капсули	Ноулз		-53 ±3	-	1.3	72

Типові криві частотної характеристики:

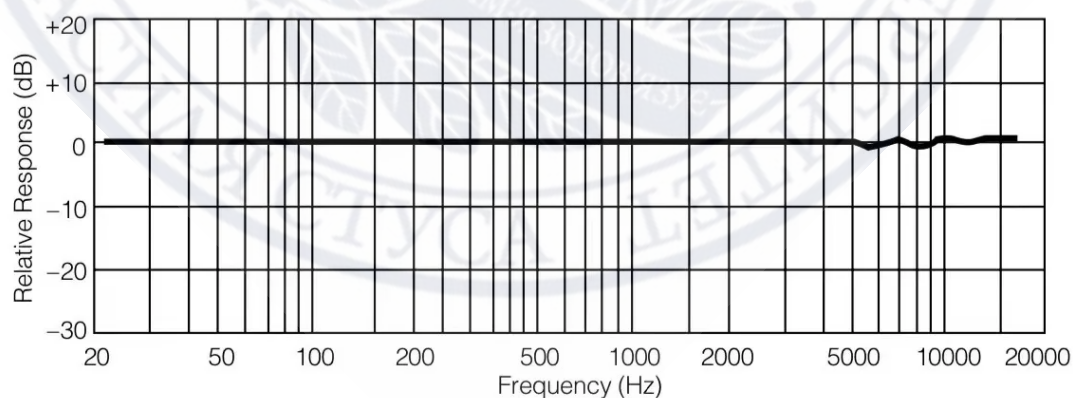


Рис 3.3 - Частотна характеристика мікрофону EM172

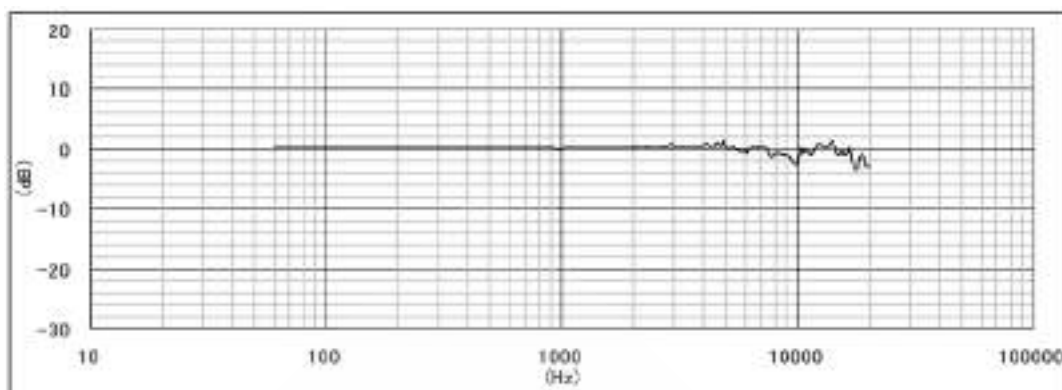


Рис 3.4 - Частотна характеристика мікрофону EM258

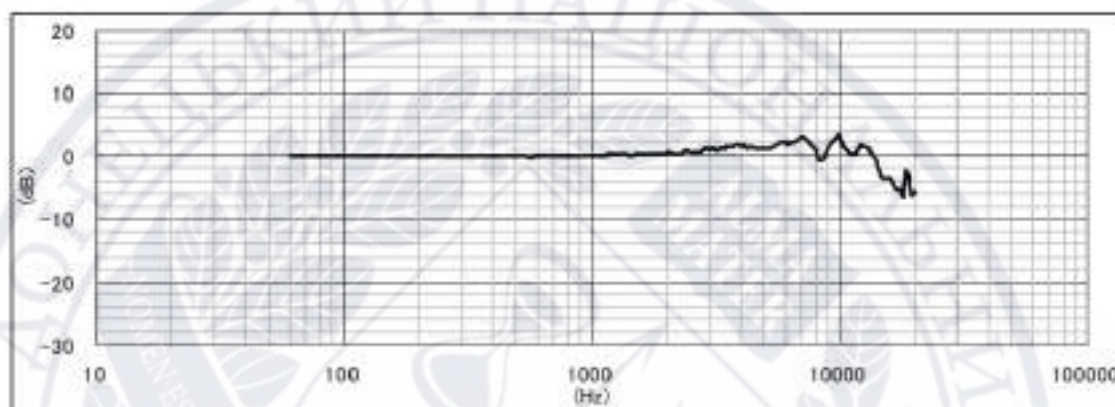


Рис 3.5 - Частотна характеристика мікрофону EM272Z1

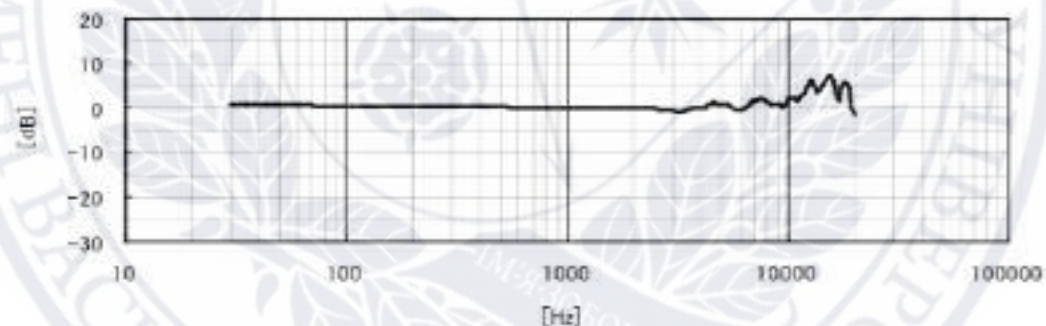


Рис 3.6 - Частотна характеристика мікрофону EM189T

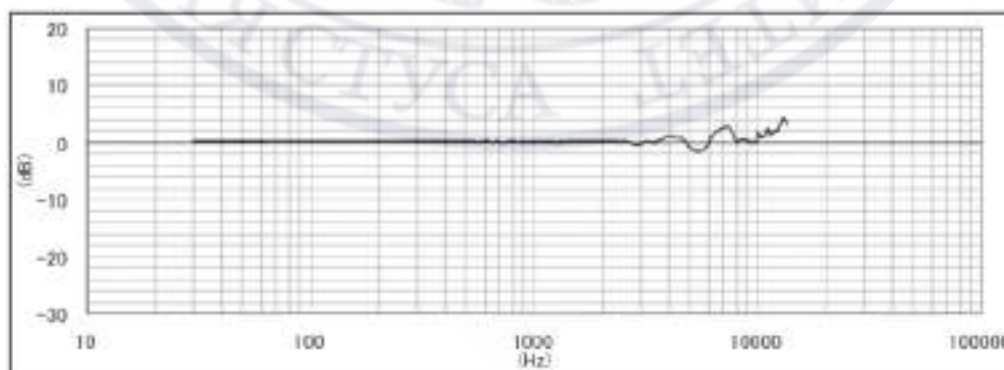


Рис 3.7 - Частотна характеристика мікрофону EM289

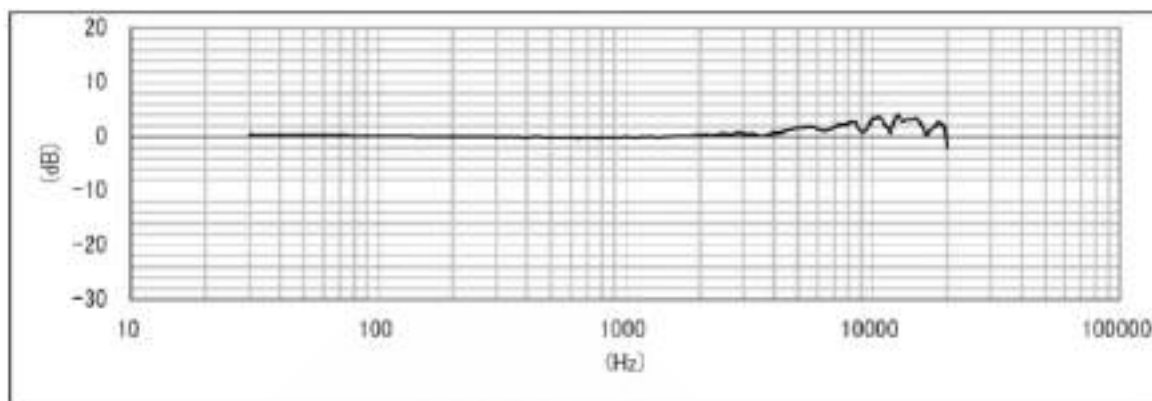


Рис 3.8 - Частотна характеристика мікрофону EM294Z1

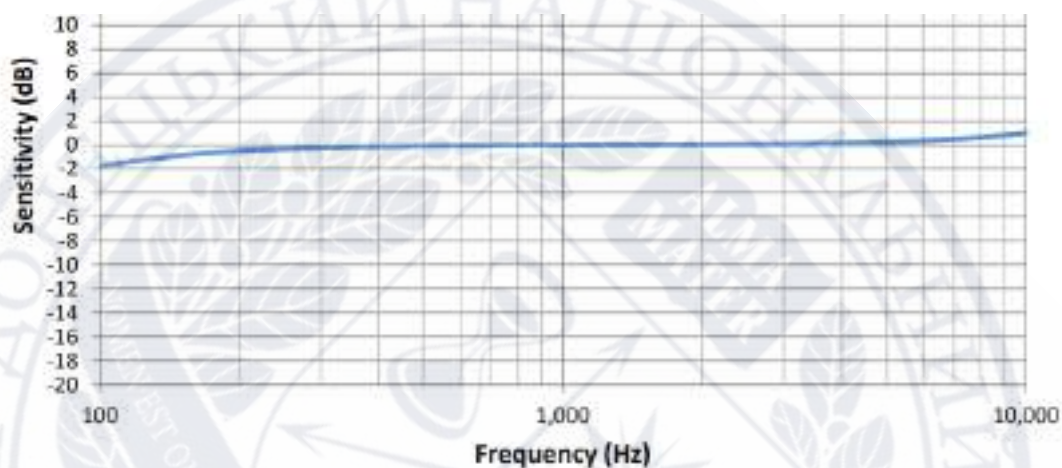


Рис 3.9 - Частотна характеристика мікрофону WM-61A

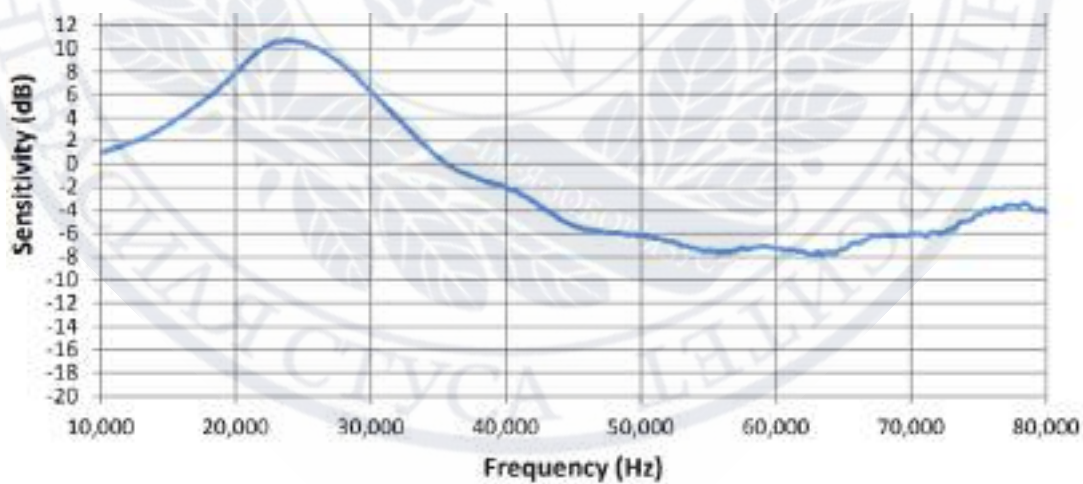


Рис 3.10 - Частотна характеристика мікрофону ICS-40180

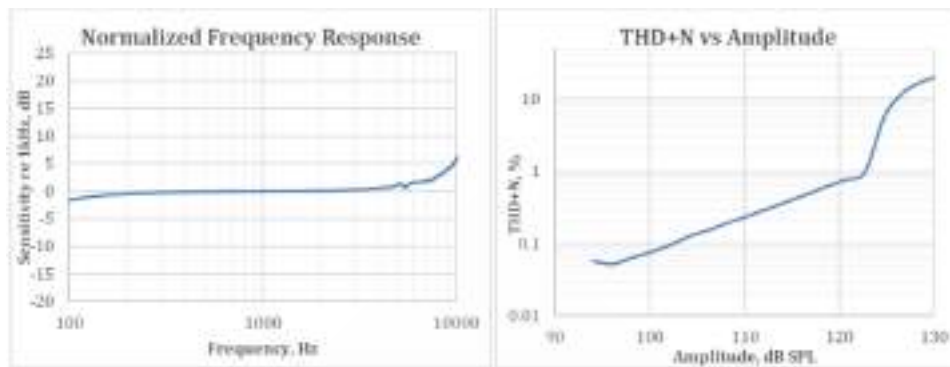


Рис 3.11 - Частотна характеристика мікрофону SPU0410LR5H-QB

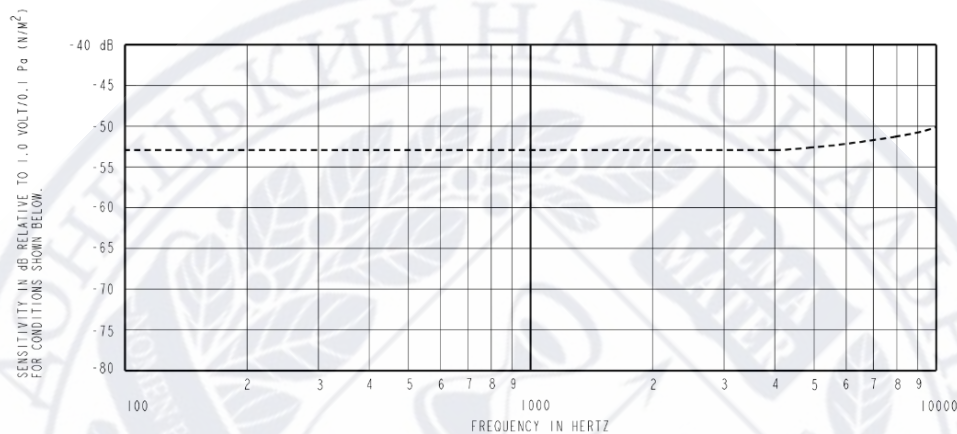


Рис 3.12 - Частотна характеристика мікрофону PMM-3738-VM1000-R

3.2.1 Висновок і рекомендації

При розробці ARU, оптимізованого як для звукового, так і для ультразвукового запису, мікрофони з високим SNR і достатньою пропускнуою здатністю є ключовими. На основі порівняльного аналізу та доступності:

- Primo EM272Z1 та AOM-5024L-HD-R — чудові електретні капсули для моніторингу звукової дальності, забезпечуючи високий рівень чутливості та чутливість за низькою вартістю.
- ICS-40730 та SPU0410LR5H-QB представляють відповідні MEMS-варіанти для височастотного та надійного використання на відкритому повітрі, з рівною характеристикою та екологічною стійкістю.
- Використання мікрофонів MEMS з низьким рівнем шуму особливо корисне у довгострокових, самотних розгортаннях завдяки їхній міцності, розміру та стабільності.

- Коли потрібен ультразвуковий запис (наприклад, ехолокація кажана), віддають перевагу мікрофонам MEMS, таким як ICS-40730 або VM1000.

- У всіх випадках акустичне екранування, захист від вітру та герметизація від вологи необхідні для підтримки роботи мікрофона на відкритому повітрі.

Для масштабних розгортань модульна конструкція з мікрофонами, які можна замінити користувачем, може бути корисною, дозволяючи польові оновлення та обслуговування. Вибір завжди має балансувати акустичну продуктивність, енергоспоживання, вартість і фізичну міцність залежно від цілей моніторингу.

3.3 Тестування придбаних зразків мікрофонів

Відсутність офіційних дилерів в Україні для кількох брендів мікрофонів, обговорених у розділі 1.4.1, ускладнює закупівлю. Щоб вирішити цю проблему, ми придбали кілька моделей електретних мікрофонів у магазині Micverve на платформі AliExpress, ритейлері, що спеціалізується на мікрофонах. Придбані моделі включають WM-61A (\$1.83) та Primo EM258 (\$6), а також два недорогі no-name мікрофони: один з меншим діаметром («no-name 1») і один з більшим («no-name 2»). Щоб перевірити, чи відповідають ці мікрофони своїм заданим характеристикам і підтвердити придатність для пасивного акустичного моніторингу, ми провели серію тестів для оцінки їхньої частотної характеристики та ультразвукової чутливості.

3.3.1 Некалібрована частотна характеристика

На рисунку 1 можна зробити такі спостереження:

А. EM258 демонструє найвищу загальну чутливість у діапазоні 200 Гц – 7 кГц, демонструючи рівні SPL щонайменше на 5 дБ вищі, ніж у всіх інших протестованих мікрофонах.

SPL (рівень звукового тиску) вимірюється в децибелах (дБ) і позначає рівень акустичного тиску, який фіксує мікрофон. Вищі значення SPL свідчать про більшу чутливість мікрофона до звуку.

Б. WM-61A та Noname #1 демонструють схожі патерни відгуків до 7 кГц, при цьому Noname #1 демонструє трохи менші коливання.

В. На частотах вище 7 кГц чутливість мікрофонів Noname значно знижується, особливо Noname #2. Тим часом WM-61A та EM258 зберігають вищу чутливість до 20 кГц, при цьому криві збігаються після ~15 кГц.

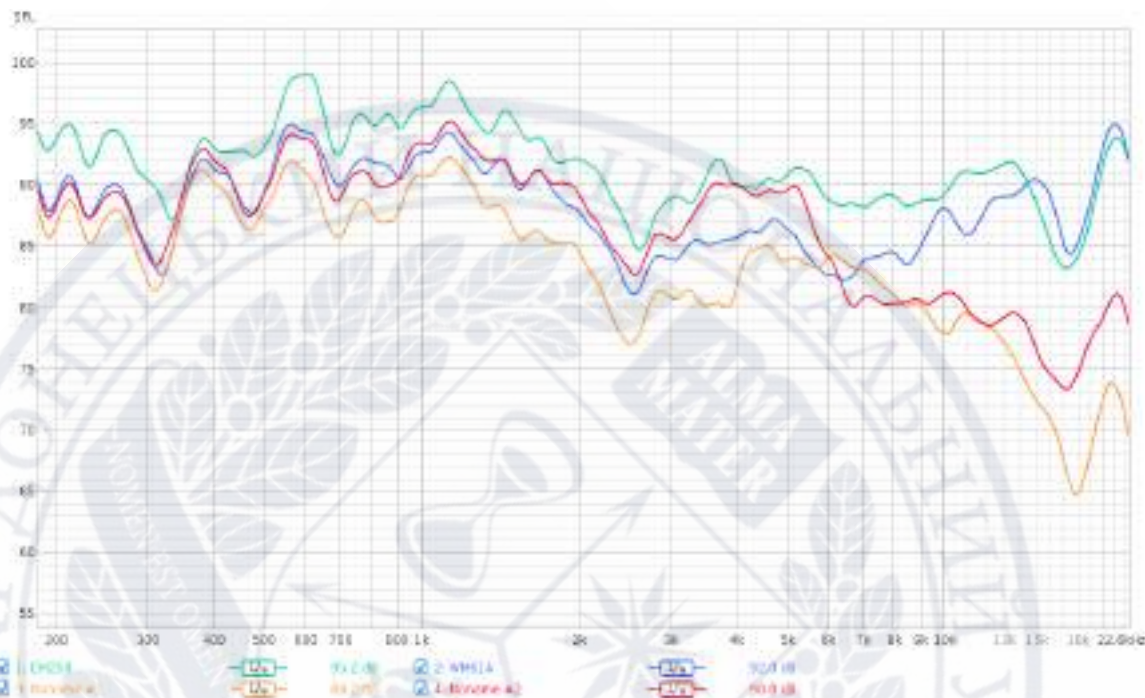


Рис 3.13 - Некалібрована частотна характеристика мікрофонів

3.3.2 Калібровані частотні характеристики

Ми перекалібрували частотні характеристики відносно кожного з еталонних мікрофонів (WM-61A та EM258), щоб чіткіше спостерігати порівняльні тенденції.

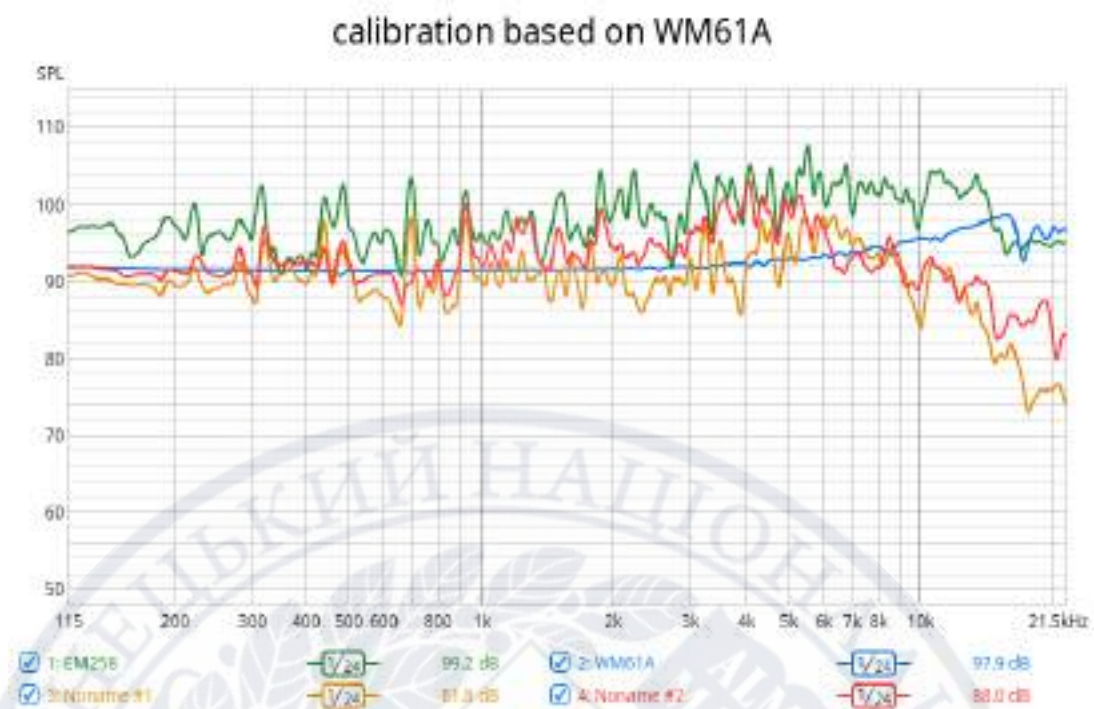


Рис 3.14 - Частотна характеристика калібрована по мікрофону WM61A

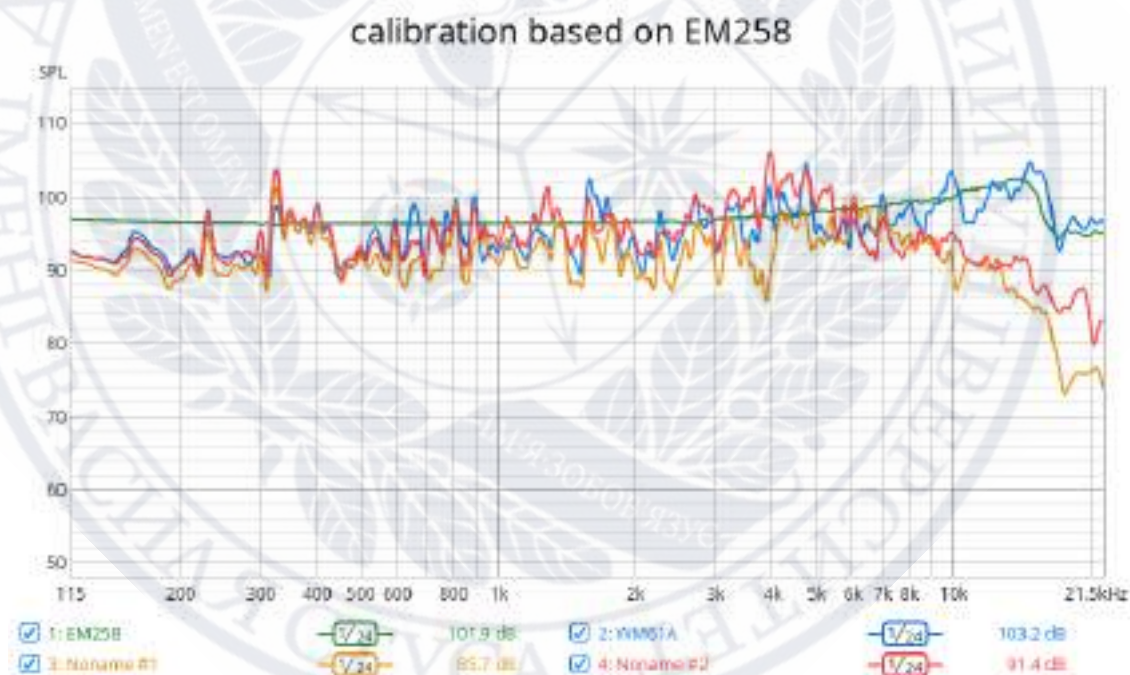


Рис 3.15 - Частотна характеристика калібрована по мікрофону EM258

В обох калібруваннях результати підтверджують, що:

- Мікрофони Noname демонструють значне зниження частоти вище ~9 кГц, що підтверджує обмежену смугу пропускання.

- EM258 та WM-61A зберігають відносно плоскі відгуки по всьому тестованому спектру, що робить їх більш придатними для широкосмугового моніторингу.

3.3.3 Тест на чутливість до ультразвуку (40 кГц)



Рис 3.16 - Ультразвуковий мікрофон HC-SR04 та його частотна характеристика

Щоб оцінити здатність мікрофонів захоплювати ультразвукові звуки, такі як ехолокація кажана, ми використали ультразвуковий датчик відстані HC-SR04 як джерело ультразвуку, який випромінює сигнал 40 кГц. HC-SR04 генерує імпульс із восьми імпульсів 40 кГц понад 200 мкс при спрацьовуванні. Для генерації безперервного сигналу 40 кГц сенсор активувався кожні 200 мкс, використовуючи Arduino UNO з малюнком, показаним нижче:

```
int trigPin = 11;
int echoPin = 12;
void setup() {
  Serial.begin(9600);
  pinMode(trigPin, OUTPUT);
  pinMode(echoPin, INPUT);
}
void loop() {
  digitalWrite(trigPin, HIGH);
  delayMicroseconds(10);
  digitalWrite(trigPin, LOW);
  delayMicroseconds(200);
}
```

Джерело ультразвуку було розташоване на відстані 30 см і 80 см від ARU. Спектрограми записів, створені за допомогою Audacity, показують великі пікові висоти на 40 кГц, що свідчить про чутливість до ультразвукових сигналів. Результати такі:

- При 30 см (рисунок 6) пікові висоти на 40 кГц становлять -61 дБ для «noname 1», -74 дБ для «noname 2», -40 дБ для WM-61A і -41 дБ для EM258 (пікові висоти оцінювалися за допомогою графіків спектра FFT в Audacity). Це свідчить про те, що WM-61A та EM258 значно чутливіші до ультразвуку 40 кГц, з піками на 20–33 дБ вищими, ніж у «noname» мікрофонів.

- При 80 см пікові висоти знижуються до -71 дБ для «noname 1», -84 дБ для «noname 2», -50 дБ для WM-61A і -50 дБ для EM258, зберігаючи подібну перевагу у 20–34 дБ для WM-61A та EM258 над «noname» моделями.

- Спектрограми також показують, що HC-SR04 випромінює чутні частоти до 3 кГц поряд із ультразвуком 40 кГц. Для WM-61A та EM258 спектрограми демонструють більшу потужність у цьому чутному діапазоні з чіткими гармоніками порівняно з слабшими, менш визначеними відгуками мікрофонів «noname». Це додатково підтверджує висновок, що WM-61A та EM258 мають вищу чутливість і ширшу частотну характеристику як у звуковому, так і в ультразвуковому діапазонах.

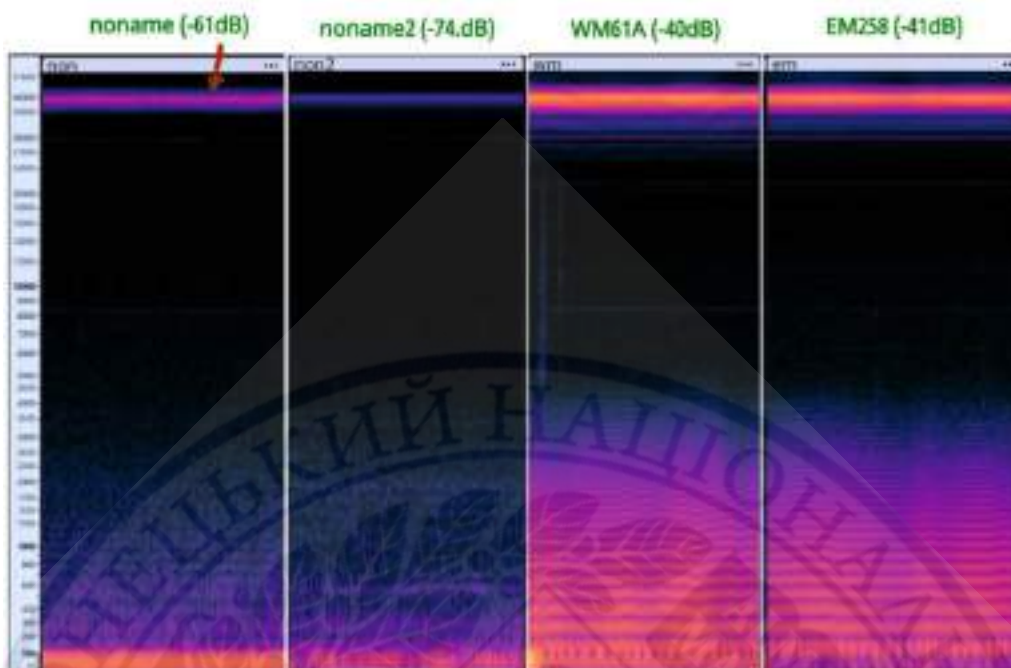


Рис 3.17 - Спектрограма чутливості мікрофонів до ультразвуку

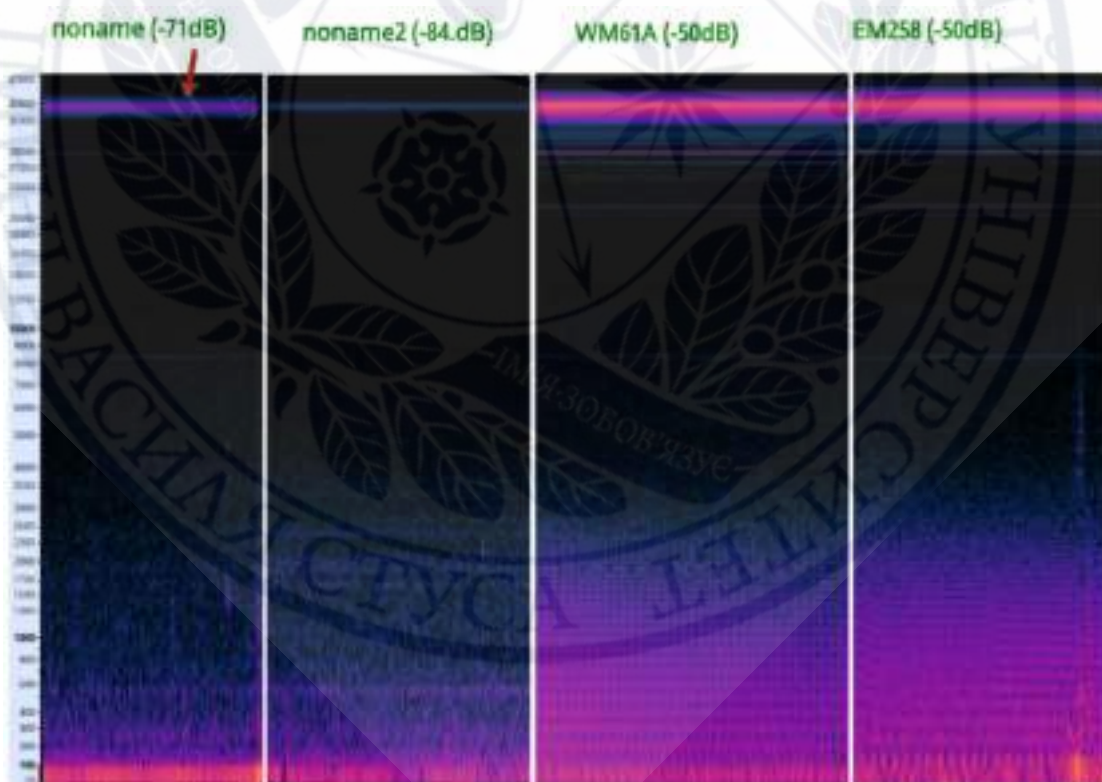


Рис 3.18 - Спектрограма чутливості мікрофонів до ультразвуку

3.3.4 Висновок

Тестування підтверджує:

- EM258 перевершує мікрофони WM-61A та noname за чутливістю, особливо нижче 7 кГц.
- WM-61A підтримує надійну продуктивність у широкому частотному діапазоні, включно з частковою ультразвуковою характеристикою до 20 кГц.
- Мікрофони Noname демонструють недостатню чутливість до високих частот і ультразвуку і не підходять для акустичного моніторингу, що потребують повноспектрального або ультразвукового захоплення.
- І EM258, і WM-61A виявляли сигнал 40 кГц від ультразвукового сенсора на обох тестових відстанях, тоді як мікрофони Noname мали значні втрати сигналу.

Таким чином, Primo EM258 і WM-61A визнані відповідними мікрофонними елементами для ARU застосувань, спрямованих на пташиний спів і моніторинг ехолокації кажанів. EM258 віддають перевагу там, де потрібна підвищена чутливість, тоді як WM-61A залишається життєздатною бюджетною альтернативою для звукового запису.

3.4 Розпізнавання видів птахів за допомогою BirdNet

як запускали (використовуючи скрипти написані на кафедрі) які види знайшло, видова різноманітність? як від часу залежить

для розпізнавання видів птахів за їх співами Ми використовували неймережу birdnet. розпізнавання здійснювалося в пакетному режимі, всі отримані аудіофайли передавались на вхід неймережі на виході отримувались результати у вигляді файлів анотації у форматі програми Raven Pro. birdnet розділяє переданий на її вхід файл на трьох секундні фрагменти комірки і для кожного фрагменту здійснює розпізнавання видів птахів, результатом є список усіх птахів які вона здатна розпізнати та так звані довірчі оцінки. довірча оцінка (confidence score) знаходиться в діапазоні від нуля до одиниці і показує рівень довіри що в цій комірці є спів що належить саме цьому

виду. для того щоб обмежити список можливих видів вводять порогове значення для довірчої оцінки, всі види що мають для поточної комірки значення довірчої оцінки менше за порогове відкидаються. додатково можна зменшити список видів вказавши інформацію про локалізацію зроблених аудіозаписів і час. в цьому випадку програма автоматично залишає тільки ті види які є характерні для даного регіону і даного періоду часу.

отримані від нейромережі результати зберігалися нами у вигляді файлів анотації в форматі Raven Pro. цей файл містить текстову таблицю стовпчики якої представляють наступну інформацію час початку та кінець комірки в якій було детектовано той чи інший вид, назву детектованого виду як наукове ім'я та як загальне ім'я українською та англійською також зберігалися значення довірчих балів всього було отримано 644 файли, таблиця представляє фрагмент такого файла:

Таблиця 3.2 - Види птахів які розпізнала нейромережа BirdNet

Вибір	Час початку	Кінець часу	Вид	Довірчий бал	Загальна назва
1	42.0	45.0	Phylloscopus collybita	38.561	Вівчарик-ковалик
2	45.0	48.0	Parus major	30.836	Синиця велика
3	51.0	54.0	Лебедь колумбійний	78.737	Лебідь чорнодзьобий
4	54.0	57.0	Turdus merula	23.668	Дрізд чорний
5	57.0	60.0	Лебедь колумбійний	36.481	Лебідь чорнодзьобий
6	60.0	63.0	Лебедь колумбійний	72.862	Лебідь чорнодзьобий

Продовження таблиці 3.2

Вибір	Час початку	Кінець часу	Вид	Довірчий бал	Загальна назва
7	63.0	66.0	Phylloscopus collybita	26.522	Вівчарик-ковалик
8	66.0	69.0	Parus major	22.527	Синиця велика
9	69.0	72.0	Phylloscopus collybita	32.64	Вівчарик-ковалик
10	93.0	96.0	Лебедь колумбійний	54.042	Лебідь чорнодзьобий
11	102.0	105.0	Ficedula hypoleuca	24.83	Мотузка мухоловка
12	105.0	108.0	Parus major	20.811	Синиця велика
13	108.0	111.0	Parus major	34.018	Синиця велика
14	123.0	126.0	Лебедь колумбійний	79.811	Лебідь чорнодзьобий
15	126.0	129.0	Parus major	62.665	Синиця велика
16	132.0	135.0	Cyanistes caeruleus	28.77	Синиця блакитна
17	132.0	135.0	Parus major	22.878	Синиця велика
18	138.0	141.0	Parus major	50.899	Синиця велика
19	150.0	153.0	Parus major	35.159	Синиця велика
20	153.0	156.0	Parus major	31.081	Синиця велика
21	153.0	156.0	Cyanistes caeruleus	21.73	Синиця блакитна

Продовження таблиці 3.2

22	168.0	171.0	Ficedula hypoleuca	24.749	Мотузка мухоловка
23	171.0	174.0	Parus major	27.492	Синиця велика
24	183.0	186.0	Dendrocopos major	66.275	Дятел звичайний
25	189.0	192.0	Pica pica	99.386	Сорока звичайна

Далі представлено результати розпізнавання BirdNet за весь період моніторингу. Величина мінімального довірчого балу була встановлена 0.2.

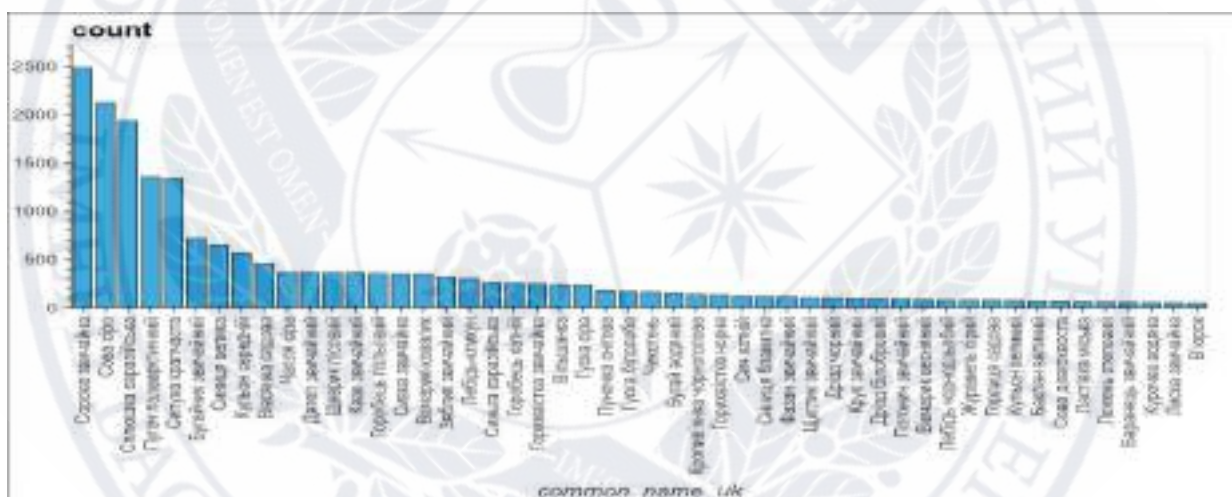


Рис 3.19 - Види які правильніше розпізнала нейромережа BirdNet



Рис 3.20 - Діаграма кількості реєстрацій для кожного зареєстрованого виду за весь період моніторингу

Динаміка реєстрацій видів на прикладі двох видів, що мають найбільшу кількість реєстрацій за весь період моніторингу (*Pica pica*, *Strix aluco*), всього ідентифіковано 175 видівдетектовано 19940 вокалізацій цих видів (17 годин зі 100 годин записів)

3.5 Оцінка точності нейромережі BirdNet

Оцінка точності нейромережі BirdNet проводилась на основі порівняння її результатів (файл анотації BirdNet) з файлом анотації звукозапису

зробленому вручну в програмі RavenPro. На основі порівняння файлів анотації ми підраховували кількість істинно позитивних (TP), істинно негативних (TN) та хибно негативних (FP) 3 секундних комірок (chunks) в яких нейромережа здійснює детекцію видів. На основі кількості TP, TN та FP ми розраховували метрики precision, recall та F1-score за наступними формулами:

Для підрахунку TP, TN та FP ми використовували наступну процедуру:

Вручну було проанатовано 5 аудіозаписів. Розмічались наступні звукові події: вокалізація пташок (видів 'Parus major', 'Pica pica', 'Phylloscopus collybita', 'Chicken', 'Phoenicurus phoenicurus', 'Dendrocopos major', 'Muscicapa striata', 'Streptopelia decaocto', 'Passer montanus'), а також звуки біогного, техногенного походження 'Dog', 'Sound1', 'Other', 'Fieldfare', 'Insect', 'Flight', 'Car', 'Technics', 'ARU'. Кількість подій для всіх розмічених аудіо записів наведена в таблиці manual_annot_summary_1.csv

manual_annot_summary_1.csv Табл. Кількість аудіо подій ідентифікованих при ручній розмітці в 5 аудіо файлах

З аудіо подій, що розмічались вручну BirdNet здатна розпізнавати наступні 8 (співи птахів): 'Parus major', 'Pica pica', 'Phylloscopus collybita', 'Phoenicurus phoenicurus', 'Dendrocopos major', 'Muscicapa striata', 'Streptopelia decaocto', 'Passer montanus'. Відповідно ми проаналізували точність детекції саме цих 8 аудіо подій. Кількість 3-секундних комірок в яких BirdNet в усіх 5 записах знайшла ці види наведена в таблиці .як можна бачити найбільшу кількість детекцій мають види (>30): 'Pica pica', 'Parus major', 'Passer montanus' bn_annot_summary_1.csv Табл Кількість детекцій видів 'Parus major', 'Pica pica', 'Phylloscopus collybita' , 'Phoenicurus phoenicurus', 'Dendrocopos major', 'Muscicapa striata', 'Streptopelia decaocto', 'Passer montanus', що були ідентифіковані BirdNet в 5ти аудіофайлах які також піддавались ручному анотуванню.

на гістограмах представлені розподіли довірчих оцінок на всіх 5 записах для декількох видів пташок. 'Pica pica', 'Parus major', 'Passer montanus'

Pica pica - всі > 20, більшість > 50

'Parus major', всі > 20 , ,більшість < 50

'Passer montanus' , всі > 20 , ,більшість < 50

використовуючи процедуру описану в методиці ... для всіх детекцій видів птахів нейроереженню BirdNet на 5 аудіозаписах, нами було обчислено:

А. кількість TP, FP та FN детекцій

Б. метрики точності, відклику та результатів Формули-1

Всі ці величини були обчислені нами для різних порогових значень confidence score в діапазоні 0-90 на графіках (нижче) показані залежності кількість TP, FP та FN детекцій від порогового значення confidence score. Можна бачити, що для видів ... Ріса ... при малих порогових значеннях confidence score кількість FP детекцій (зайві детекції, детекції неіснуючого співу, співу немає але нейромережа його детектувала) є малою. . переважна більшість детекцій є TP (вірні детекції - є спів і його детектовано) та FN (спів є але нейромережа його не детектувала). Причому $FN > TP$.

Метою збільшення порогового значення Confidence Score (CS) є зменшення кількості FP. Але разом з тим, збільшуючи поріг CS ми відсіюємо TP, що мають значення CS менше за порогове. Причому ці TP детекції стають FN. Задача - знайти баланс між кількістю видалених FP та TP. Зменшення кількості FP (при використанні високих порогових значень балу впевненості (CS)) збільшує точність (precision), але одночасне зменшення TP знижує повноту (recall). Задача вибору оптимального порогового значення CS зводиться до знаходження максимального CS при якому точність зростає а повнота впаде до припустимих значень. На скільки потрібно збільшити точність і наскільки припустимо зменшити повноту залежить від пріоритетів користувача та дослідницьких цілей. Наприклад, низька повнота не є проблематичною, якщо дані використовуються для оцінок чисельності популяції, скоригованих на недосконале виявлення (наприклад, моделювання заселеності). У той час як для оцінки поведінки або щільності потрібна висока повнота.

На графіках (іще нижче) наведені залежності precision, recall та F1-score від порогового значення CS.

Оскільки FP від початку мало, зміна FP мало впливають на відповідні метрики (на збільшення точності). В той же час вибір великих порогових значень може призводити до значного падіння повноти. Таким чином при виборі порога слід орієнтуватись в першу чергу на повноту. Потрібно збільшити поріг до тієї межі де точність ще зростає (де FP не зменшили до 0), а повнота падає не значним чином. Те саме можна зробити використовуючи F1-оцінку. F1-оцінка є сумою точності та повноти значить в нашому випадку ми маємо збільшувати порогове значення CS до тих пір поки F1 оцінка не почне зменшуватися (це буде відбуватись тоді коли точність вийде на плато, а повнота продовжить зменшуватися)

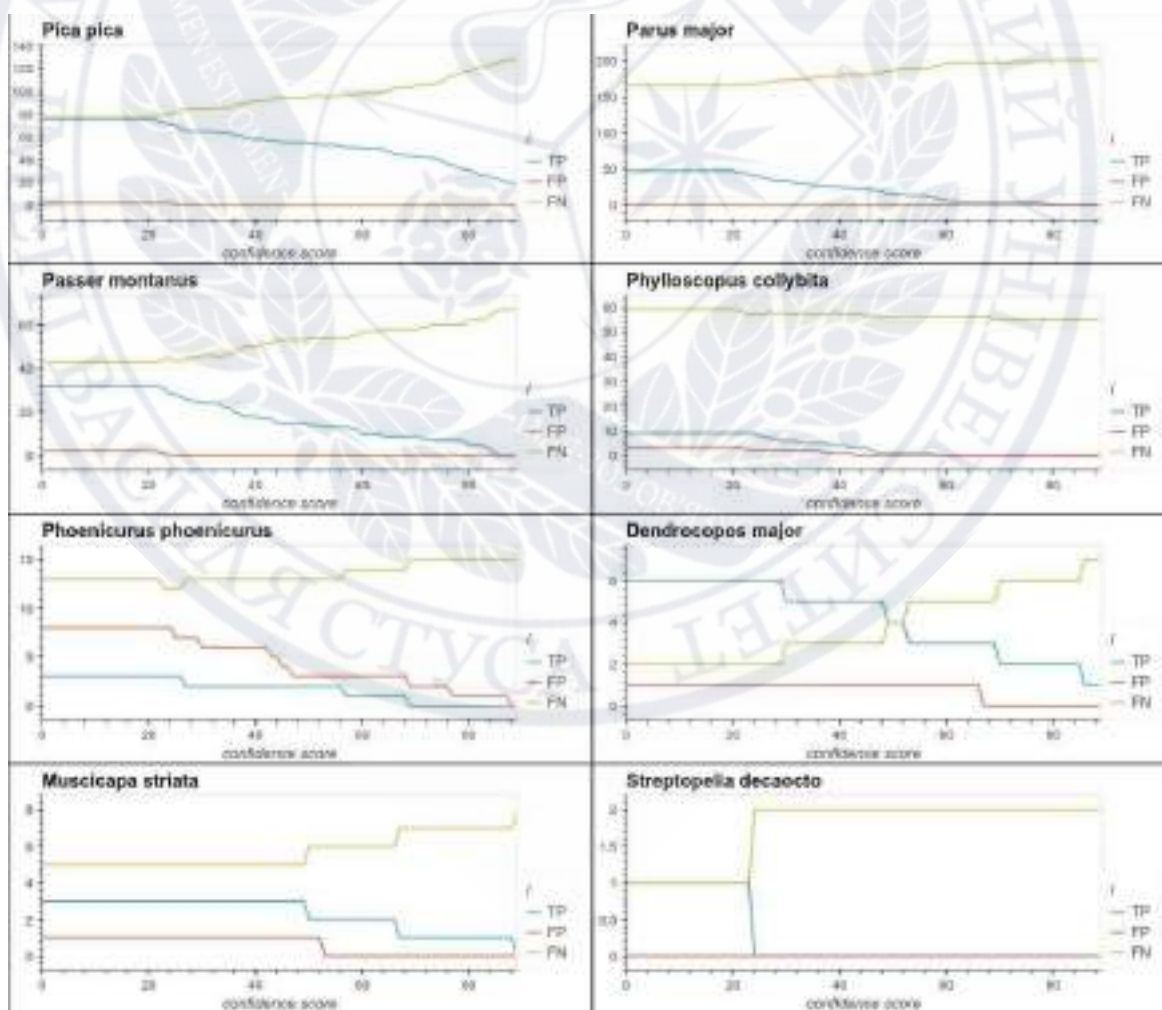


Рис 3.21 - Залежність повноти до кількості правильних детекцій

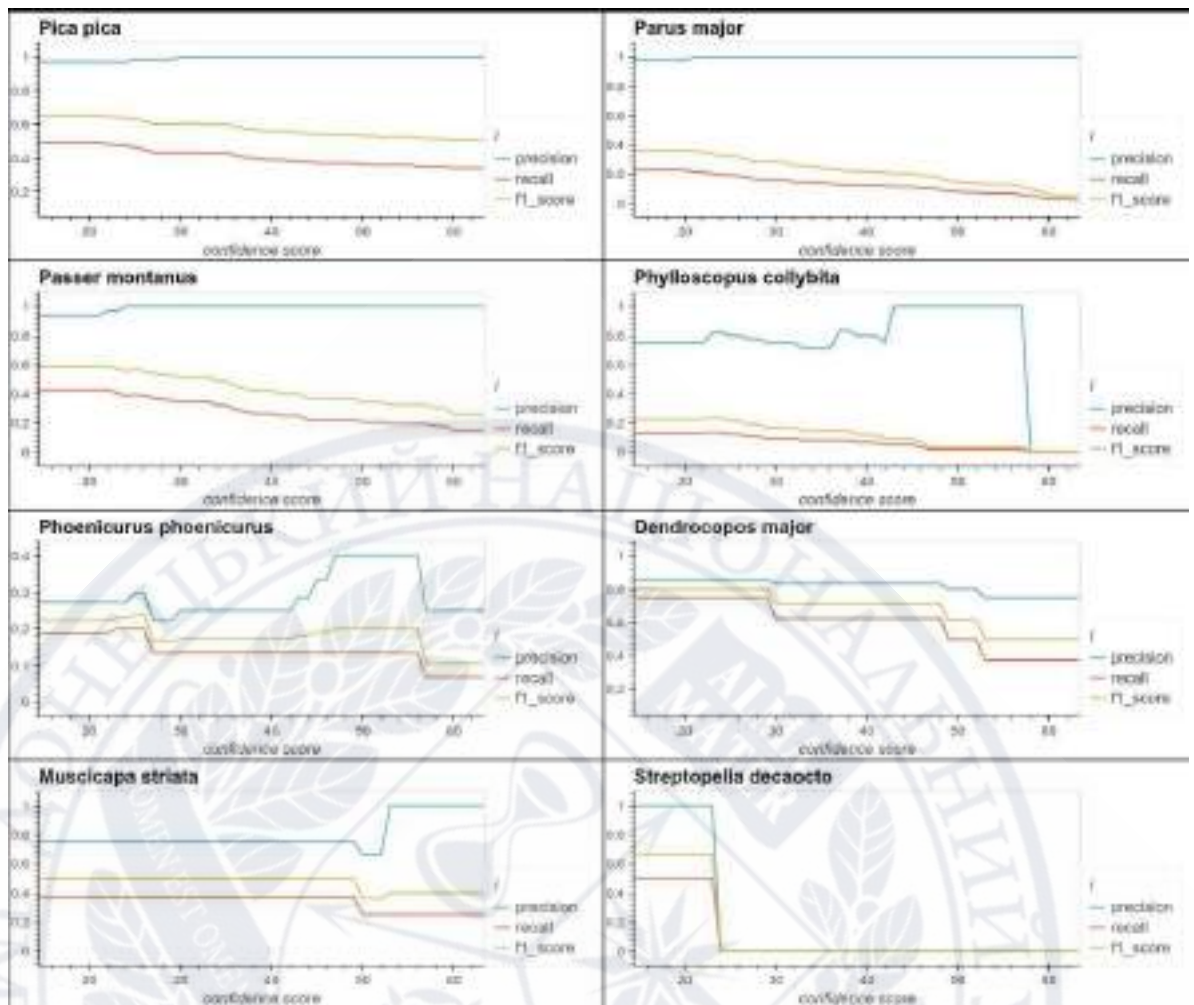


Рис 3.22 - Залежність повноти до кількості правильних детекцій інших видів

3.6 Вибір граничних значень довірчої оцінки BirdNet

Загальна характеристика поведінки TP, FP і FN залежно від порогу CS

Оптимальне порогове значення довірчої оцінки (Confidence Score, CS) є ключовим параметром, що визначає баланс між істинно позитивними (TP), хибнопозитивними (FP) та хибнонегативними (FN) детекціями нейромережі BirdNET. Використання одного універсального порогу не є доцільним, оскільки якість роботи нейромережі істотно залежить від характеристик конкретного акустичного середовища. У рамках цього дослідження точність BirdNET оцінювалася шляхом порівняння автоматичних анотацій з вручну створеними розмітками в RavenPro, виконаними на тих самих аудіофайлах.

Обидві анотації були приведені до однакового формату з розбиттям запису на 3-секундні фрагменти, що дозволило обчислити значення TP, FP та FN для кожного порогового значення CS.

Отримані залежності продемонстрували характерну загальну тенденцію: при низьких значеннях CS кількість FP була невеликою для всіх видів, тоді як співвідношення TP і FN характеризувалося значною перевагою FN. Це означає, що за низького CS BirdNET зазвичай пропускає істотну частину вокалізацій, хоча при цьому майже не генерує хибнопозитивних виявлень. Для середніх порогових значень (близько 20–50), де зосереджена більшість значень CS у таких видів, як *Pica pica*, *Parus major* і *Passer montanus*, спостерігалось поступове зменшення FP, яке супроводжувалося частковою втратою TP. У цьому діапазоні відбувається найбільш суттєвий перехід від низької точності до помірної, без різкого падіння повноти. Подальше збільшення CS до високих значень (понад 50–60) приводило до швидкого зменшення кількості TP, тоді як FP наближалися до нуля. У результаті recall різко знижувався, тоді як precision досягав плато й переставав змінюватися.

Компроміс між точністю та повнотою при збільшенні порога CS

Підвищення порогу CS має двоїстий вплив на результати детекції. З одного боку, зростання порогового значення дозволяє ефективно зменшувати кількість FP, підвищуючи точність класифікації. З іншого боку, значна частина істинних детекцій (TP) відсіюється разом з FP, що веде до збільшення частки FN та падіння повноти. Особливістю цього дослідження було те, що кількість FP у вихідних даних була невеликою, тому їхнє зменшення при збільшенні CS майже не впливало на зростання точності. Натомість зниження TP за високих CS мало суттєвіший вплив, призводячи до помітного погіршення recall. Отже, у нашому випадку головним чинником, який визначає оптимальність порога, є не зростання точності, а характер зміни повноти та загальної узгодженості між precision і recall.

Оскільки надто високі пороги завжди призводять до істотної втрати TP, вибір CS є пошуком компромісу між мінімізацією FP та уникненням різкого

падіння recall. Цей компроміс може бути різним залежно від цілей дослідження. Наприклад, для оцінювання чисельності популяцій, де головним завданням є виявлення динаміки, певне зниження повноти є допустимим. Натомість поведінкові дослідження або задачі оцінювання щільності популяцій можуть потребувати максимальної повноти, навіть якщо це збільшує кількість FP.

Використання F1-оцінки як критерію для вибору оптимального порога

F1-score є ефективним інструментом оцінювання якості роботи моделі в умовах, коли необхідно збалансувати точність та повноту. У межах цього дослідження поведінка F1-оцінки слугувала основним критерієм для визначення оптимального CS. При збільшенні CS F1-оцінка зростала до певного моменту, досягаючи максимуму, після якого починала знижуватися. Причиною цього було те, що precision досягав стадії насичення через мінімальну кількість FP, тоді як recall продовжував падати внаслідок різкого скорочення TP. Таким чином, момент відхилення F1-оцінки від максимуму чітко відображав межу, після якої збільшення CS призводить до втрати коректних детекцій без помітного виграшу в точності.

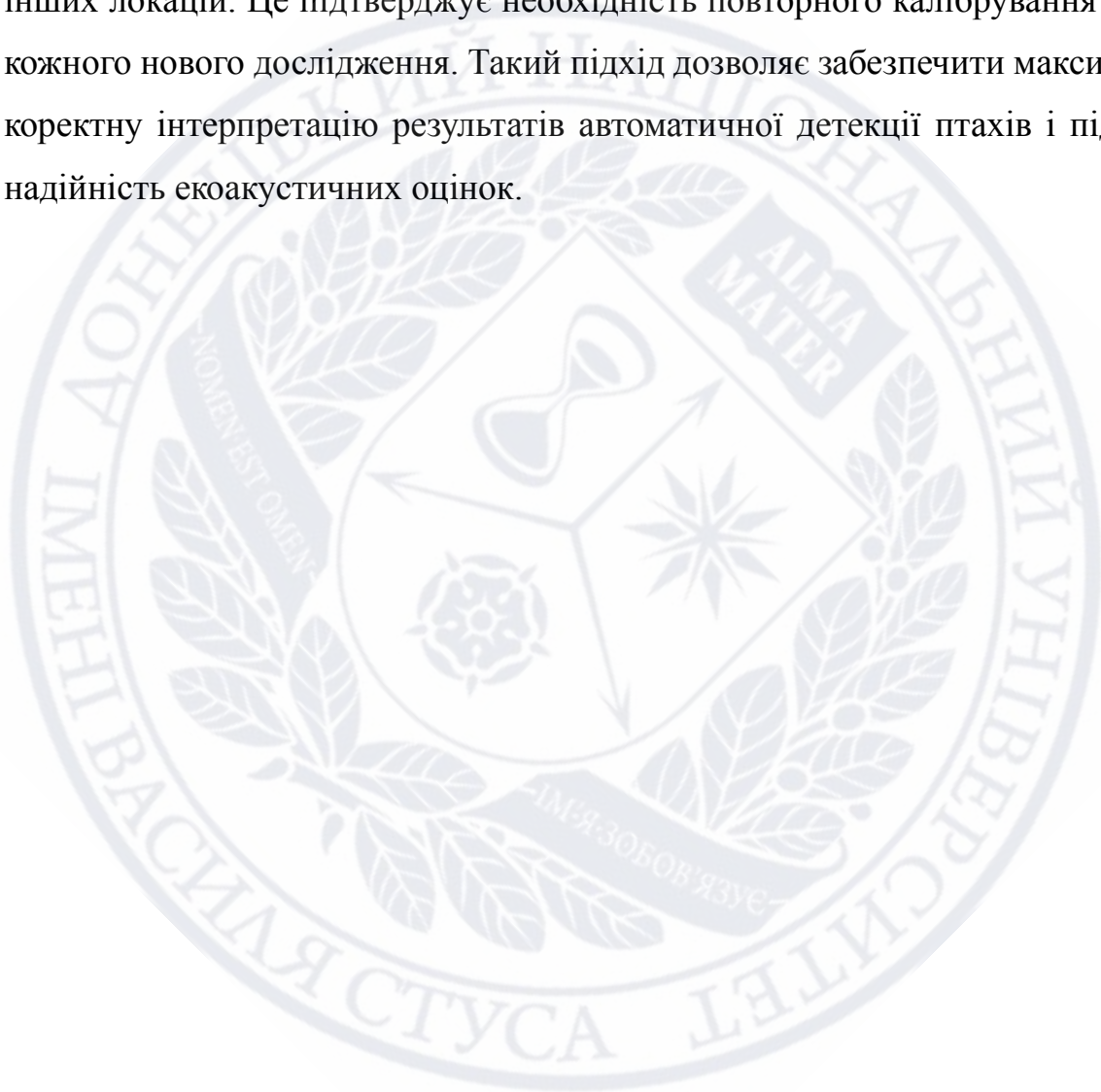
Аналіз виявив, що оптимальні значення CS для видів, представлених у нашій вибірці, здебільшого знаходилися в інтервалі від 20 до 40. У цьому діапазоні FP зменшувалися, але TP ще залишалися на достатньо високому рівні, щоб забезпечити максимальні або близькі до максимальних значення F1-оцінки. Пороги понад 50–60 у більшості випадків призводили до різкого падіння recall і, відповідно, до зниження F1-score, попри майже повну відсутність FP.

Узагальнений висновок про вибір граничного порога CS

З урахуванням отриманих даних вибір оптимального граничного значення CS повинен базуватися на точці, у якій F1-оцінка або досягає максимуму, або перестає зростати. Це значення є верхньою межею, після якої збільшення порога призводить до непропорційної втрати TP. У проведеному дослідженні

такою межею стали порогові значення, що переважно лежать у діапазоні 20–40. Саме цей діапазон забезпечує найкращий баланс між зменшенням кількості FP та збереженням високого рівня TP, що є критично важливим у задачах пасивного акустичного моніторингу.

Визначений оптимальний поріг CS є специфічним для конкретного акустичного середовища, у якому проводилися записи, і може відрізнятися для інших локацій. Це підтверджує необхідність повторного калібрування CS для кожного нового дослідження. Такий підхід дозволяє забезпечити максимально коректну інтерпретацію результатів автоматичної детекції птахів і підвищує надійність екоакустичних оцінок.



ВИСНОВКИ

1. **За результатами літературного огляду** встановлено, що вибір мікрофона є критичним фактором для ефективного пасивного акустичного моніторингу дикої природи. Аналіз наукових публікацій і технічної документації показав, що для запису вокальної активності птахів і кажанів ключове значення мають низький власний рівень шуму, широка та рівномірна амплітудно-частотна характеристика, висока чутливість і стабільність параметрів у польових умовах. Огляд доступних на ринку моделей дозволив визначити групу мікрофонів, потенційно придатних для застосування в недорогих системах акустичного моніторингу.

2. **У ході вимірювання амплітудно-частотних характеристик** досліджуваних мікрофонів встановлено їхні реальні частотні межі, рівномірність підсилення та чутливість у діапазонах, характерних для вокалізацій птахів і ультразвукових сигналів кажанів. Отримані результати дали змогу оцінити придатність кожної моделі для конкретних завдань акустичного моніторингу та виявити обмеження, пов'язані з ослабленням сигналу на високих частотах або підвищеним рівнем шуму.

3. **Планування та проведення польових записів** підтвердило можливість застосування розробленого пристрою для автономної реєстрації акустичної активності птахів і кажанів у реальних умовах середовища. Записи охоплювали різні періоди доби та умови шумового фону, що дозволило зібрати репрезентативний набір аудіоданих для подальшого аналізу. Отримані результати засвідчили стабільну роботу пристрою та його придатність для довготривалого моніторингу.

4. **Автоматична детекція видів птахів** за допомогою нейронної мережі BirdNET продемонструвала високу ефективність аналізу великих обсягів аудіоданих. Нейромережа успішно ідентифікувала значну частину вокалізацій птахів без необхідності ручної обробки, що підтверджує доцільність використання сучасних методів машинного навчання в задачах екоакустичного моніторингу.

5. **Оцінка точності детекції та аналіз довірчих оцінок** для обраних видів птахів показали, що оптимальні порогові значення confidence score залежать від акустичних умов запису та рівня шумового фону. Визначення граничних значень дозволило зменшити кількість хибнопозитивних виявлень і підвищити надійність результатів автоматичної ідентифікації. Отримані дані підтверджують необхідність адаптації параметрів BirdNET до конкретних умов дослідження.



СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Acevedo M. A., Villanueva-Rivera L. J. Using automated digital recording systems as effective tools for the monitoring of birds and amphibians // *Wildlife Society Bulletin*. 2006. Vol. 34, № 1. P. 211–214.
2. Blumstein D. T., Mennill D. J., Clemins P. et al. Acoustic monitoring in terrestrial environments using microphone arrays: Applications, technological considerations and prospectus // *Journal of Applied Ecology*. 2011. Vol. 48, № 3. P. 758–767. DOI: 10.1111/j.1365-2664.2011.01993.x.
3. Darras K. F. A., Deppe F., Fabian Y. et al. High microphone signal-to-noise ratio enhances acoustic sampling of wildlife // *PeerJ*. 2020. Vol. 8. e9955. DOI: 10.7717/peerj.9955.
4. Darras K., Kolbrek B., Knorr A. et al. Assembling cheap, high-performance microphones for recording terrestrial wildlife: The Sonitor system // *F1000Research*. 2021. Vol. 7. 1984. DOI: 10.12688/f1000research.17511.3.
5. Furnas B. J., Callas R. L., Zeller K. A. Using automated recorders and occupancy models to monitor common forest birds across a large geographic region // *Journal of Wildlife Management*. 2017. Vol. 81, № 2. P. 325–337.
6. Kahl S., Wood C. M., Eibl M., Klinck H. BirdNET: A deep learning solution for avian diversity monitoring // *Ecological Informatics*. 2021. Vol. 61. 101236. DOI: 10.1016/j.ecoinf.2021.101236.
7. Kasten E. P., Gage S. H., Fox J., Joo W. The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology // *Ecological Informatics*. 2012. Vol. 12. P. 50–67.
8. Merchant N. D., Fristrup K. M., Johnson M. P. et al. Measuring acoustic habitats // *Methods in Ecology and Evolution*. 2015. Vol. 6, № 3. P. 257–265. DOI: 10.1111/2041-210X.12330.
9. Mydlarz C., Nacach S., Roginska A. et al. The implementation of MEMS microphones for urban sound sensing // *Proceedings of the 137th Audio Engineering Society Convention*. 2014. P. 740–748.

10. Mydlarz C., Salamon J., Bello J. P. The implementation of low-cost urban acoustic monitoring devices // *Applied Acoustics*. 2017. Vol. 117. P. 207–218. DOI: 10.1016/j.apacoust.2016.06.010.
11. Mydlarz C., Shamon C., Baglione M., Pimpinella M. The design and calibration of low-cost urban acoustic sensing devices // *Proceedings of EuroNoise*. 2015.
12. URL:
https://www.conforg.fr/eurnoise2015/output_directory/data/articles/000497.pdf
(дата звернення: 27.06.2025).
13. Pieretti N., Farina A., Morri D. A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI) // *Ecological Indicators*. 2011. Vol. 11, № 3. P. 868–873.
14. Raven Pro: Interactive Sound Analysis Software (Version 1.6). Ithaca, NY: Cornell Lab of Ornithology, 2019.
15. Solo (Whytock & Christie, 2017) Sonitor system (K. Darras et al., 2021) EcoPi (Wood et al., 2024) (*Recording Technology - ecoPi*, n.d.) Недорогі міські акустичні сенсорні пристрої (Mydlarz et al., 2015) (Mydlarz et al., 2014) (Mydlarz et al., 2017)
16. Sueur J., Farina A., Gasc A., Pieretti N., Pavoine S. Acoustic indices for biodiversity assessment and landscape investigation // *Acta Acustica united with Acustica*. 2014. Vol. 100, № 4. P. 772–781.
17. Vlad G. Best microphones for recording nature and wildlife sounds // *earth.fm*. 2022.
18. URL: <https://earth.fm/recording-advice/21633/> (дата звернення: 27.06.2025).
19. Whytock R. C., Christie J. Solo: An open source, customizable and inexpensive audio recorder for bioacoustic research // *Methods in Ecology and Evolution*. 2017. Vol. 8, № 3. P. 308–312. DOI: 10.1111/2041-210X.12678.

20. Wood C. M., Günther F., Rex A. et al. Real-time acoustic monitoring facilitates the proactive management of biological invasions // *Biological Invasions*. 2024. Vol. 26, № 12. P. 3989–3996. DOI: 10.1007/s10530-024-03426-y.
21. Zach Poff. DIY wildlife microphones and bat detectors.
22. URL: <https://zachpoff.com>.

