

**Володимир Вікторович Казимир¹, Андрій Іванович Роговенко²,
Костянтин Олександрович Ткаченко³**

¹доктор технічних наук, професор кафедри інформаційних та комп'ютерних систем

Національний університет «Чернігівська політехніка» (Чернігів, Україна)

E-mail: vvkazymyr@gmail.com. ORCID: <https://orcid.org/0000-0001-8163-1119>. ResearcherID: Q-2925-2016

²кандидат технічних наук, доцент кафедри інформаційних та комп'ютерних систем

Національний університет «Чернігівська політехніка» (Чернігів, Україна)

E-mail: arogovenko@gmail.com. ORCID: <https://orcid.org/0000-0003-4594-5692>. ResearcherID: G-3926-2014

³магістр кафедри інформаційних та комп'ютерних систем

Національний університет «Чернігівська політехніка» (Чернігів, Україна)

E-mail: kostiantintkachenko@gmail.com. ORCID: <https://orcid.org/0009-0000-0927-0080>

МЕТОД РОЗПІЗНАВАННЯ АКУСТИЧНИХ СИГНАЛІВ ЗА ДАНИМИ СПЕКТРАЛЬНОГО АНАЛІЗУ ЗА ДОПОМОГОЮ НЕЙРОННИХ МЕРЕЖ

У статті запропоновано метод розпізнавання акустичних сигналів на основі спектрального аналізу та нейронних мереж для детекції цільових об'єктів в умовах негативного співвідношення сигнал/завада. Досліджено специфічну задачу акустичної ідентифікації, коли цільовий сигнал має нижчий рівень потужності порівняно з фоновими завадами від джерела аналогічної природи. На основі попереднього аналізу гармонічної структури досліджуваних сигналів визначено частотний діапазон 700-850 Гц як оптимальний для розрізнення цільового та фонового компонентів. Розроблено кастомний аналізатор для вилучення 21 спектральної ознаки з обраного частотного діапазону з використанням часових вікон тривалістю 1 секунда з перекриттям 0,5 секунди. Запропоновано архітектуру повнозв'язної нейронної мережі з двома прихованими шарами для класифікації векторів ознак. Експериментальна валідація методу продемонструвала досягнення точності детекції 88 % при рівні хибних спрацьовувань 4 %. Метод забезпечує компактне представлення акустичної інформації та обчислювальну ефективність за рахунок фокусування на попередньо визначеному інформативному частотному діапазоні замість обробки повного спектра сигналу.

Ключові слова: акустичне розпізнавання; спектральний аналіз; нейронні мережі; низький SNR; частотна селекція; векторні ознаки.

Рис.: 8. Бібл.: 14.

Актуальність теми дослідження. Розвиток технологій безпілотних літальних апаратів супроводжується зростанням їх застосування у цивільних та комерційних сферах, що одночасно створює нові виклики для забезпечення безпеки повітряного простору. Традиційні системи виявлення повітряних цілей, розроблені для великих пілотованих літальних апаратів, виявляються недостатньо адаптованими для детекції малорозмірних безпілотних об'єктів через їх низький радіолокаційний переріз розсіювання та здатність до польоту на низьких висотах [1].

Акустичні методи детекції набувають особливого значення в умовах, де радіолокаційні та оптичні системи мають обмеження. Водночас переважна більшість існуючих систем акустичного розпізнавання призначена для стаціонарних умов застосування з контрольованим акустичним середовищем [2]. Практичні потреби вимагають розробки систем детекції, здатних функціонувати на рухомих платформах, які самі є джерелами інтенсивних акустичних завад.

Постановка проблеми. Основною проблемою, яка розглядається у дослідженні, є розпізнавання акустичних сигналів в умовах, коли цільовий сигнал має нижчий рівень потужності порівняно з фоновими завадами від власної платформи детекції. Традиційні методи акустичного аналізу, що працюють з повним спектром сигналу або стандартними акустичними дескрипторами, виявляються неефективними за таких умов через домінування завадового компонента [3].

Існуючі підходи до акустичної детекції зосереджуються переважно на обробці широкосмугових сигналів із формуванням спектрограм великої розмірності, що вимагає значних обчислювальних ресурсів і не завжди забезпечує необхідну точність в умовах

складної акустичної обстановки. Відсутність методів, орієнтованих на виділення вузьких інформативних частотних діапазонів, обмежує можливості створення компактних і ефективних систем детекції.

Додатковою проблемою є необхідність забезпечення оперативності обробки при мінімальних вимогах до обчислювальних ресурсів, що особливо актуально для мобільних платформ з обмеженим енергоспоживанням. Це вимагає розробки спеціалізованих алгоритмів, здатних поєднувати детальний спектральний аналіз обраних частотних діапазонів із компактним представленням отриманих ознак.

Аналіз останніх досліджень і публікацій. Дослідження в галузі акустичного розпізнавання сигналів характеризуються широким застосуванням методів глибинного навчання для розв'язання задач детекції та класифікації звукових об'єктів. Переважна більшість наукових праць з проблем акустичного розпізнавання сигналів зосереджується на використанні або методів спектрального аналізу [4], або нейромережових підходів [5; 6; 7] як автономних рішень.

Водночас дослідження, спрямовані на розв'язання задач детекції та класифікації звукових об'єктів, характеризуються широким застосуванням методів глибинного навчання. Так, у роботі [5] розглянуто систему глибинного навчання для аналізу акустичних сигналів промислового обладнання. Автори здійснюють перетворення звукових сигналів у мел-спектрограми та застосовують нейромережну архітектуру DenseNet-169 для класифікації зображень спектрограм. Система забезпечила точність від 97,17 % до 99,87 % за різних рівнів шумових завад, що підтверджує доцільність підходу, заснованого на повному перетворенні аудіосигналу у візуальне представлення.

Дослідження акустичної детекції безпілотних літальних апаратів представлене в роботі [6], де здійснено порівняння трьох методів вилучення ознак: спектрограми, фільтр-банку (Fbank) та мел-частотних кепстральних коефіцієнтів (MFCC). Автори встановили оптимальну продуктивність за MFCC у середовищах з низьким співвідношенням сигнал/шум, застосовуючи модифіковану модель ResNet10_CBAM. Водночас підхід залишається орієнтованим виключно на нейромережові методи без інтеграції класичного спектрального аналізу.

Комплексний підхід до детекції дронів запропоновано в дослідженні [8], де автори поєднують радіочастотні та акустичні ознаки для підвищення точності класифікації. Система використовує рекурентні нейронні мережі з LSTM-шарами для обробки часово-частотних характеристик, зокрема MFCC та вейвлет-перетворення. Такий підхід демонструє переваги мультимодального аналізу, забезпечуючи стійкість до шумових завад та покращену продуктивність в умовах низького співвідношення сигнал/шум.

Системи детекції дронів на основі акустичних радарів досліджуються в роботі [4], де розглянуто можливості пасивного виявлення цілей без необхідності прямої видимості. Акустичні сенсори виявляються придатними для детекції як керованих вручну, так і автономних дронів, заповнюючи функціональні прогалини в системах захисту повітряного простору.

Виділення недосліджених частин загальної проблеми. Проведений аналіз літературних джерел виявляє обмеженість досліджень акустичного розпізнавання в специфічних умовах, коли система детекції розміщується на рухомій платформі, що сама є джерелом акустичних завад. Переважна більшість існуючих методів розроблена для стаціонарних систем детекції або для сценаріїв, де фонові шуми мають відмінну від цільового сигналу природу походження. Водночас недостатньо вивченими залишаються випадки, коли необхідно розрізняти акустичні сигнали від об'єктів аналогічного типу в умовах, коли один з них створює домінуючий акустичний фон для детекції іншого.

Додатково встановлено брак досліджень, що зосереджуються на виділенні вузьких частотних діапазонів для розв'язання задач акустичної класифікації з подальшим застосуванням компактних векторів ознак. Існуючі роботи переважно використовують широкосмугову обробку сигналу з формуванням спектрограм великої розмірності.

Метою дослідження є розробка та дослідження методу розпізнавання акустичних сигналів, що поєднує спектральний аналіз обмеженого частотного діапазону з нейромережевою класифікацією для детекції цільових об'єктів в умовах низького співвідношення сигнал/шум. Досягнення цієї мети передбачає вирішення таких завдань: обґрунтування вибору оптимального частотного діапазону на основі аналізу гармонічної структури досліджуваних сигналів, розробка методу вилучення компактного набору спектральних ознак з обраного діапазону, синтез архітектури нейронної мережі для класифікації векторів ознак та експериментальна валідація запропонованого підходу.

Виклад основного матеріалу. Запропонований метод реалізує послідовний алгоритм обробки акустичних сигналів, що складається з чотирьох основних етапів (рис. 1).

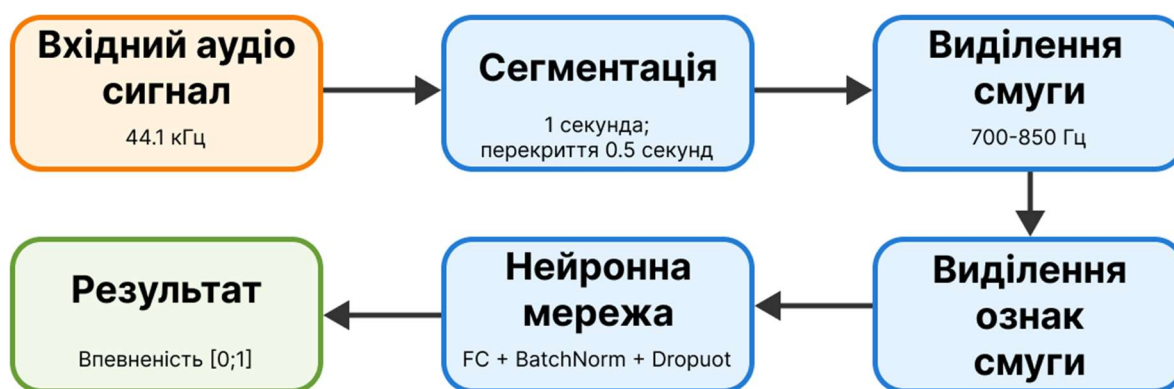


Рис. 1. Схема алгоритму обробки акустичних сигналів

Джерело: розроблено авторами.

На першому етапі здійснюється сегментація вхідного аудіосигналу на фрагменти тривалістю 1 секунда з перекриттям 0,5 секунди відповідно до методики, описаної в [6]. Такі параметри часового вікна забезпечують компроміс між темпоральною роздільністю системи та достатньою статистикою для надійного спектрального аналізу.

Другий етап передбачає фільтрацію сигналу в спеціально визначеному частотному діапазоні. Відфільтрований сигнал містить виключно ту частину спектра, де співвідношення енергії цільового та фонового компонентів є найбільш сприятливим для розрізнення.

На третьому етапі з відфільтрованого сигналу вилучається 21 числова ознака. Ці ознаки формують вектор, що містить інформацію про ключові характеристики сигналу в обраному частотному діапазоні.

Четвертий етап передбачає класифікацію отриманого вектора ознак за допомогою повнозв'язної нейронної мережі, яка видає значення впевненості в діапазоні від нуля до одиниці щодо належності аналізованого фрагмента до класу цільових сигналів. Фінальне бінарне рішення про наявність або відсутності цільового об'єкта приймається шляхом порівняння отриманого значення з попередньо встановленим порогом.

Вибір частотного діапазону для аналізу здійснювався на основі детального спектрального дослідження акустичних характеристик як цільових сигналів, так і фонових завад. Метою було знайти смуга, де цільовий звук є переважаючим над фоновими завадами. Для початкового аналізу було обрано смугу 550-650 Гц, де зосереджена основна енергія обох типів сигналів. Для побудови спектральних характеристик було використано метод

Велча, який забезпечує надійну оцінку спектральної густини потужності шляхом усереднення періодограм сегментів сигналу. Спектральна густина потужності розраховувалася за формулою (1), яка більш детально описана в роботі [9]:

$$\hat{S}_{xx}(f) = \frac{1}{F_s U K} \sum_{i=0}^{K-1} |X_i(f)|^2, \quad (1)$$

де $\hat{S}_{xx}(f)$ – оцінка спектральної густини потужності сигналу;

F_s – частота дискретизації (Гц);

U – нормувальний коефіцієнт, що враховує потужність вікна;

K – кількість сегментів;

$X_i(f)$ – спектр i -го сегмента після віконної обробки.

Результат побудови спектральних характеристик та розрахунків наведено на рисунку 2.

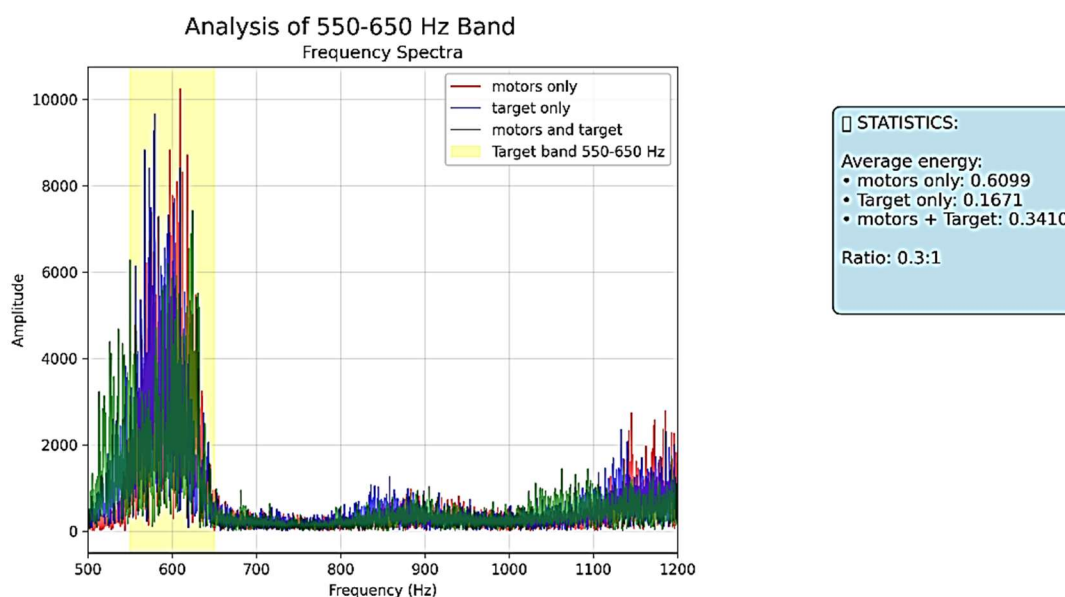


Рис. 2. Спектральний аналіз цільового сигналу та фонових завад у частотному діапазоні 550-1200 Гц

Джерело: розроблено авторами.

Як видно з рис. 2, у діапазоні 550-650 Гц спостерігається значне перевищення рівня фонових завад над цільовим сигналом. Статистичний аналіз показав, що середня енергія фонових шумів (motors only) становить 0,1760, тоді як енергія цільового сигналу (target only) дорівнює лише 0,1671. Комбінований сигнал (motors and target) має енергію 0,3410, що дає співвідношення сигнал/завада приблизно 0,3:1. Такий низький рівень співвідношення значно ускладнює детекцію цільового об'єкта через домінування завадового компонента в аналізованій частотній смузі.

Отримані результати підтвердили необхідність пошуку альтернативних частотних діапазонів із більш сприятливим енергетичним балансом між цільовим сигналом та фоновими завадами.

Цільовий об'єкт дослідження в ідеальних умовах має основну частоту звуку близько 100 Гц. Враховуючи нестабільність джерела звуку та варіації режимів роботи, було прийнято рішення розширити аналізований діапазон до 150 Гц для забезпечення покриття можливих флуктуацій основної частоти та її гармонік.

Для систематичного пошуку діапазону з оптимальним співвідношенням сигнал/завада було розроблено алгоритм, який здійснює сканування частотної області від 600 Гц до 1300 Гц з кроком 50 Гц, аналізуючи смуги шириною 150 Гц. Вибір початкової частоти

600 Гц зумовлено тим, що в цій області зосереджена основна енергія як шумових завад, так і гармонік цільового сигналу. Верхня межа 1300 Гц відповідає зоні, де енергія обох типів сигналів поступово спадає.

Алгоритм використовує той самий підхід розрахунку спектральної густини потужності за методом Велча, описаний вище у формулі (1). Для кожної смуги обчислювалося співвідношення енергії цільового сигналу до енергії фонових завад згідно з формулою (2):

$$R = \frac{E_{\text{target}}}{E_{\text{noise}}}, \quad (2)$$

де R – співвідношення сигнал/завада;

E_{target} – енергія цільового сигналу в аналізованій смузі;

E_{noise} – енергія фонових завад в тій самій смузі.

Результат систематичного аналізу наведено на рис. 3.

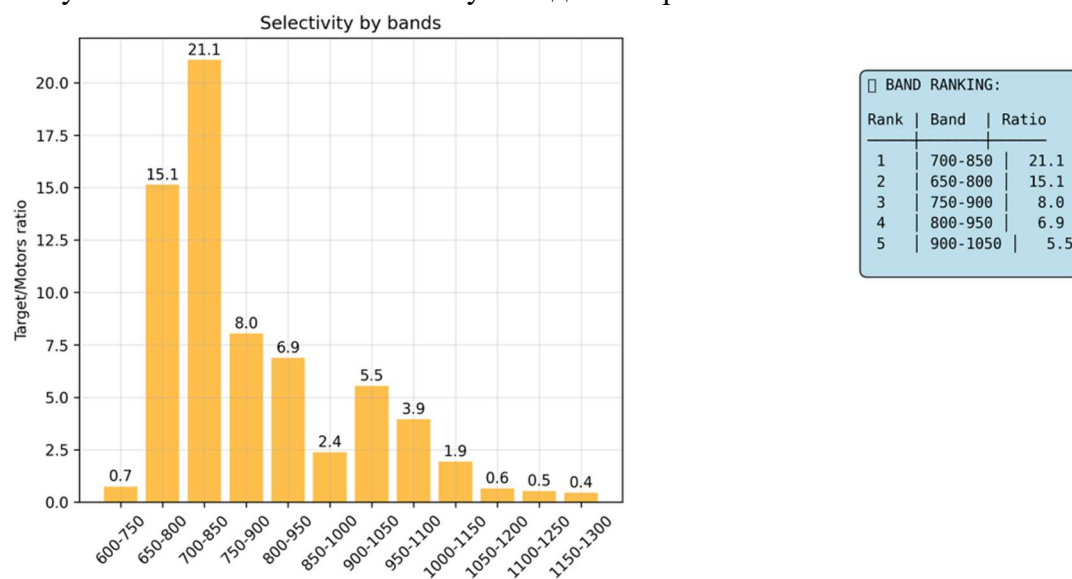


Рис. 3. Результати систематичного пошуку оптимального частотного діапазону
Джерело: розроблено авторами.

Як видно з рис. 3, аналіз виявив оптимальну смугу 700-850 Гц з найвищим співвідношенням сигнал/завада 21,1:1. Друга за ефективністю смуга 650-800 Гц показала співвідношення 15,1:1. Значне перевищення цих показників порівняно з початково аналізованою смугою 550-650 Гц (співвідношення 0,3:1) підтверджує доцільність використання діапазону 700-850 Гц для подальшого аналізу та видалення ознак.

Для підтвердження логічності виявленого оптимального діапазону було проведено детальний аналіз спектральних характеристик окремих компонентів акустичного середовища. Аналіз здійснювався на основі чотирьох фрагментів тривалістю 1 секунда для кожного з наступних варіантів звукового наповнення: виключно фонові завади (звук бортової системи носія мікрофонів), виключно цільовий сигнал, а також штучно створене поєднання цих двох об'єктів шляхом відтворення звуку цілі під час роботи бортової системи носія.

Результати аналізу спектральних характеристик цільового сигналу наведено на рис. 4.

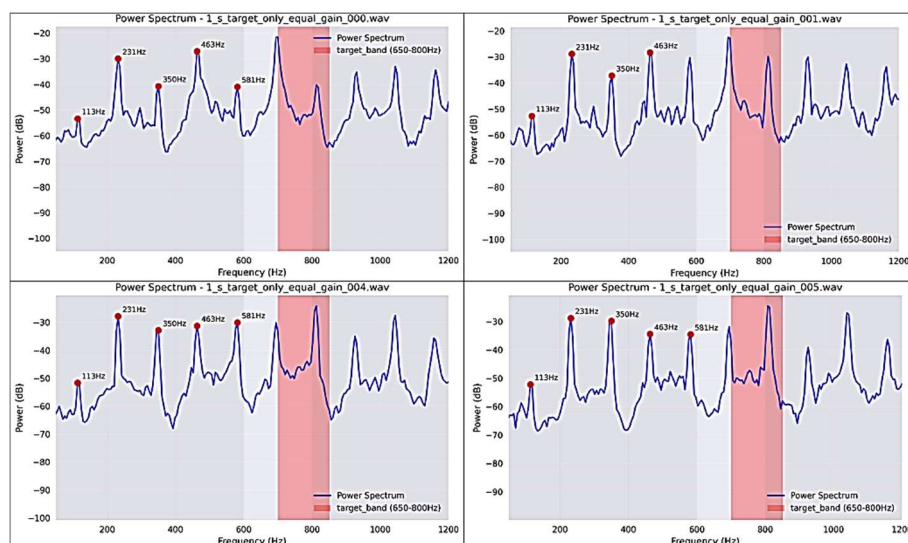


Рис. 4. Спектральні характеристики цільового сигналу з виділенням діапазону 700-850 Гц
Джерело: розроблено авторами.

Як видно з отриманих спектрограм, виділена смуга 700-850 Гц захоплює спектральні компоненти, що відповідають частині шостої та повністю сьомій гармонічній складовій цільового об'єкта. Ці гармонічні компоненти проявляються у вигляді характерних спектральних піків із відповідними частотними розширеннями, що обумовлено нестабільністю режиму роботи джерела.

Аналіз спектральних характеристик фонових завад представлено на рис. 5.

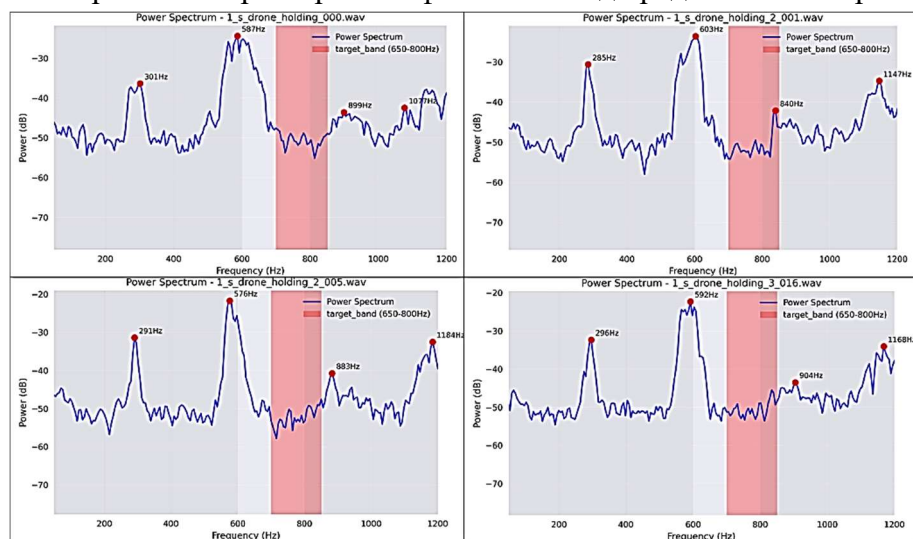


Рис. 5. Спектральні характеристики фонових завад із позначенням робочого діапазону
Джерело: розроблено авторами.

Спектральний аналіз показує, що обрана смуга 700-850 Гц потрапляє на область спектрального спаду між другою та третьою гармоніками бортової системи носія. Це створює сприятливі умови для виділення цільового сигналу, оскільки енергія завадового компонента в цьому діапазоні є мінімальною.

Результати аналізу штучно створених умов, де присутні обидва типи акустичних сигналів, наведено на рис. 6.

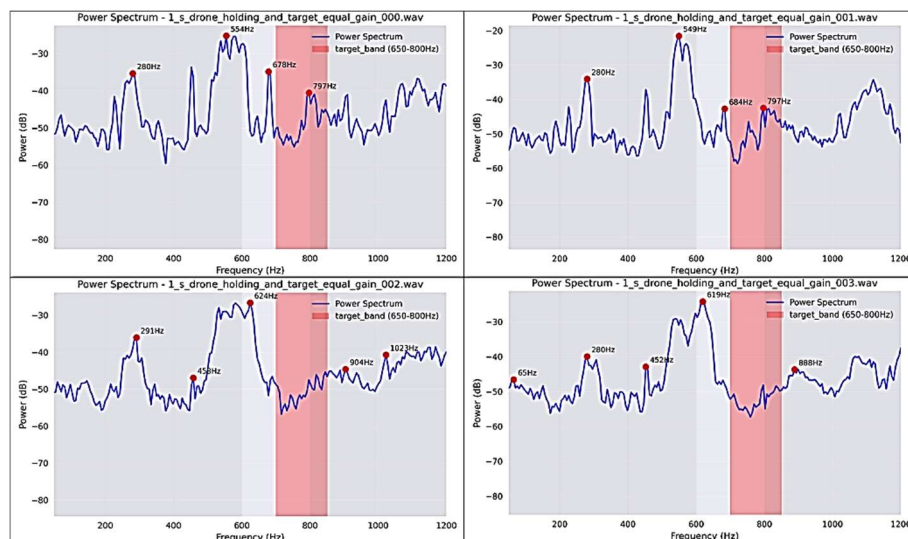


Рис. 6. Спектральні характеристики комбінованого сигналу: ціль і завади

Джерело: розроблено авторами.

У спектральній області між другою та третьою гармоніками завадового сигналу чітко проглядаються локальні піки, які є проявом цільового сигналу. Ці піки зберігають свою структуру навіть за наявності домінуючих фонових завад, що підтверджує можливість їх виділення та аналізу.

Таким чином, теоретичний аналіз підтверджує, що систематичний алгоритм пошуку оптимального частотного діапазону виявив смугу 700-850 Гц не випадково. Фізичною основою такого вибору є збіг частотного розташування високих гармонік цільового об'єкта з областю мінімальної енергії фонових завад від бортової системи носія, що створює оптимальні умови для розрізнення.

Після визначення оптимального частотного діапазону 700-850 Гц наступним етапом методу є видалення інформативних ознак із відфільтрованого сигналу. Сформований набір ознак включає три групи параметрів:

1. Основні спектральні ознаки:

- RMS енергія сигналу – середньоквадратичне значення амплітуди;
- пікова амплітуда – максимальне абсолютне значення амплітуди в аналізованому фрагменті;
- частота перетинів нуля – кількість переходів сигналу через нульовий рівень, нормована на довжину фрагмента;
- амплітуда спектрального піку – максимальне значення модуля Фур'є-перетворення в аналізованому діапазоні;
- позиція піку – частота максимуму, нормована відносно центру діапазону (775 Гц);
- ширина піку – частотний діапазон на рівні половини максимальної амплітуди, нормований на ширину смуги;
- статистичні моменти спектру: середнє значення, дисперсія чи стандартне відхилення, асиметрія та ексцес.

2. MFCC, які забезпечують представлення спектральної оболонки сигналу з урахуванням особливостей людського слухового сприйняття. Для діапазону 700-850 Гц обчислюється вісім параметрів налаштуванням мел-фільтрів на вузький частотний діапазон. Для кожного фрагмента використовується усереднене за часом значення кожного коефіцієнта.

3. Додаткові характеристики:

- гармонічна чистота – відношення амплітуди спектрального піку до сумарної енергії спектру в смузі;

- спектральна різкість – середня крутизна градієнтів навколо спектрального піку;
- часова когерентність – автокореляція огинаючої сигналу між сусідніми часовими вікнами.

Розрахунок основних спектральних ознак здійснювався за методиками, описаними в [10]. MFCC коефіцієнти з першого по восьмий обчислювалися згідно з підходом, представленим в [11]. Додаткові характеристики розраховувалися на основі методів, викладених у [12] для гармонічної чистоти та спектральної різкості, а також методів, наведених у [13], для часової когерентності.

Сформований вектор ознак забезпечує комплексну характеристику акустичного сигналу в обраному частотному діапазоні. Він використовується як вхідний тензор для нейронної мережі класифікації.

Враховуючи компактність вхідних даних та необхідність функціонування на автономних системах з обмеженими обчислювальними ресурсами, було розроблено повнозв'язну нейронну мережу (Fully Connected Neural Networks – FCNN) на основі лінійних шарів. Архітектура мережі представлена на рисунку 7.

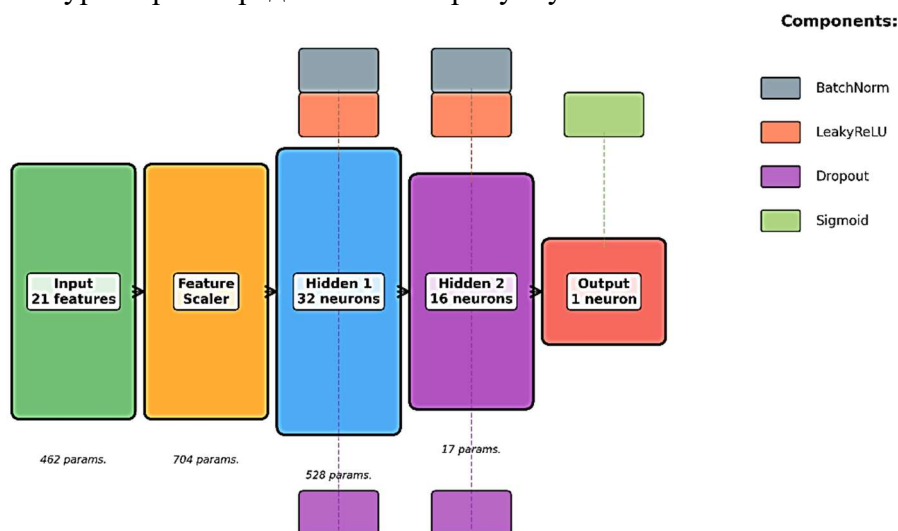


Рис. 7. Архітектура повнозв'язної нейронної мережі для класифікації векторів ознак
Джерело: розроблено авторами.

Запропонована нейромережна модель складається з послідовності лінійних шарів із регуляризацією та нормалізацією:

- Вхідний шар: приймає 21-вимірний вектор ознак.
- Попередня нормалізація: Feature Scaler для стандартизації вхідних даних.
- Перший прихований шар: лінійний шар з 32 нейронами, функція активації LeakyReLU, BatchNorm та Dropout.
- Другий прихований шар: лінійний шар з 16 нейронами з аналогічною конфігурацією.
- Вихідний шар: лінійний шар з 1 нейроном та сигмоїдною функцією активації для бінарної класифікації.

Загальна кількість параметрів моделі становить 1710. Реалізація виконана з використанням бібліотеки PyTorch 2.0 у середовищі Python 3.10. Для ініціалізації ваг застосовувалася методика Kaiming для LeakyReLU активації.

Методологія експериментального дослідження. Експериментальне дослідження проводилося з використанням акустичних записів, отриманих в умовах реального функціонування системи детекції на рухомій платформі. Цільові акустичні сигнали характе-

ризувалися основною частотою близько 100 Гц, що відповідає роботі чотирициліндрового двотактного двигуна досліджуваного об'єкта, з можливими варіаціями залежно від режиму роботи. Фонові завади створювалися квадрокоптером-носієм системи детекції, що генерував інтенсивні акустичні сигнали в широкому частотному діапазоні з домінуючими компонентами в низькочастотній області.

Всі акустичні дані записувалися з частотою дискретизації 44,1 кГц, що забезпечувало аналіз сигналів у діапазоні до 22 кГц.

Навчання нейронної мережі проводилося протягом 300 епох з механізмом раннього зупинення (early stopping), який спрацював на 204 епосі для запобігання перенавчанню. Розподіл датасету здійснювався у співвідношенні 60 % для навчання, 20 % для валідації та 20 % для тестування. Навчальна вибірка містила 650 позитивних фрагментів (наявність цільового сигналу) та 850 негативних фрагментів (відсутність цільового сигналу).

Параметри навчання для обраної архітектури нейронної мережі включали: швидкість навчання (learning rate) встановлена на рівні 0,001 для забезпечення стабільної конвергенції, розмір пакета (batch size) – 32 елементи, що є допустимим через невеликий розмір вхідного тензора.

Оцінювання продуктивності моделі здійснювалося за стандартними метриками для задач бінарної класифікації [14]:

- Accuracy – загальна точність класифікації
- Precision – точність позитивного класу
- Recall – повнота (чутливість)
- F1-Score – гармонійне середнє Precision та Recall.

Основною метрикою для оцінювання якості моделі була обрана Accuracy на тестовому наборі як найбільш репрезентативна для збалансованого датасету.

Фінальні результати на тестовій вибірці:

- Accuracy: 90,85 %;
- Precision: 86,26 %;
- Recall: 93,39 %;
- F1-Score: 89,68 %.

Процес навчання та динаміка зміни метрик представлені на рис. 8.

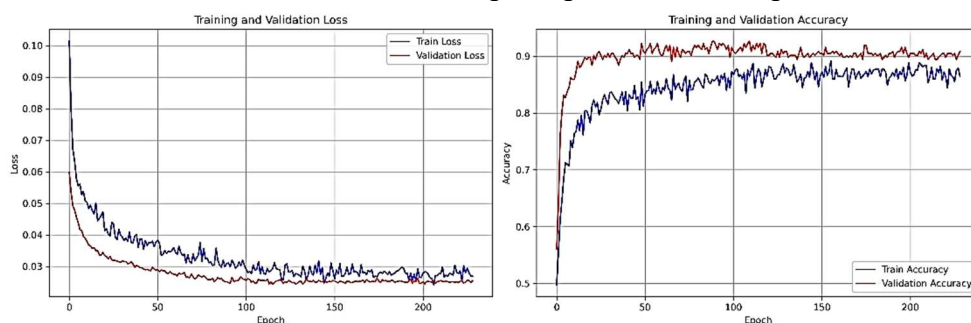


Рис. 8. Динаміка зміни метрик точності під час навчання моделі

Джерело: розроблено авторами.

Отримані результати демонструють здатність запропонованого методу до детекції цільових сигналів навіть в умовах низького співвідношення сигнал/завада.

Аналіз результатів дослідження. Оцінка ефективності запропонованого методу для задачі бінарної класифікації в умовах низького співвідношення сигнал/завада проводилась на тестовому наборі, що містив 650 позитивних та 850 негативних зразків.

Показник Ассурасу (точність), що відображає загальну частку правильно класифікованих фрагментів серед усіх тестових зразків, склав 90,85 %. Цей результат вказує на здатність розробленої моделі нейронної мережі до ефективної класифікації обох класів.

Precision (влучність), що характеризує частку дійсно позитивних зразків серед тих, що були класифіковані як позитивні, виявився на рівні 86,26 %.

Recall (повнота), що показує частку правильно виявлених позитивних зразків від їх загальної кількості в тестовій вибірці, дорівнює 93,39 %. Високе значення цієї метрики свідчить про низький рівень пропусків цільових сигналів – менше 7 % дійсно позитивних зразків залишилися невиявленими.

F1-Score, що являє собою гармонічне середнє між Precision та Recall, забезпечуючи збалансовану оцінку ефективності моделі для позитивного класу, становить 89,68 %. Близькість цього значення до показників Precision та Recall підтверджує стабільність роботи класифікатора.

Висновки. У ході дослідження розроблено метод розпізнавання акустичних сигналів для детекції цільових об'єктів в умовах, коли їх енергія нижча за енергію фонових завад від власної платформи детекції. Метод ґрунтується на послідовному виконанні чотирьох етапів: сегментація аудіосигналу на фрагменти тривалістю 1 секунда з перекриттям 0,5 секунди, фільтрація сигналу в оптимальному частотному діапазоні, вилучення 21 ознаки сигналу та нейромережева класифікація.

Визначення оптимального частотного діапазону здійснювалось шляхом систематичного сканування смуг шириною 150 Гц у діапазоні 600-1300 Гц з кроком 50 Гц. Для досліджуваного випадку обраний діапазон 700-850 Гц забезпечив співвідношення сигнал/завада 21,1:1 завдяки збігу з шостою та сьомою гармоніками цільового об'єкта – звуку чотирициліндрового двотактного двигуна – та розташуванню між другою і третьою гармоніками завадового сигналу квадрокоптера-носія. Запропонований алгоритм пошуку оптимального частотного діапазону є універсальним і може бути адаптований для інших комбінацій цільових об'єктів та джерел завад.

Експериментальна валідація на датасеті з 1500 зразків показала точність 90,85 %, що підтверджує працездатність методу в умовах низького співвідношення сигнал/завада. Високе значення Recall (93,39 %) свідчить про мінімізацію пропусків цільових об'єктів.

Заява про використання генеративного ШІ та технологій на основі ШІ в процесі написання тексту статті

Під час написання статті автори не використовували генеративний ШІ. Використовувалися технології ШІ Gemini для допомоги в пошуку існуючих рішень проблеми. Після використання цього інструменту автори переглянули та відредагували зміст за потреби і взяли на себе повну відповідальність за зміст публікації.

Список використаних джерел

1. Famili, A., Stavrou, A., Wang, H., Park, J.-M., & Gerdes, R. (2024). Securing Your Airspace: Detection of Drones Trespassing Protected Areas. *Sensors*, 24(2), 14. <https://doi.org/10.3390/s24020579>.
2. Wilshin, S. (2023). Seeing with sound; surface detection and avoidance by sensing self-generated noise. *International Journal of Micro Air Vehicles*, 15(1), 1-12. <https://doi.org/10.1177/17568293221148377>.
3. Vashishtha, G., & Kumar, R. (2023). Editorial: Advances in noise reduction and feature extraction of acoustic signal. *Frontiers in Physics*, 11, 1287956. <https://doi.org/10.3389/fphy.2023.1287956>.
4. Zakaria, N. A. Z., Mohd Rozaimi, F. N. S., Enche Ab Rahim, S. A., & Khan, Z. I. (2024). Signal analysis for drone detection and characterization using acoustic radar. *Journal of Physics Conference Series*, 2922(1), 012003. <https://doi.org/10.1088/1742-6596/2922/1/012003>.

5. Yurdakul, M., & Tasdemir, S. (2023). *Acoustic signal analysis with deep neural network for detecting fault diagnosis in industrial machines. International Journal of Applied Engineering Research*, 12(12), 3461-3469. <https://doi.org/10.48550/arXiv.2312.01062>.
6. Liu, Z., Fan, K., Chen, Y., Xiong, L., Ye, J., Fan, A., & Zhang, H. (2025). Deep learning-based acoustic recognition of UAVs in complex environments. *Drones*, 9(6), 389. <https://doi.org/10.3390/drones9060389>.
7. Казимир, В., Роговенко, А., & Карась, О. (2024). Використання штучних нейронних мереж для аудіо класифікації на мінікомп'ютерній платформі. *Технічні науки та технології*, (4(38)), 134-149. [https://doi.org/10.25140/2411-5363-2024-4\(38\)-134-149](https://doi.org/10.25140/2411-5363-2024-4(38)-134-149).
8. Frid, A., Ben-Shimol, Y., Manor, E., & Greenberg, S. (2024). Drones detection using a fusion of RF and acoustic features and deep neural networks. *Sensors*, 24 (8), 2427. <https://doi.org/10.3390/s24082427>.
9. Same, M. H., Gandubert, G., Gleeton, G., Ivanov, P., & Landry, R. (2020). Simplified welch algorithm for spectrum monitoring. *Applied Sciences*, 11(1), 86. <https://doi.org/10.3390/app11010086>.
10. Madanian, S., Chen, T., Adeleye, O., Templeton, J. M., Poellabauer, C., Parry, D., Schneider, S. L. (2023). Speech emotion recognition using machine learning – A systematic review. *Intelligent Systems with Applications*, 20, 200266. <https://www.sciencedirect.com/science/article/pii/S2667305323000911>.
11. Mukkamala, S. N. V. J., & Radhika, Y. (2020). A review: Music feature extraction from an audio signal. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(2), 974-976.
12. Prabhakar, S. K., Rajaguru, H., & Won, D.-O. (2024). Coherent feature extraction with swarm intelligence based hybrid adaboost weighted ELM classification for snoring sound classification. *Diagnostics*, 14(17), 1857. <https://doi.org/10.3390/diagnostics14171857>.
13. Lee, J., & Oxenham, A. J. (2025). Temporal coherence effects on voice attribution in multi-speaker stream segregation. *JASA Express Letters*, 5(5). <https://doi.org/10.1121/10.0036672>.
14. Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-56706-x>.

References

1. Famili, A., Stavrou, A., Wang, H., Park, J.-M., & Gerdes, R. (2024). Securing Your Airspace: Detection of Drones Trespassing Protected Areas. *Sensors*, 24(2), 14. <https://doi.org/10.3390/s24020579>.
2. Wilshin, S. (2023). Seeing with sound; surface detection and avoidance by sensing self-generated noise. *International Journal of Micro Air Vehicles*, 15(1), 1-12. <https://doi.org/10.1177/17568293221148377>.
3. Vashishtha, G., & Kumar, R. (2023). Editorial: Advances in noise reduction and feature extraction of acoustic signal. *Frontiers in Physics*, 11, 1287956. <https://doi.org/10.3389/fphy.2023.1287956>.
4. Zakaria, N. A. Z., Mohd Rozaimi, F. N. S., Enche Ab Rahim, S. A., & Khan, Z. I. (2024). Signal analysis for drone detection and characterization using acoustic radar. *Journal of Physics Conference Series*, 2922(1), 012003. <https://doi.org/10.1088/1742-6596/2922/1/012003>.
5. Yurdakul, M., Tasdemir, S. (2023). Acoustic Signal Analysis with Deep Neural Network for Detecting Fault Diagnosis in Industrial Machines. *International Journal of Applied Engineering Research*, 12(12), 3461-3469. <https://doi.org/10.48550/arXiv.2312.01062>.
6. Liu, Z., Fan, K., Chen, Y., Xiong, L., Ye, J., Fan, A., Zhang, H. (2025). Deep Learning-Based Acoustic Recognition of UAVs in Complex Environments. *Drones*, 9(6), 389. <https://doi.org/10.3390/drones9060389>.
7. Kazymyr, V., Rohovenko, A., & Karas, O. (2024). Vykorystannia shtuchnykh neironnykh merezh dlia audio klasyfikatsii na minikompiuternii platformi. [The use of artificial neural networks for audio classification on a minicomputer platform.], *Tekhnichni nauky ta tekhnolohii - Technical Sciences and Technologies*, (4(38)), 134-149. [https://doi.org/10.25140/2411-5363-2024-4\(38\)-134-149](https://doi.org/10.25140/2411-5363-2024-4(38)-134-149).
8. Frid, A., Ben-Shimol, Y., Manor, E., Greenberg, S. (2024). Drones Detection Using a Fusion of RF and Acoustic Features and Deep Neural Networks. *Sensors*, 24(8), 2427. <https://doi.org/10.3390/s24082427>.
9. Same, M. H., Gandubert, G., Gleeton, G., Ivanov, P., Landry Jr., R. (2021). Simplified Welch Algorithm for Spectrum Monitoring. *Applied Sciences*, 11(1), 86. <https://doi.org/10.3390/app11010086>.
10. Madanian, S., Chen, T., Adeleye, O., Templeton, J. M., Poellabauer, C., Parry, D., Schneider, S. L. (2023). Speech emotion recognition using machine learning – A systematic review. *Intelligent Systems with Applications*, 20, 200266. <https://www.sciencedirect.com/science/article/pii/S2667305323000911>.

11. Mukkamala, S. N. V. Jitendra, & Radhika, Y. (2020). A Review: Music Feature Extraction from an Audio Signal. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(2), 974-976.
12. Prabhakar, S. K., Rajaguru, H., & Won, D. O. (2024). Coherent Feature Extraction with Swarm Intelligence Based Hybrid Adaboost Weighted ELM Classification for Snoring Sound Classification. *Diagnostics*, 14(17), 1857. <https://doi.org/10.3390/diagnostics14171857>.
13. Lee, J., & Oxenham, A. J. (2025). Temporal coherence effects on voice attribution in multi-speaker stream segregation. *JASA Express Letters*, 5(5), 054401. <https://doi.org/10.1121/10.0036672>.
14. Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, 14, 6086. <https://doi.org/10.1038/s41598-024-56706-x>.

Отримано 17.09.2025

UDC 004.2

Volodymyr Kazymyr¹, Andrii Rohovenko², Konstantin Tkachenko³

¹Doctor of Sciences, Professor, Professor of the Department of Information and Computer Systems
Chernihiv Polytechnic National University (Chernihiv, Ukraine)

E-mail: vvkazymyr@gmail.com. **ORCID:** <https://orcid.org/0000-0001-8163-1119>. **ResearcherID:** Q-2925-2016

²PhD in Technical Sciences, Associate Professor of the Department of Information and Computer Systems
Chernihiv Polytechnic National University (Chernihiv, Ukraine)

E-mail: arogoenko@gmail.com. **ORCID:** <https://orcid.org/0000-0003-4594-5692>. **ResearcherID:** G-3926-2014

³Student of the Department of Information and Computer Systems
Chernihiv Polytechnic National University (Chernihiv, Ukraine)

E-mail: kostiantintkachenko@gmail.com. **ORCID:** <https://orcid.org/0009-0000-0927-0080>

METHOD OF ACOUSTIC SIGNAL RECOGNITION BASED ON SPECTRAL ANALYSIS USING NEURAL NETWORKS

The paper proposes a method for recognizing acoustic signals based on spectral analysis and neural networks for detecting target objects under negative signal-to-noise ratio conditions. A specific task of acoustic identification is investigated, where the target signal has a lower power level compared to background noise from a source of similar nature. Based on a preliminary analysis of the harmonic structure of the studied signals, the frequency range of 700-850 Hz was determined as optimal for distinguishing target and background components.

A custom analyzer was developed to extract 21 spectral features from the selected frequency range using time windows of 1 second with a 0.5-second overlap. An architecture of a fully connected neural network with two hidden layers was proposed for classifying the feature vectors. Experimental validation of the method demonstrated a detection accuracy of 88 % with a false alarm rate of 4 %.

The method provides a compact representation of acoustic information and computational efficiency by focusing on a predefined informative frequency range instead of processing the full signal spectrum.

Keywords: acoustic recognition; spectral analysis; neural networks; low SNR; frequency selection; feature vectors.

Fig.: 8. Ref.: 14.