

ДОСЛІДЖЕННЯ РОЗПІЗНАВАННЯ ДРОНІВ ЗА ЇХ АКУСТИЧНИМ ВИПРОМІНЮВАННЯМ З ВИКОРИСТАННЯМ ПОВНОЗВ'ЯЗНИХ НЕЙРОННИХ МЕРЕЖ

Вступ

У сучасних умовах широкого розповсюдження безпілотних літальних апаратів (БПЛА), або дронів, питання їх своєчасного виявлення та розпізнавання набуває особливої актуальності. Дрони використовуються як у цивільних, так і у військових цілях, а в окремих випадках становлять загрозу безпеці критичної інфраструктури та громадської безпеки [1–3]. Традиційно для виявлення дронів застосовуються радіолокаційні та оптичні методи [4–6]. Проте ці підходи мають низку суттєвих обмежень, які знижують їхню ефективність в умовах реального середовища.

Радіолокаційні системи виявлення демонструють високу точність, проте малорозмірні дрони з малими ефективними поверхнями розсіювання часто залишаються поза межами їхньої чутливості, особливо у складних рельєфних або міських умовах [4, 5]. Крім того, дрони можуть використовувати матеріали та конструкції, що знижують радіолокаційну помітність. Оптичні методи, зокрема відеоспостереження та інфрачервона зйомка, залежать від погодних умов, освітлення та наявності прямої видимості, що також значно обмежує їх застосування у нічний час або за несприятливої погоди [6].

На цьому фоні акустичні методи виявлення та розпізнавання дронів виділяються рядом переваг. Акустичне випромінювання, що виникає під час роботи електродвигунів та обертання гвинтів, є унікальним для кожного типу дрона і може бути виявлене навіть при відсутності прямої видимості. Такі сигнали менш залежні від погодних умов і можуть застосовуватись у міських або лісистих умовах, де інші методи втрачають ефективність [7–9].

З огляду на складність та варіативність акустичних сигналів, особливо важливою є розробка ефективних методів їх автоматичного аналізу, які можуть ефективно виділяти та аналізувати загальні ознаки таких сигналів в умовах дії зовнішніх шумів [10]. В цьому контексті нейронні мережі, зокрема повнозв'язні (fully connected) моделі, демонструють високу здатність до розпізнавання складних шаблонів у спектральних представленнях звуку. Тому дослідження застосування повнозв'язних нейронних мереж для задач розпізнавання дронів за їх акустичним випромінюванням є актуальним напрямом сучасної науково-технічної діяльності.

Метою цього дослідження були: аналіз ефективності використання повнозв'язних нейронних мереж для виявлення дронів на різних відстанях та оцінка апаратної складності реалізації алгоритму розпізнавання на сучасних мікропроцесорних платформах.

Аналіз спектральних особливостей власного акустичного випромінювання дронів

Існує багато моделей дронів за конструктивними особливостями. Але всі вони мають зазвичай від 1 до 8 гвинтів, які саме і створюють акустичне випромінювання.

При побудові методів ефективного виявлення дронів за їх власним акустичним випромінюванням аналізуються часові або спектральні особливості такого випромінювання. З точки зору конструкції дронів їх акустичні випромінювання представляють собою періодичні коливання, пов'язані з обертанням пропелерів та роботою двигунів. Наряду з періодичними складовими спостерігаються імпульсні складові, зумовлені турбулентними потоками повітря та взаємодією конструктивних елементів дрона з навколишнім середовищем [6, 11].

Для аналізу сигналів у часовій області, наприклад, застосовуються рекурентні нейронні мережі (RNN), які дозволяють ефективно розпізнавати акустичні сигнали, навіть в умовах

впливу шумів, але вимагають для практичної реалізації значних обчислювальних ресурсів та великих об'ємів пам'яті [12]. Такі вимоги пов'язані з великою кількістю часових відліків акустичних сигналів, що значно навантажує обчислювальне ядро.

У спектральній площині присутні чітко виражені основні частоти, що відповідають обертанню двигунів, а також їх вищі гармоніки. Акустичне випромінювання дронів у спектральній площині можна охарактеризувати, як широкосмугове, що охоплює частоти від десятків герць до 8–10 кГц [6–8], хоча найчастіше цей діапазон обмежений частотами до 2–5 кГц. Приклади результатів спектрального аналізу записів власних акустичних випромінювань дронів DJI Phantom 3 та Syma X5SW, що мають по 4 гвинта, на відстанях 8 та 40 м наведено на рис. 1.

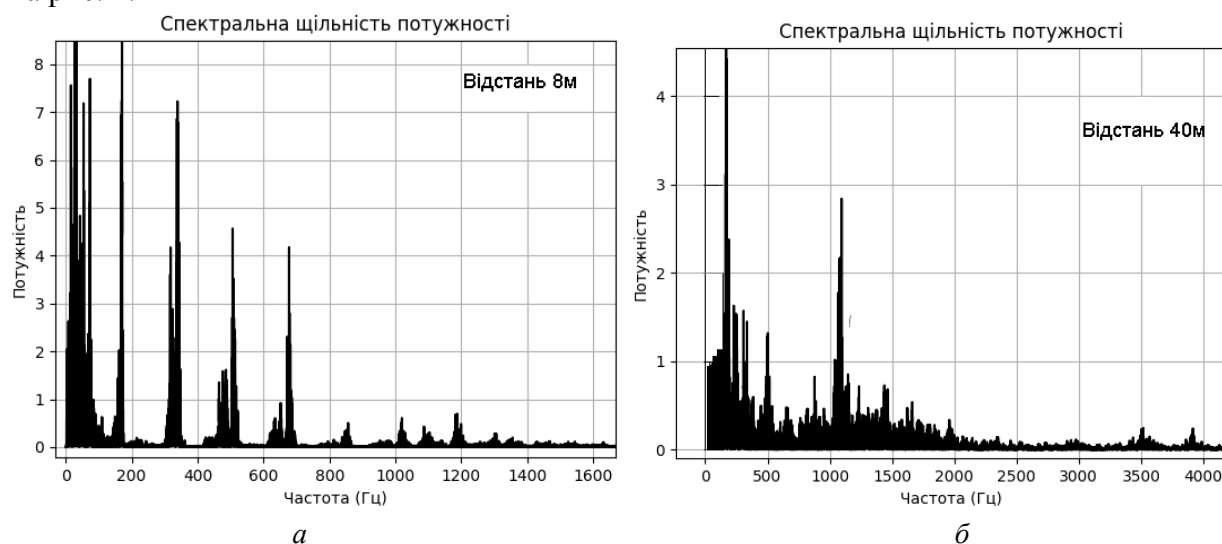


Рис. 1. Приклади спектральної щільності потужності власного акустичного випромінювання дронів на відстанях 8 та 40 м

Аналіз прикладів на рис. 1 показує, що коли дрон висить на постійній висоті і гвинти обертаються не дуже швидко, то спектральна щільність потужності обмежується частотою 1,6 кГц. При польоті на великих швидкостях частота обертання двигунів проаналізованих дронів збільшується і максимальна частоти гармонік досягають 8 кГц. На невеликих відстанях (8 м) відношення корисний сигнал/зовнішні шуми достатньо високе і у спектрі можна побачити порядку 9–10 гармонік основної частоти 185 Гц. Із зростанням відстані до 40 м відношення сигнал/шум значно зменшується, тобто рівень шумових спектральних компонентів стає близьким до рівню гармонічних складових. Кількість гармонічних складових, що суттєво перевищують рівень шумів зменшується 6–7. Таке зменшення можна пояснити більшим затуханням високочастотних складових спектра у атмосфері у порівнянні з низькочастотними. Отже власні дослідження спектрального складу та особливостей спектрів акустичних випромінювань дронів співпадають з проаналізованими науковими роботами.

Перехід у спектральну площину дозволяє значно зменшити об'єм даних, що треба обробляти, що особливо актуально для подальшого використання розпізнавання цієї інформації з використанням нейронних мереж. При реалізації такого переходу найпоширенішими та найефективнішими є перетворення Фур'є та Вейвлет перетворення [13]. Вейвлет перетворення забезпечує ефективний аналіз у складних заводових обставинах, дуже ефективно при аналізі нестационарних сигналів, дозволяє робити одночасно аналіз у частотній та часовій площинах. Однак складність обирання базових функцій розкладання, велика кількість обчислень значно обмежують практичне використання. Тому для сигналів з періодичним характером більш якісніше використання перетворення Фур'є. Після виконання перетворення Фур'є кількість спектральних компонентів у діапазоні частот від 0 Гц до 5 кГц залишається великою для безпосередньої обробки. Тому на практиці існують методи вторинної обробки

спектрів, що дозволяють суттєво зменшити кількість значень у інформаційному векторі, що описує спектральні характеристики акустичного сигналу.

Найефективнішим методом формування вектора ознак за результатами обчислення спектральної щільності потужності є метод Mel-Frequency Cepstral Coefficients (MFCC) [14, 15]. В основу цього метода покладено модель чутливості вуха людини, що сприймає звуки нелінійно у частотній області. Обробка сигналу починається з розбиття вхідного сигналу $x(n)$ на короткі сегменти довжиною N із перекриттям M , що дорівнює 25–50 % ширини вікна. Для аналізу мовних сигналів довжина вікна складає 20–40 мс, але з точки зору накопичення енергії корисного слабкого сигналу довжину вікна треба обирати значно більшою. З метою зменшення спектральних витіків виконується згладження одним із вікон $w(n)$: Ганна, Хемінга і т.і. Далі обчислюється швидко перетворення Фур'є (ШПФ) і спектр потужності $P(k)$:

$$P(k) = \left| \sum_{n=0}^{N-1} x(n) w(n) e^{-j2\pi kn/N} \right|^2, k = 0, 1, \dots, N-1 \quad (1)$$

Далі обчислюється шкала частот Mel, що відповідає частотній характеристиці вуха людини:

$$f_{mel} = 2595 \cdot \log_{10} \left(1 + \frac{f}{1000} \right) \quad (2)$$

Після цього задається набір трикутних фільтрів, рівномірно розташованих у шкалі Mel. Вихід кожного фільтра M обчислюється як зважена сума потужності спектральних складових:

$$S_m = \sum_{k=f_{m-1}}^{f_{m+1}} P(k) H_m(k) \quad (3)$$

де $H_m(k)$ – коефіцієнти трикутного фільтра, а f_{m-1} і f_{m+1} – межі фільтра.

Далі виконується обчислення кепстральних коефіцієнтів. До отриманих значень S_m застосовується дискретне косинусне перетворення (DCT), що дозволяє зменшити кореляцію між коефіцієнтами:

$$C_n = \sum_{m=1}^M \log S_m \cos \left[\frac{\pi n (m-0.5)}{M} \right], n = 1, 2, \dots, L, \quad (4)$$

де L – кількість коефіцієнтів MFCC.

За результатами досліджень різних вчених кількість коефіцієнтів, що потрібна для класифікації типу акустичного сигналу значно відрізняється. Так для аналізу мовних сигналів зазначається, що потрібно 8–14 коефіцієнтів, в інших випадках обирають до 20 коефіцієнтів. Однак для задач виявлення дронів дослідники не обґрунтували кількості коефіцієнтів MFCC, яка б була оптимальною з точки зору максимізації імовірності вірного виявлення та мінімізації похибкового прийняття інших звуків, як звуків дронів. Також для розпізнавання акустичних випромінювань дронів на фоні шумів важливим є накопичення енергії сигналу при прийомі цього сигналу з великих відстаней. Тому можна зробити припущення, що часові вікна, що обробляються, треба значно збільшувати, але таке збільшення повинно бути допустимим з точки зору часових затримок при обробці сигналів, а також не потребувати занадто великих вимог до апаратної платформи при практичній реалізації.

Розробка алгоритму досліджень та нейронної мережі

Відомі Інтернет платформи, такі як Kaggle, Roboflow надають для досліджень готові датасети із значеннями MFCC коефіцієнтів для різних моделей дронів та напрямів руху, що значно спрощує процес навчання нейронних мереж [16, 17]. Але суттєвий недолік таких датасетів – це відсутність інформації о тривалостях часового вікна, вікна згладження, відстанях між мікрофоном та дроном, частоти дискретизації та параметрів мікрофонів. Все це не

дозволяє точно визначити вплив цих факторів на результат навчання та подальшого розпізнавання. Для вирішення виявлених вище та проаналізованих проблем було вирішено створити власну базу даних записів акустичних випромінювань дронів DJI Phantom 3 та Syma X5SW при польотах на різних відстанях в межах від 5 до 100 м. Для запису акустичного випромінювання цих дронів використовувались: конденсаторний мікрофон Superlux ECM-999 та зовнішня звукова карта U-Phoria UM2. Частота дискретизації була обрана 44100 Гц, а формат зберігання аудіоданих – wav, щоб уникнути втрат інформації. Далі була створена база даних обчислених коефіцієнтів MFCC, що відповідають кількості відліків ШПФ від 2048 до 16384, а в часовій області вікнам від 46 до 372 мс. Для кожного часового вікна обчислюється 32 MFCC коефіцієнта за формулами (1) – (4). На вхід нейронної мережі подається 8, 12, 16, 20, 24, 28 або 32 коефіцієнти, що забезпечує пошук їх оптимальної кількості при тестуванні нейронної мережі.

При навчанні нейронних мереж результат навчання описується сукупністю параметрів [6, 18]: TP – кількість вірно розпізнаних об'єктів, TN – кількість нерозпізнаних об'єктів; FP – кількість помилкових позитивних передбачень, FN – вірних нерозпізнавань. Тобто якісно навчена нейронна мережа повинна забезпечити не тільки високу ймовірність вірного розпізнавання, але і мінімально можливі ймовірності прийняття фонових звуків за акустичні випромінювання дрона. Для цього була підготовлена база аудіозаписів фонових звуків Ground (вулиць міста, природи), а також джерел звуків, які за часовими характеристиками мають близький характер, тобто теж містять періодичні складові: літак, автобус, автомобіль, гелікоптер, мотоцикл, поїзд, фура. Всього було підготовлено 853 аудіофайла з тривалістю не менше 10 с. З цих файлів для навчання нейронної мережі було обрано 70 %, для валідації – 10 %, а для тестування навченої нейронної мережі – 20 %.

Для досліджень були спроектовані 3 архітектури нейронних мереж з використанням фреймворку Tensorflow [19]. Кожна з них має від 8 до 32 входів, що відповідає кількості коефіцієнтів MFCC, а також 9 виходів, що відповідають 9 класам об'єктів розпізнавання (дрон, Ground, літак, автобус, автомобіль, гелікоптер, мотоцикл, поїзд, фура). В першій архітектурі є 3 проміжних шари з кількістю нейронів в шарах 1024, 265, 64. В другій архітектурі є 4 проміжних шари з кількістю нейронів в шарах 1024, 512, 128, 32. Такі архітектури є простими з точки зору реалізації, але мають проблеми з процесом навчання. Тому в третю архітектру були додані додаткові шари BatchNormalization, що нормалізує активації нейронів для прискорення навчання та стабілізації градієнтів, та Dropout шари, що випадково вимикає частину нейронів під час навчання для запобігання перенавчанню. В трьох проміжних шарах цієї архітектури міститься 512, 128 та 32 нейрона. Остання архітектура наведена на рис. 2.

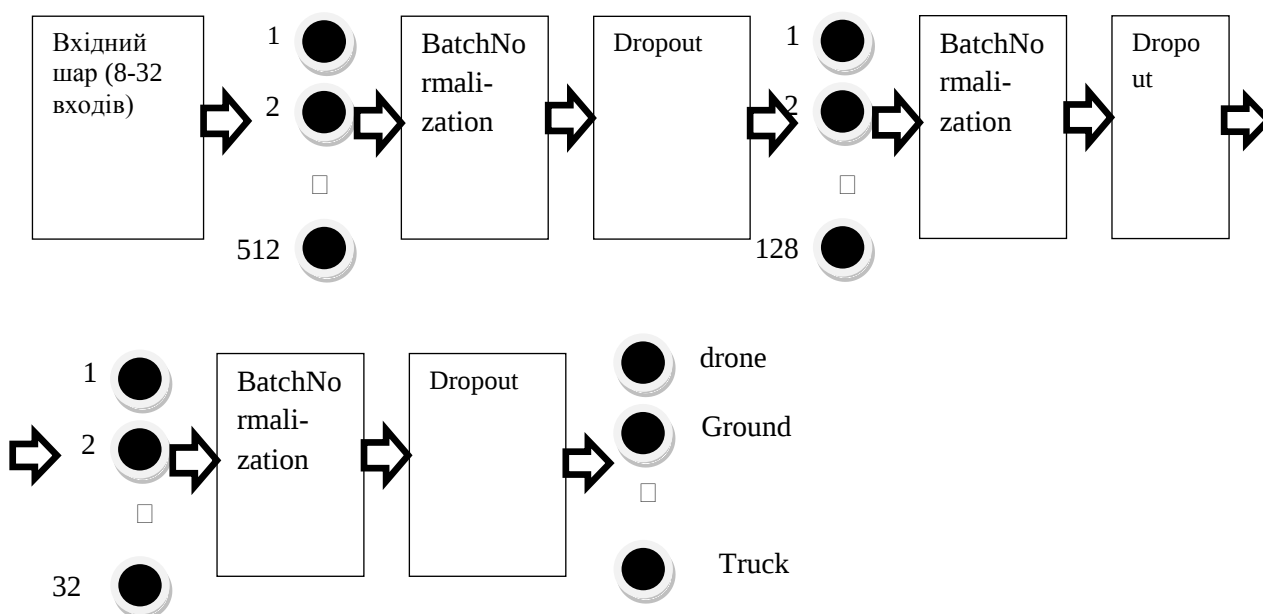


Рис. 2. Архітектура нейронної мережі із захистом від перенавчання

В усіх нейронах проміжних шарів використовується функція активації ReLU, що найбільш відповідає біологічному аналогу математичного нейрона та потребує невеликого процесорного навантаження, а в вихідному шарі використовується функція активації softmax. Навчання нейронних мереж виконувалось з використанням двох найбільш ефективних оптимізаторів Adam та RMSprop [20]. RMSprop працює за рахунок збереження експоненційного середнього квадратів градієнтів (дисперсії) та поділу градієнта на квадратний корінь цієї величини, що дозволяє уникнути проблем із занадто великими або малими оновленнями ваг. Він добре працює для рекурентних нейронних мереж і задач, де спостерігаються шумні або нестаціонарні градієнти. Adam, на відміну від RMSprop, поєднує два механізми: експоненційне згладжування середніх значень градієнтів (моменту) як у методу Momentum, і експоненційне згладжування квадратів градієнтів, як у RMSprop. Це дозволяє Adam швидко збігатися, добре працювати при нестабільних градієнтах і мінімізувати ручне налаштування гіперпараметрів. Завдяки поєднанню переваг двох підходів, Adam зазвичай забезпечує швидшу та стабільнішу збіжність порівняно з RMSprop, особливо в задачах з великими наборами даних або складними функціями втрат.

Навчання архітектур нейронних мереж та аналіз результатів навчання

Для навчання розроблених архітектур нейронних мереж використовувалась бібліотека tensorflow. Навчання проводилось впродовж 150 епох, але з функцією автоматичного припинення навчання, коли впродовж 8 епох немає зменшення втрат. Розмір одного батча 400 векторів MFCC. Навчання проводилось багаторазово для різної кількості коефіцієнтів MFCC, різних архітектур та оптимізаторів. За результатами кожного навчання було отримано матрицю помилок, приклад якої наведено на рис. 3.

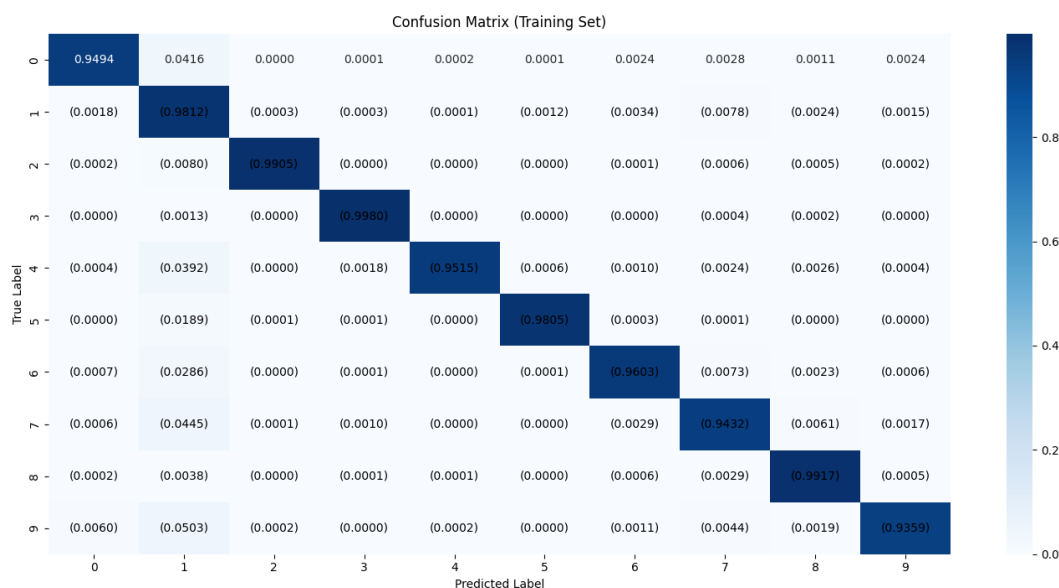


Рис. 3. Матриця помилок для архітектури №3, оптимізатора Adam, кількість часових відліків 8192 та 28 кількості коефіцієнтів MFCC

Як можна побачити, з матриці помилок навчання було виконано достатньо якісно, тобто ймовірність вірного розпізнавання знаходиться в межах від 93,5 до 99,8 %, ймовірність помилкового розпізнавання інших класів сигналів, як дронів до 0,6 %. Аналіз аналогічних матриць для різної кількості MFCC коефіцієнтів в діапазоні від 8 до 36 показує, що значення ймовірності вірного розпізнавання може суттєво змінюватись. Наприклад, при довжині вікна 8192 відліки та використанні оптимізатора Adam діапазон ймовірностей вірного розпізнавання знаходиться в межах від 0,81 до 0,987, тобто складає 17,7 %.

Після навчання проводилось тестування кожної навченої нейронної мережі на незалежній сукупності аудіозаписів. За результатами цього тестування було отримано декілька груп залежностей: 1) залежності ймовірності вірного розпізнавання від кількості MFCC коефіцієнтів; 2) залежності ймовірності вірного розпізнавання від довжини часового вікна та оптимізатора; 3) залежності ймовірності вірного розпізнавання від архітектури нейронної мережі. На рис. 4 наведено залежності ймовірності вірного розпізнавання від кількості MFCC коефіцієнтів.

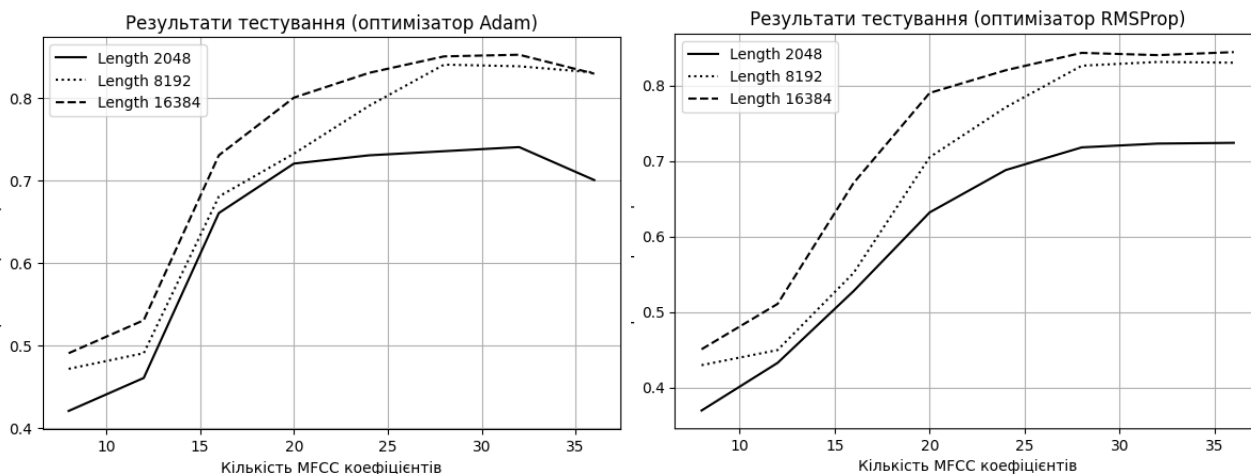


Рис. 4. Залежності ймовірності вірного розпізнавання від кількості MFCC коефіцієнтів

Аналізуючи залежності, наведені на рис. 4, можна зробити висновки, що існує оптимальне значення кількості коефіцієнтів MFCC для подальшої обробки нейронними пов'язаними мережами, що забезпечує максимізацію ймовірності вірного розпізнавання. Для часових вікон з кількістю відліків 2048 оптимальним є використання 32 коефіцієнтів MFCC, із зростанням довжини вікна до 16384 відліків достатньо використовувати 28 коефіцієнтів MFCC, бо подальше збільшення їх числа не дозволяє збільшувати ймовірність вірного розпізнавання. Використання малого числа коефіцієнтів 8–12 поступається використанню оптимального числа коефіцієнтів до 38 %. Якщо порівнювати між собою залежності, наведені на рис. 4, а, б, то можна побачити, що оптимізатор Adam забезпечує на 4,3 % більшу ймовірність вірного розпізнавання при малій кількості коефіцієнтів MFCC (8–10) та на 0,9–1,7 % більшу ймовірність при оптимальних значеннях числа коефіцієнтів MFCC. Отже, оптимізатор Adam можна рекомендувати до практичного використання. Також порівнюючи використання різних довжин часових вікон можна зробити висновок, що малі часові вікна довжиною 2048 відліків значно програють більш довгим (8192 та 16384 відліки) за ймовірністю вірного розпізнавання. Цей програш становить 10,8 % використанню вікна довжиною 8192 відліки. В той самий час збільшення довжини вікна з 8192 до 16384 відліків покращує ймовірність вірного розпізнавання до 1 %. Тобто, якщо кінцевою метою системи розпізнавання є максимізація ймовірності вірного розпізнавання, то слід використовувати часові вікна довжиною 16384 відліки, а при реалізації компромісу між максимізацією ймовірності вірного розпізнавання та апаратними вимогами слід використовувати часові вікна 8192 відліки.

На рис. 5 показані залежності сумарних ймовірностей похибкового розпізнавання різноманітних акустичних джерел як випромінювання дрона. Аналіз цих залежностей показує, що оптимальною кількістю коефіцієнтів MFCC є 28–32 коефіцієнта, коли забезпечується мінімум ймовірностей похибкового розпізнавання. Оптимальною довжиною часової виборки є 8192 відліки. Використання більших виборок навіть збільшує ймовірностей похибкового розпізнавання. Найкращі результати знов забезпечує використання оптимізатора Adam.

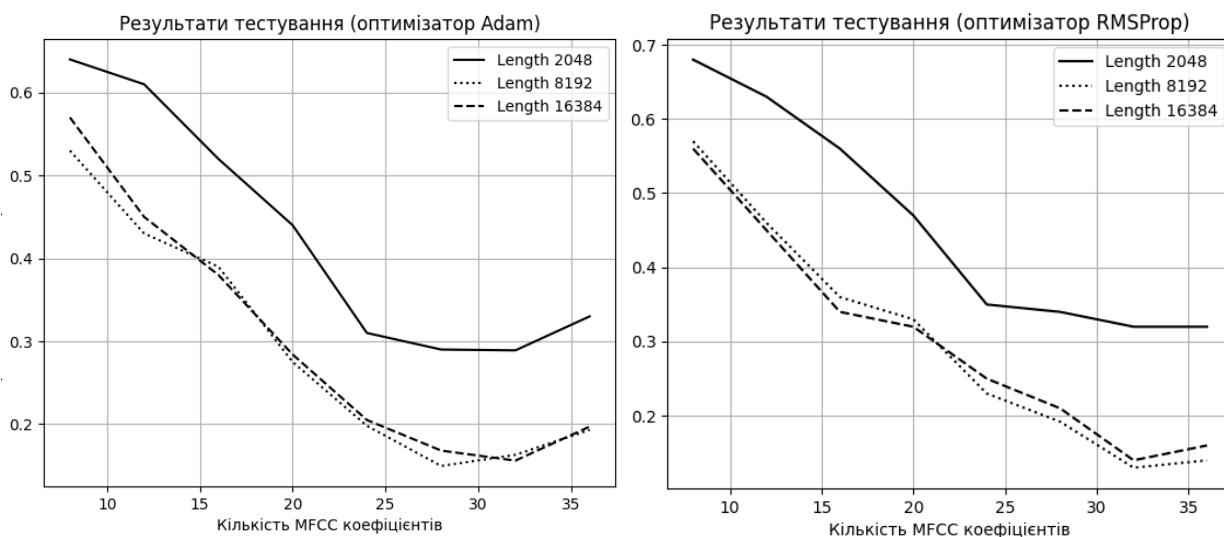


Рис. 5. Залежності сумарних ймовірностей похибкового розпізнавання 8 акустичних джерел як випромінювання дрона

Аналогічні дослідження були проведені для першої та другої архітектур нейронних мереж. Результати цих досліджень для часового вікна з кількістю відліків 16384 та оптимізатора Adam наведені у табл. 1.

Таблиця 1

Номер архітектури	Ймовірність вірного розпізнавання TP		Ймовірність похибкового розпізнавання FP	
	Adam	RMSprop	Adam	RMSprop
1	0,734	0,732	0,1773	0,1884
2	0,772	0,758	0,1628	0,1716
3	0,882	0,879	0,1495	0,1523

Аналіз даних табл. 1 показує, що архітектури нейронних мереж без використання шарів, що запобігають перенавчанню та покращують навчання програють архітектурі з такими шарами до 14,8–11 %.

Для архітектури нейронної мережі №3, оптимізатора Adam та кількості відліків у часовому вікні 16384 було проведено тестування ефективності роботи нейронної мережі в діапазоні польотів дрону від 5 до 100 м з GPS трекінгом польоту дронів та подальшої сумісної обробки трекінгу та акустичного випромінювання. Приклад такого розпізнавання показаний на рис. 6.

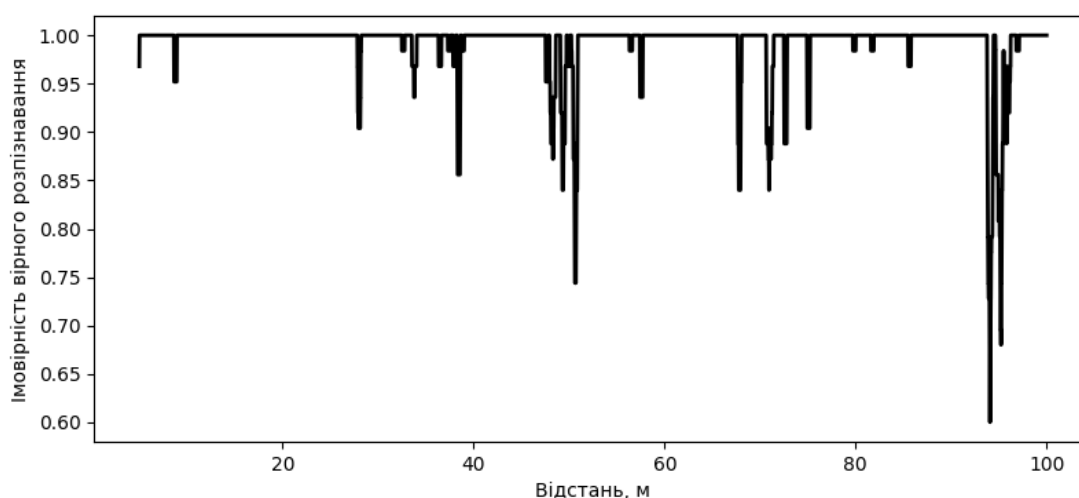


Рис. 6. Приклад розпізнавання дрону на відстанях від 5 до 100 м

Як можна побачити з даних рис. 6 та аналогічних обробок польотів дронів, якісне розпізнавання (практично без помилок) можливе на відстанях до 50 м. Із зростанням дальності зростає кількість помилок і на відстанях порядку 100 м вона складає до 23 %, що потребує використання додаткових вторинних методів обробки, до яких можна віднести методи трекінгу та прийняття рішення по сукупності результатів обробки впродовж інтервалів часу, що відповідають тривалості 3–5 часових вікон.

Крім аналізу якості роботи розроблених нейронних мереж було проаналізовано кількість параметрів нейронної мережі №3 та кількість потрібної оперативної пам'яті на реалізацію роботи навченої нейронної мережі на спеціалізованих процесорах для вбудовування нейронних мереж, таких як STM32MP157. При використанні 16-розрядного вбудованого в мікропроцесор аналого-цифрового перетворювача, каналу DMA з подвійним буфером для зберігання одразу 32768 відліків акустичного сигналу, а також обробки 32 коефіцієнтів потрібно не менше 762 Кбайт оперативної пам'яті. Тобто використання повнозв'язних нейронних мереж дозволяє значно зменшити вимоги до апаратної платформи при реалізації розпізнавання акустичних випромінювань дронів і зменшити вартість подібних пристроїв.

Висновки

1. Проведено аналіз сучасних методів розпізнавання акустичних випромінювань дронів в часовій області та частотній. Обґрунтовано ефективність використання MFCC для подальшого розпізнавання векторів цих коефіцієнтів повнозв'язними нейронними мережами.

2. Розроблено 3 архітектури нейронних мереж, виконано акустичні записи польотів дронів на різних відстанях, а також же 8 різних джерел фонових звуків та звуків з періодичними часовими складовими, проведено навчання розроблених архітектур за допомогою створеного датасету акустичних записів, а також тестування навчених архітектур на незалежній частині датасету.

3. Доказано, що існує оптимальне значення кількості коефіцієнтів MFCC для розпізнавання акустичних випромінювань дронів. Воно дорівнює 28–32, при якому забезпечується максимізація ймовірності вірного розпізнавання до 0,882 та мінімізація похибкового розпізнавання до 0,1495. Оптимальною довжиною часового вікна при розрахунку коефіцієнтів MFCC є 372 мс. Використання оптимізатора Adam дає вигаш 0,9–1,7 % в ймовірності вірного розпізнавання у порівнянні з оптимізатором RMSProp. Використання шарів Batch Normalization та Dropout при створенні архітектури нейронної мережі дозволяє збільшити ймовірність вірного розпізнавання на 11–14,8 %.

4. На відстанях до 50 м реалізується розпізнавання акустичних випромінювань дронів з високою ймовірністю вірного розпізнавання. На відстанях порядку 100 м похибки розпізнавання складають порядку 23 %. Для зменшення впливу цих похибок розпізнавання дронів, а також похибкового розпізнавання інших джерел звуку, як дронів, можливо використання методів вторинної обробки, що є предметом подальших наукових досліджень

5. Використання повнозв'язних нейронних мереж при розпізнаванні акустичних випромінювань дронів можливо реалізувати на спеціалізованих процесорних платформах типу STM32MP157, що значно зменшують собівартість таких рішень завдяки невеликим вимогам до оперативної пам'яті порядку 1 Мбайт.

Список літератури:

1. Agapiou A. Drones in Construction: A Comparative International Review of the Legal and Regulatory Landscape // *Management Procurement and Law*. 2020. Vol. 174, No.3. P.1–8. DOI:10.1680/jmapl.19.00041
2. Pyrgies J. The UAVs threat to airport security: risk analysis and mitigation // *Journal of Airline and Airport Management*. 2019. Terrassa. Vol. 9, Iss.2. P.63–96.
3. Mowafaq SA, Muhyeeddin A, Al-Batah MS. AI in the Sky: Developing Real-Time UAV Recognition Systems to Enhance Military Security // *Data and Metadata*. Nazarre. 2024. Vol. 3. P.1–19. <https://doi.org/10.56294/dm2024.417>
4. Aouladhadj D., Kpre E., Deniau V., Kharchouf A., Gransart C., Gaquière C. Drone Detection and Tracking Using RF Identification Signals // *Sensors*. Basel. 2023. Vol. 23. P.1–24. <https://doi.org/10.3390/s23177650>

5. Yousaf J., Zia H., Alhalabi M., Yaghi M., Basmaji T., Shehhi E.A., Gad A., Alkhedher M., Ghazal M. Drone and Controller Detection and Localization: Trends and Challenges// Appl. Sci. Basel. 2022. Vol.12, Iss.24, P.1–22. <https://doi.org/10.3390/app122412612>
6. Zubkov O.V., Sheiko S.O., Oleynikov V.M., Kartashov V.M., Babkin S.I. Investigation Of The Yolov5 Algorithm Efficiency For Drone Recognition // Telecommunications and Radio Engineering. Danbury. 2024. Vol.83, Iss.1. P. 65–79. DOI: 10.1615/TelecomRadEng.2023048987
7. Oleynikov V., Zubkov O., Kartashov V., Korytsev I., Babkin S., Sheiko S. Investigation of the efficiency of detection and recognition of small-sized unmanned aerial vehicles by their acoustic radiation. Radiotekhnika. 2018. Vol.4, No.195. P.209–217. <https://doi.org/10.30837/rt.2018.4.195.21>
8. Paszkowski W., Gola A., Świć A. Acoustic-Based Drone Detection Using Neural Networks – A Comprehensive Analysis// Advances in Science and Technology Research Journal. Lublin. 2024. Vol.18, Iss.1. P.36–47. <https://doi.org/10.12913/22998624/175863>
9. Tejera-Berengue D., Zhu-Zhou F., Utrilla-Manso M., Gil-Pita R., Rosa-Zurera M. Analysis of Distance and Environmental Impact on UAV Acoustic Detection // Electronics. Basel. 2024. Vol. 13. <https://doi.org/10.3390/electronics13030643>
10. Othman E., Cibilić I., Poslončec-Petrić V., Saadallah D. Investigating Noise Mapping in Cities to Associate Noise Levels with Sources of Noise Using Crowdsourcing Applications// Urban Sci. 2024. Vol.8, Iss.1. <https://doi.org/10.3390/urbansci8010013>
11. Dombrovski M., Deaconu M., Cristea L., Frigioescu T.F., Cican G., Badea ., Totu G. Acoustic Analysis of a Hybrid Propulsion System for Drone Applications// Acoustics. Basel. 2024. Vol. 6, Iss.3. P. 698–712. <https://doi.org/10.3390/acoustics6030038>
12. Islam M., Haque M., Islam S., Mia Z.A., Rahman M. DCNN-LSTM Based Audio Classification Combining Multiple Feature Engineering and Data Augmentation Techniques Intelligent // Computing & Optimization Springer. Cham. 2021. Vol. 371. P.227–236. DOI:10.1007/978-3-030-93247-3_23
13. Wei N., Gu J.X., Gu F., Chen Z., Li G., Wang T., Ball A.D. An Investigation into the Acoustic Emissions of Internal Combustion Engines with Modelling and Wavelet Package Analysis for Monitoring Lubrication Conditions // Energies. Basel. 2019. Vol.12, Iss.4. P. 1–14. <https://doi.org/10.3390/en12040640>
14. Zrar K.A., Abdulbasit K.A. Mel Frequency Cepstral Coefficient and its Applications: A Review // IEEE Access. Piscataway. 2022. Vol.10. P.122136–122158. DOI:10.1109/ACCESS.2022.3223444
15. Sithara A., Abraham T., Mathew D. Study of MFCC and IHC Feature Extraction Methods With Probabilistic Acoustic Models for Speaker Biometric Applications // Procedia Computer Science. Amsterdam. 2018. Vol. 143. P. 267–276. DOI:10.1016/j.procs.2018.10.395
16. Fazal M.A., Baig M.A., Manj W.A., Faraz Z., Mallah G.A. Implementation of Deep Learning for Acoustic Classification // Journal of Xidian University. Christchurch. 2023. Vol. 17, Iss. 8. P.1653–1673. DOI:10.37896/jxu17.8/138
17. Pham L., Ngo D., Salovic D.c, Jalali A., Schindler A., Nguyen P. X., Tran K., Vu H. Lightweight deep neural networks for acoustic scene classification and an effective visualization for presenting sound scene contexts // Applied Acoustics. Basel. 2023. Vol. 211. P.723–731. DOI:10.1016/j.apacoust.2023.109489
18. Jung H.K., Choi G.S. Improved YOLOv5: Efficient Object Detection Using Drone Images under Various Conditions // Appl. Sci. Basel. 2022. Vol. 12. P.1–16. DOI: 10.3390/app12147255
19. Pang B., Nijkamp E., Nian Y.W. Deep Learning With TensorFlow: A Review // Journal of Educational and Behavioral Statistics. 2019. Vol. 45, Iss. 2. P.415–421. <https://doi.org/10.3102/107699861987276>
20. Soujanya B., Sitamahalakshmi T. Optimization with ADAM and RMSprop in Convolution neural Network (CNN): A Case study for Telugu Handwritten Characters // International Journal of Emerging Trends in Engineering Research. Pawan. 2020. Vol. 8. No. 9. P.5116–5121.

Надійшла до редколегії 12.06.2025

Відомості про авторів:

Зубков Олег Вікторович – канд. техн. наук, професор, Харківський національний університет радіоелектроніки, завідувачий кафедрою мікропроцесорних технологій і систем, Україна; e-mail: oleh.zubkov@nure.ua; ORCID: <https://orcid.org/0000-0002-8528-6540>

Бойко Наталія Вікторівна – Харківський національний університет радіоелектроніки, асистент кафедри мікропроцесорних технологій і систем, Україна; e-mail: natalia.boiko@nure.ua

Мачоніс Тадас Сігігасович – Харківський національний університет радіоелектроніки, аспірант кафедри мікропроцесорних технологій і систем, Україна; e-mail: tadas.machonis@nure.ua