

КОШЛАНЬ МИКОЛА

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-1477-124x>

ШАХОВСЬКА НАТАЛІЯ

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-6875-8534>

Nataliya.b.shakhovska@lpnu.ua

МЕЛЬНИКОВА НАТАЛІЯ

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-2114-3436>

## МЕТОДИ ТА ЗАСОБИ РОЗПІЗНАВАННЯ ЛІТАЛЬНИХ ОБ'ЄКТІВ

Проведено аналіз результатів застосування нового алгоритму класифікації зображень, який був випробуваний на датасеті, що включає зображення об'єктів у небі. Отримані результати свідчать про високу точність цього алгоритму у класифікації зображень, і він ефективно працює у паралельному режимі. Для оцінки ефективності запропонованого алгоритму було проведено порівняльний аналіз швидкості обчислень при різних кількостях потоків. Один із можливих напрямків для майбутніх досліджень може включати розширення обсягу дослідження на інші типи зображень і використання більш складних архітектур нейронних мереж для покращення результатів. Отримані результати можуть бути застосовані у розробці систем автоматичної класифікації зображень в різних галузях, таких як медицина, безпека, маркетинг та інші.

Ключові слова: нейронні мережі, CNN, точність, розпаралелення.

KOSHLAN N., SHAKHOVSKA N., MEKNYKOVA N.

Lviv Polytechnic National University

## METHODS AND MEANS OF RECOGNITION OF FLYING OBJECTS

This research project revolves around the development of a sophisticated system dedicated to the recognition of objects in the sky. The analysis of the results of the application of a new image classification algorithm was carried out, which was tested on a dataset that includes images of objects in the sky. The obtained results demonstrate the high accuracy of this algorithm in image classification, and it works effectively in parallel mode. To evaluate the efficiency of the proposed algorithm, a comparative analysis of the speed of calculations with different amounts of flows was carried out. It was found that parallel computations significantly reduce the execution time of the algorithm compared to the sequential approach. One possible direction for future research could include extending the scope of the study to other types of images and using more complex neural network architectures to improve the results. The obtained results can be applied in the development of automatic image classification systems in various fields, such as medicine, security, marketing, and others. Leveraging cutting-edge neural networks, state-of-the-art machine learning techniques, and advanced image processing technologies, this system exhibits the remarkable capability to autonomously detect and categorize a diverse array of objects within Earth's atmosphere. The dataset comprises photos of airplanes, helicopters, missiles, fighter jets, drones. One of the pivotal objectives of this endeavor is to ensure not only the efficiency but also the user-friendliness of the system. This dual focus is aimed at facilitating a seamless integration of this technology into a plethora of domains, including meteorology for more accurate weather forecasting, security for enhanced surveillance and monitoring capabilities, and scientific research within the realms of astronomy and aerospace technology, enabling groundbreaking discoveries and innovations. In conclusion, this research represents a significant leap forward in our quest to gain a deeper understanding of our ever-evolving skies and harness this knowledge for various practical applications. It underscores the transformative potential of modern technology in addressing multifaceted challenges and expanding the horizons of human knowledge.

Keywords: neural networks, CNN, accuracy, concurrency

### Вступ

Процес виявлення повітряних цілей може здатися простим через несумісність літальних апаратів, вертольотів, БПЛА та дронів щодо фону ясного неба. Коли враховуються зміни фону, зміна яскравості неба ускладнює процес. Зміни у формах і типах хмар додають цьому процесу ще одну складність. Процес відстеження безпосередньо залежить від процесу виявлення та типу потоку даних. Практичні системи, що використовуються для відео виявлення та супроводу повітряних цілей, є ручними, а ручні конструкції мають деякі недоліки порівняно з автоматичними. Для програм відеоспостереження, інструкцій, регіональної безпеки та оборони в щільному середовищі процес автоматичного виявлення та відстеження може бути радше обов'язковим, ніж перевагою.

Задачею роботи є дослідження проблеми розпізнавання та відстеження об'єктів на відео неба, а також створення базової моделі для розв'язання даної задачі. Результатом виявлення об'єктів на дорозі вважається:

- Розпізнавання типу об'єкту
- Робота моделі з використанням переданого відео.

### Аналіз літературних джерел

У дослідженні [1] запропоновано алгоритм автоматичного виявлення та супроводу відео потоків повітряних цілей. Наземна рухома камера знімає відеодані, і в русі перебувають не тільки літаючі об'єкти, але, ймовірно, і камера. Хоча виявлення та відстеження рухомих об'єктів за допомогою рухомих датчиків є відносно важким завданням, це поширений випадок у сценаріях реального життя. Системи відео детектування та стеження мають один або кілька рухомих відео датчиків, коли один або кілька літаючих

апаратів знаходяться в зоні дії. Запропонований алгоритм включає етап обробки зображення для виявлення та етап відстеження для ініціації та продовження. Було проведено оціночне дослідження фактичних відеоданих і встановлено, що запропонований метод дає успішні результати для процесів виявлення, формування доріжок і продовження.

Дослідження [2] присвячено зображенням з туманною сценою які страждають від низької контрастності, що знижує видимі якість сцени, що робить виявлення об'єктів більш складним завданням. Низька контрастність може бути результатом туманної погоди під час отримання зображення. Усунення затемнення — це процес видалення туману з фотографії туманної сцени. Добре відомими методами в цій галузі є Single-image dehazing на основі попереднього темного каналу. Стаття [3] має на меті систематичний огляд методів відстеження та виявлення об'єктів у разі загроз для додатків на основі GANN. Вибрані дослідження базувалися на різних факторах, таких як рік публікації, метод, реалізований у статті, надійність вибраних алгоритмів і розмір набору даних. Кожне дослідження підсумовується, призначаючи його для одного з двох попередньо визначених завдань: застосування GANN або використання традиційних методів машинного навчання (ML).

У статті [4] автори пропонують новий метод виявлення відео об'єктів, який використовує як контекст руху об'єкта, так і просторово-часові агреговані характеристики для підвищення ефективності виявлення відео об'єктів. По-перше, контекст руху об'єкта витягується за допомогою оператора кореляції між картами ознак двох сусідніх кадрів. На додаток до створення контексту руху, карти просторових функцій для  $N$  суміжних кадрів агрегуються, щоб підвищити якість карти функцій із закритою мережею уваги. У дослідженні [5] розповідається, що виявлення відео об'єктів є важливим напрямком досліджень комп'ютерного зору. Завдання виявлення відео об'єктів полягає у виявленні та класифікації рухомих об'єктів у послідовності зображень. На основі детектора об'єктів статичного зображення більшість існуючих методів виявлення об'єктів відео використовують унікальну часову кореляцію відео для вирішення проблеми пропущеного виявлення та помилкового виявлення, спричиненого оклюзією та розмиттям рухомого об'єкта. Широко використовується інша модель виявлення відео об'єктів, керована мережею оптичного потоку. Агрегація ознак суміжних кадрів виконується шляхом оцінки поля оптичного потоку. У [6] пропонується метод виявлення відео об'єктів, який поєднує функції руху та YOLO. У цьому методі використовується метод фільтрації відеокадрів без ознак руху та сегментації відеокадрів із ознаками руху, щоб зменшити тиск міркувань моделі алгоритму YOLO. У цьому процесі відеокадри, що містять рухомі об'єкти, спочатку отримуються модулем виявлення рухомих об'єктів. По-друге, рухома ціль буде розпізнана модулем розпізнавання цікавого об'єкта. Нарешті, модуль фонового прийняття рішень записує та аналізує результати виявлення, щоб отримати оновлення фонової моделі або результат.

#### Методи та моделі

Щоб навчити систему було використано датасет з фотографій об'єктів в небі. Датасет складався з об'єктів п'яти типів: airplane (1 425 фото), helicopter (1 388 фото), missile (1 335 фото), fighter\_jet (1 868 фото), drone (1333 фото). Було використано фото двох типів PNG та JPEG. Після обробки, система формує вихідні дані у вигляді відео, у підписі якого є додана інформація з такими даними: Object recognition – назва виявленого об'єкта, FPS – кількість кадрів в секунду. Аугментація даних відіграє важливу роль у підвищенні стійкості моделей у різних умовах та середовищах, а також у запобіганні перенавчанню, коли модель стає надто адаптованою до конкретного набору даних. Вона допомагає зробити моделі більш універсальними та ефективними в реальних сценаріях використання.

Для оптимізації розпізнавання зображень з відеоряду у статті запропоновано розпаралелення обробки кадрів. Підхід полягає у розділенні обробки відео на паралельні задачі, де кожен кадр відеоряду обробляється окремо. Це дозволяє ефективно використовувати ресурси обчислювального обладнання та забезпечує швидку обробку кадрів. Застосовуючи розпаралелення обробки кадрів, можливо досягти високої продуктивності та точності розпізнавання зображень у відеоряді. Було проведено аналіз різних підходів для оптимізації та встановлено, що більшість підходів базуються на глибоких нейронних мережах, які забезпечують високу точність розпізнавання [7], але при цьому потребують значних обчислювальних ресурсів. Однак, на сьогодні, відсутній універсальний метод для цієї задачі, що би забезпечував високу точність при низьких витратах на обчислювальні ресурси. Для реалізації згорткової нейронної мережі було покращено модель шляхом комбінування кількох шарів згортки, ректифікації, пулінгу та дропауту. Також важливим кроком є правильний підбір функцій активації, що потенційно покращують точність класифікації. Пошарова будова мережі:

1. Conv2D: шар згортки з 32, 64, 96 та 128 фільтрами відповідно.
2. Activation: шар активації за допомогою функції ReLU, яка розгладжує нелінійність ознак.
3. BatchNormalization: шар нормалізації пакету даних для зменшення внутрішньокласової дисперсії.
4. MaxPooling2D: шар зменшення розміру зображення, застосовується після кожного блоку Conv2D.
5. Dropout: рандомно відкидає частину нейронів, що допомагає уникнути перенавчання моделі.
6. Flatten: шар розгортання, що перетворює матричний вихід із шарів згортки в вектор.
7. Dense: пов'язаний шар, що з'єднує всі вхідні вузли з усіма вихідними вузлами.
8. Activation: шар активації за допомогою функції ReLU.
9. BatchNormalization: шар нормалізації пакету даних.
10. Dropout: рандомно відкидає частину нейронів.

11. Dense: пов'язаний шар, що з'єднує всі вхідні вузли з усіма вихідними вузлами.

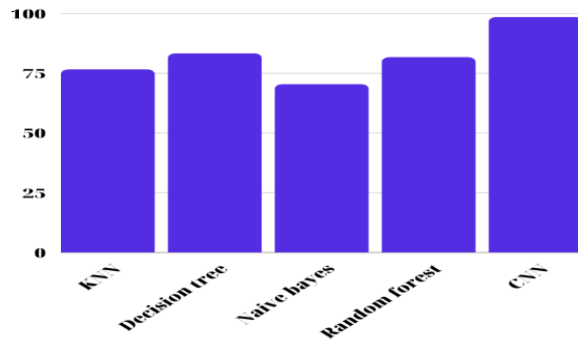
Нарешті, модель компілюється з використанням оптимізатора Adam та втрати "binary\_crossentropy". Модель тренується з використанням методу fit\_generator. Основні параметри навчання включають кількість епох, розмір пакету та кількість кроків на епоху. Після навчання графік відображається для відслідковування втрат та точності навчання та валідації.

Для розпаралелення використовується модуль multiprocessing, а саме його функція Pool. Після того, як програма завантажила модель, створюється об'єкт пулу процесів: pool = Pool(). Наступним кроком програма читає кадри з веб-камери в циклі за допомогою webcam.read(), і передає кожен з них у функцію process\_frame() для обробки. Замість того, щоб обробляти кадри послідовно, вони оброблюються паралельно на різних процесорах. Для цього функція map() викликає функцію process\_frame() для кожного кадру. Результатом роботи функції map() є список оброблених кадрів, які далі відображаються відразу на екрані з використанням cv2.imshow().

**Результати експерименту**

Для тренування моделі та визначення об'єкта у небі було обрано датасет, що називається "Flying planes" [8]. Цей набір даних складається з близько 6500 фотографій різних об'єктів у небі, і використовується для навчання та тестування алгоритмів комп'ютерного зору та машинного навчання.

Було навчено 5 різних типів моделей. Як результат було отримано різні показники точності (accuracy) рис. 1. Отже, модель CNN показала найвищу точність серед усіх досліджених моделей, досягнувши 95.15%. Це свідчить про те, що CNN є дуже потужною та ефективною моделлю для завдань класифікації об'єктів на зображеннях, особливо коли дані мають великий обсяг та складну структуру. Під час крос-валідації нашої моделі було використано серію метрик для оцінки її ефективності під час кожного із збірних тестувань. Зазвичай ці метрики обчислюються окремо для кожного фолду (частини даних у крос-валідації) та подальше об'єднуються для отримання загального результату.



**Рис. 1. Результати точності для різних моделей**

Нижче наведено результати крос-валідації:

Таблиця 1

Результати крос-валідації

Фолд	Accuracy	Precision	Recall	Specificity	F1-score
1	0.9112	0.9106	0.9482	0.9462	0.9467
2	0.922	0.9235	0.9391	0.9473	0.9322
3	0.9354	0.9223	0.9374	0.9724	0.9362
4	0.9572	0.9248	0.9531	0.9346	0.9546
5	0.9515	0.9369	0.9668	0.9582	0.9618
Середнє	0.93566	0.9236	0.9489	0.95174	0.9463

Виведемо графік відношення метрик до кількості епох для більшої наглядності процесу навчання моделі (рис. 2. та рис. 3).

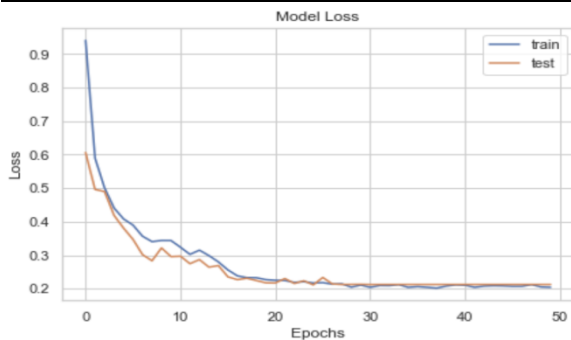


Рис. 2. Відношення метрики Loss до кількості епох

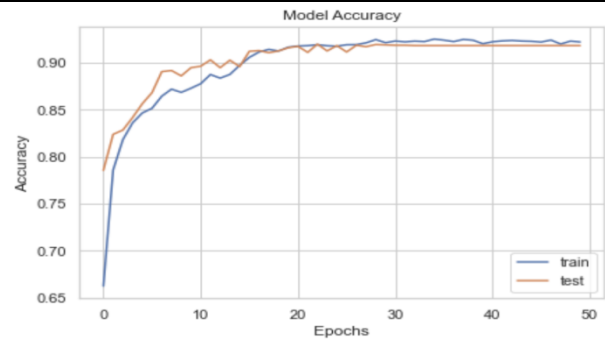


Рис. 3. Відношення метрики Accuracy до кількості епох

Наступним кроком є використання навченої моделі для розпізнавання об'єктів на тестовому відео. Для цього, перш за все, потрібно визначити оптимальну кількість потоків при розпаралелюванні обробки кадрів. Були проведені експерименти з використанням одного, двох, чотирьох та восьми потоків. В результаті експериментів було отримані результати, які показують значне покращення у швидкості в наслідок використання паралельної обробки кадрів (таблиця 2).

Таблиця 2

Показники ефективності на різній кількості потоків

К-сть потоків	1	2	4	8
Час обробки 1 кадру	~329 мсек	~238 мсек	~163 мсек	~90 мсек
Кількість оброблених кадрів за 30 секунд	91	126	183	331

### Висновок

Порівняно з результатами досліджень [4, 5], наша модель показала вищу точність класифікації. Однак вона демонструє меншу точність, ніж модель з дослідження [6]. Треба відзначити, що наша модель забезпечує більшу швидкість обробки кадрів на секунду, оскільки ми вдосконалили процес обробки кадрів в режимі онлайн відео шляхом використання розпаралелення, що дозволило нам значно підвищити темп обробки кадрів на секунду. Порівнюючи результати, виявилось, що при використанні однопотокowego підходу програма обробляла всього близько 90 кадрів за 30 секунд, у той час як використання восьми потоків істотно покращило продуктивність – досягнуто близько 331 кадрів за 30 секунд, що еквівалентно 11 кадрам на секунду. Один із можливих напрямків для майбутніх досліджень може включати розширення обсягу дослідження на інші типи зображень і використання більш складних архітектур нейронних мереж для покращення результатів.

### References

1. Pakfiliz G. A New Method for Surface-to-Air Video Detection and Tracking of Airborne Vehicles. IEEE Transactions on Aerospace and Electronic Systems, vol. 56, no. 5, pp. 3908-3922, Oct. 2020.
2. Sahu G., Seal A., Krejcar O., and Yazidi A. Single image dehazing using a new color channel. IEEE Access, vol. 8, pp. 82613-82622, 2020.
3. Al Jaber S. M., Patel A., and AL-Masri A. N. Object tracking and detection techniques under GANN threats: A systemic review. Journal of Ambient Intelligence and Humanized Computing, vol. 12, no. 3, pp. 2857-2872, Mar. 2021
4. Kim J., Koh J., and Choi J. W. Video Object Detection Using Motion Context and Feature Aggregation. Journal of Image Processing and Information Sciences, vol. 10, no. 2, pp. 369-379, Apr. 2020.
5. You H., Lu Y., and Tang H. Improved Feature Extraction and Similarity Algorithm for Video Object Detection. Journal of Electronic Imaging, vol. 30, no. 6, pp. 063020, Nov. 2021.
6. Xiao C., Liu P., Zhou Y. Research on Video Object Detection Methods Based on YOLO with Motion Features. Journal of Computer Science and Information Security, v.18, no. 3, pp. 69-76, Mar. 2020.
7. Ma X., Liu W. (2018). Remote sensing image classification using adaptive sparse representation and stacked denoising autoencoder. Remote Sensing, 10(4), 538. doi: 10.3390/rs10040538
8. Kaggle. Flying Vehicles Dataset. URL: <https://www.kaggle.com/datasets/eabdul/flying-vehicles>