

## МЕТОД ІДЕНТИФІКАЦІЇ БОЙОВИХ МАШИН НА ОСНОВІ YOLO

Марія Назаркевич<sup>1</sup>, Назар Олексів<sup>2</sup>

Національний університет “Львівська політехніка”,  
кафедра інформаційних систем та мереж, Львів, Україна

<sup>1</sup>E-mail: mariia.a.nazarkevych@lpnu.ua, ORCID 0000-0002-6528-9867

<sup>2</sup>E-mail: nazar.t.oleksiv@lpnu.ua, ORCID 0000-0001-7821-3522

© Назаркевич М., Олексів Н., 2024

Запропоновано метод розпізнавання контурів об'єктів у відеопотоці даних. Дані потрібно завантажити за допомогою відеокамери у режимі реального часу та здійснити розпізнавання об'єктів. Використано мережу YOLO – метод ідентифікації та розпізнавання об'єктів у реальному часі. Розпізнані об'єкти будуть записані у відеопослідовності із зазначенням контурів об'єктів. Підхід, який запропоновано в проєкті, обґрунтовано синтезує методи штучного інтелекту, теорії комп'ютерного зору, з одного боку, і розпізнавання образів – з іншого; він дає змогу отримати керуючі впливи та математичні функції для прийняття рішень у кожен момент часу з можливістю аналізування впливу зовнішніх факторів та прогнозування перебігу процесів. Детально описано встановлення нейромережі, наведено її характеристики та обґрунтовано її можливості. Висвітлено підходи до комп'ютерного бачення відокремлення об'єктів. Відомі методи нарощування областей, методи, що ґрунтуються на кластеризації, виділення контурів, методи з використанням гістограми. Запропоновано побудувати систему швидкої ідентифікації бойових машин на основі розроблених новітніх методів фільтрації зображень із використанням методів глибинного навчання. Ідентифікація машини потребуватиме на 10–20 % менше часу завдяки розробленій інформаційній технології виявлення об'єктів із використанням швидкозмінної інформації.

**Ключові слова:** YOLO; штучний інтелект; розпізнавання зображень; машинне навчання; ідентифікація; класифікатор.

### Постановка проблеми

Відеоспостереження – поширений засіб вирішення завдань, пов'язаних з безпекою та перебігом різних подій. Одне із головних завдань відеоспостереження – виявлення [1], відстеження [2] та ідентифікація [3] рухомих об'єктів. Однією з найпопулярніших нейромереж є YOLO (з англ. Yolo only look once), яка розпізнає об'єкти в режимі реального часу та відстежує їх переміщення.

Відеокамери розташовані біля нас і записують дані про нас. Тому повсякчас виникає потреба розпізнавати дані та об'єкти та формувати знання про них. Для того, щоб розпізнати дані, потрібно виконати етап попереднього опрацювання, тобто здійснити поліпшення якості зображення – підвищення контрастності, виділення меж, усунення розмитості, фільтрацію. Потім відбуваються операції підготовки графічних зображень – виділення об'єктів, сегментація, виділення контурів.

Поєднання запису відеокамер щодо переміщення воєнних об'єктів та YOLO створює систему, яка не лише відновлює втрачені контури бойових машин, але й ідентифікує об'єкти з високою

точністю. У статті розглянуто способи побудови методу ідентифікації бойових машин, описано методи виділення контурів у режимах сильного зашумлення.

### Аналіз останніх досліджень та публікацій

Щорічно збільшується попит на розпізнавання об'єктів у відеопотоці [4] у режимі реального часу в різних галузях, серед яких геодезія, військові місії, пошукові, розвідувальні, бойові завдання, також із метою аналізу місцевості для виявлення масштабів пошкодження після стихійних лих. Тому виникає потреба забезпечити здатність автоматичного аналізу та опрацювання отриманих візуальних даних, що досліджено у [5].

Особливістю даних для опрацювання ключових кадрів у відеопослідовності, які необхідно враховувати, є зображення з низькою якістю, що розглянуто в [6]. Низьку якість ключових кадрів спричиняють погодні умови, вібрації апаратів та дефекти обладнання тощо. У праці [7] розроблено класифікацію реальних відео і вказано, що зображення можуть суттєво відрізнятися один від одного через умови освітлення, оклюзію, фон, деформацію, зміну масштабу. Неякісні знімки, розглянуті в [8], часто містять шуми та спотворення, які іноді ускладнюють аналіз. Нейромережі можуть бути чутливими до цих шумів і надавати неточні результати.

Пришвидження прогресу згорткових нейронних мереж (CNN), зменшення часу розпізнавання, підвищення середньої точності (mAP), яке сягає 76,4 зі збереженням швидкості зміни кадрів за секунду (FPS) Faster R-CNN 5–18 (тобто повільніше), розглянуто в [9]. Тому застосування згорткових нейронних мереж для відстеження рухомих об'єктів є актуальним.

Проте для навчання нейромереж потрібна велика кількість якісних навчальних даних [10]. Однак у випадку з низькоякісними знімками може бути важко зібрати достатню кількість навчальних прикладів, що може вплинути на якість моделі. Опрацювання низькоякісних зображень потребуватиме більше обчислювальних ресурсів і тривалого часу для навчання. Іншою проблемою, яка виникає під час розпізнавання, є неоднорідність даних, що зазначено в [11]. Якість ключових кадрів може відрізнятися. Наявність нейромережі дасть змогу покращити якість розпізнавання рухомих об'єктів. Для подолання цих проблем важливо використовувати методи покращення якості зображень, попередню обробку даних, вдосконалені архітектури нейромереж.

Стеження – це визначення місцезнаходження рухомих об'єктів [12] або кількох об'єктів у часі за допомогою відеокамери. Алгоритм аналізує відеокадри і надає результат про положення рухомих об'єктів відносно кадру. Зокрема відбувається попереднє оброблення, що передбачає підвищення контрасту, усунення розмитості, створення меж об'єкта та фільтрацію (рис. 1).

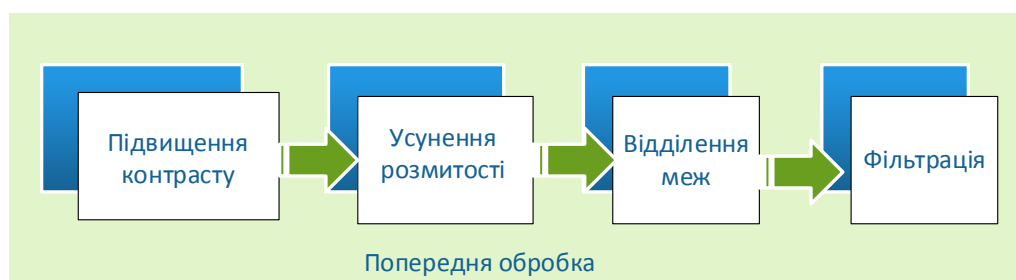


Рис. 1. Схема попереднього опрацювання відеоряду із захопленням об'єкта

Основною проблемою відстеження є зіставлення позицій цільового об'єкта в послідовних кадрах, особливо якщо об'єкт рухається швидше порівняно з частотою кадрів. Тому системи стеження зазвичай використовують модель руху [13], як може змінюватися зображення цільового об'єкта під час різних рухів. Відділення об'єктів один від одного передбачає етапи виділення об'єктів, сегментацію та виділення контурів (рис. 2).

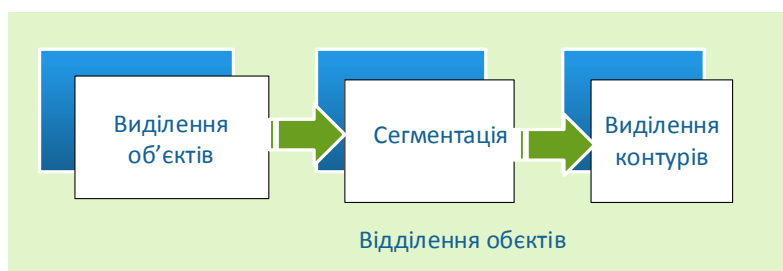


Рис. 2. Схема виділення об'єктів

Приклади таких простих рухів:

- стеження за плоским об'єктом [14];
- афінне перетворення або гомографія зображення об'єкта.

Метою може бути жорсткий тривимірний об'єкт, модель руху якого визначається зовнішнім виглядом залежно від його положення в просторі та орієнтації.

Для стиснення відео ключові кадри розділяють на макроблоки. Модель руху – це серія ключових кадрів, де кожен макроблок трансформується за допомогою вектора руху.

Зображення деформованого об'єкта можна покрити сіткою і рух об'єкта визначатиметься положенням вершин цієї сітки.

Коли об'єкт потрібно шукати та зіставляти із заданим, новий набір ключових точок витягується в тестове зображення, два набори зіставляють та обчислюють оцінку подібності.

### Вибір об'єкта

Для вибору об'єкта з відеопотоку використовують попиксельний, поблочний та методи на основі мінімізації енергетичного функціоналу [15] (рис. 3).

Попиксельні методи виділення об'єктів обробляють всі точки зображення. Вони високоточні, але чутливі до шуму.

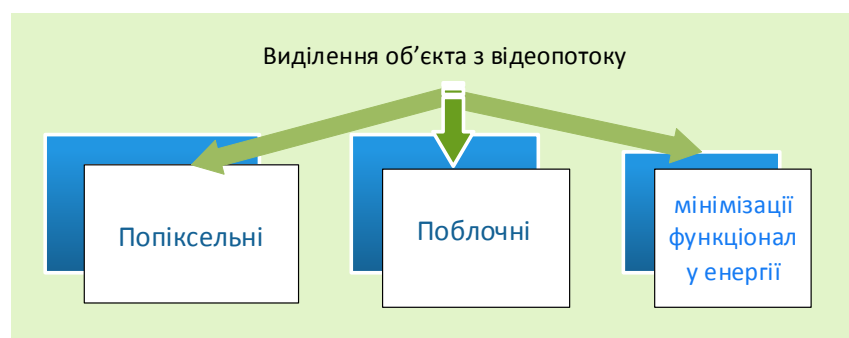


Рис. 3. Способи виділення об'єкта з відеопотоку

Методи перетворення відтінків сірого виконують сегментацію [16] – поділ цифрового зображення на кілька наборів пікселів. Сегментацію зображення зазвичай застосовують для виділення об'єктів і меж. Точніше, сегментація зображення – це призначення таких міток кожному пікселю зображення, що пікселі із однаковими мітками мають спільні візуальні характеристики.

Методи на основі мінімізації енергетичного функціоналу обробляють не окремі пікселі [17], а групи пікселів, об'єднаних у блоки. Якщо блок містить межу, то на таких ділянках межу об'єкта визначають неточно.

Недоліком методів [18] на основі мінімізації енергетичного функціоналу є низька швидкодія, що небажано для їх реалізації.

### Методи нарощування областей

Методи цієї групи основані на використанні локальних особливостей зображення [19]. Ідея методу розширення областей полягає в аналізі спочатку початкової точки, потім сусідніх з нею точок за критерієм однорідності аналізованих точок тієї чи іншої групи. В ефективніших варіантах методу відправною точкою є не окремі пікселі, а поділ зображення на кілька невеликих ділянок. Кожну область потім перевіряють на однорідність, і якщо результат тесту негативний, відповідну область поділяють на менші ділянки.

Порогова сегментація та сегментація [20] за критерієм однорідності на основі середньої яскравості (рис. 4) часто не дають бажаних результатів.



Рис. 4. Способи нарощування областей у відеопотоці даних

Така сегментація зазвичай призводить до значної кількості невеликих регіонів. Найефективніші результати дає сегментація за критерієм однорідності за текстурою [21].

### Виділення контурів

На відео часто спостерігаються неоднорідні об'єкти, тому доводиться стикатися із завданням визначення периметрів, кривизни, факторів форми, питомої поверхні об'єктів тощо. Усі ці завдання пов'язані з аналізом контурних елементів об'єктів.

Методи виділення контурів на зображенні можна розділити на три основні класи:

- 1) методи високочастотної фільтрації [22];
- 2) методи просторової диференціації [23];
- 3) методи функціональної апроксимації [24] (рис. 5).

Спільним для всіх цих методів є визначення границі як області різкого зменшення функції яскравості зображення ( $i$ ), яка вирізняється введеною математичною контурною моделлю.

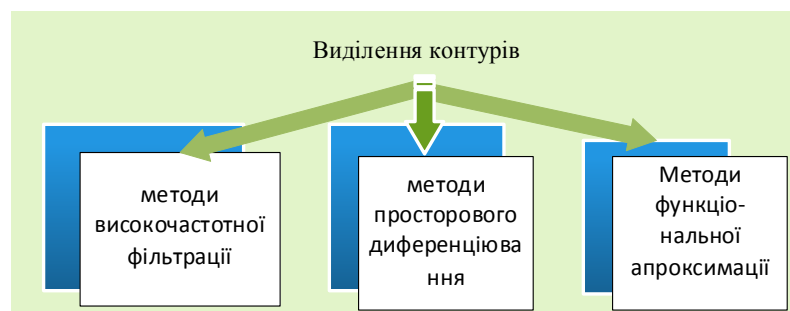


Рис. 5. Способи виділення контурів у відеопотоці

Залежно від завдань до алгоритмів виділення контурів ставлять такі вимоги: виділені контури мають бути тонкими, без розривів і закритими. Виділення контурів ускладнюється необхідністю

застосування алгоритмів потоншення та видалення пропусків. Однак контури, які є не замкнутими, непридатні для процедур аналізу.

### **Методи на основі кластеризації**

Метод *K*-середніх – це ітераційний метод, який використовують для поділу зображення на *K* кластери. Базовий алгоритм наведено нижче [25]:

Крок 1. Виберемо *K* центрів кластерів довільно.

Крок 2. Помістимо кожен піксель зображення у кластер, центр якого найближче до цього пікселя.

Крок 3. Обчислимо центри кластерів, усереднивши усі пікселі у кластері.

Крок 4. Повторимо кроки 2 і 3 до конвергенції, щоб пікселі залишилися в одному кластері.

Відстань зазвичай визначають як суму квадратів та абсолютних значень відмінностей між пікселями і центрами кластерів. Різниця найчастіше залежить від кольору, яскравості, текстури та розташування пікселів або збалансованої комбінації цих факторів.

### **Методи із використанням гістограм**

Методи на основі гістограми [26] є ефективними порівняно з іншими методами сегментації, оскільки потребують одного піксельного проходу.

Гістограму обчислюють одночасно для всіх пікселів зображення, а її мінімуми та максимуми використовують для пошуку кластерів. Для порівняння ефективності методу необхідні колірні характеристики та яскравість.

Метод з використанням гістограми можна вдосконалювати, та використовувати рекурсивно до кластерів на зображенні, щоб розділити їх на менші кластери. Процес поступово повторюється з меншими і меншими кластерами, поки нові кластери не будуть з'являтися.

Методи, основані на використанні гістограм, можна адаптувати для кількох кадрів, зберігаючи перевагу швидкості за один прохід.

### **Мережа YOLO – для виявлення об'єктів**

You-Only-Look-Once (YOLO) [27] – система виявлення відеооб'єктів, яка може працювати в режимі реального часу з дуже високою частотою кадрів, із загальним обмеженням 45 кадрів на секунду та частотою кадрів до 155 кадрів за секунду. YOLO спочатку випустили у 2015 р. з дослідженнями Facebook.

YOLO складається із двох основних частин: детектора класів і детектора рамок. Детектор класу визначає, які об'єкти є на зображенні. Детектор кадрів визначає розташування об'єктів на зображенні. Детектор класу працює за допомогою регресійної нейронної мережі, яка навчається передбачати значення змінної, щоб навчитися передбачати ймовірність наявності певного об'єкта на зображенні.

Детектор класу YOLO – це регресійна нейронна мережа із 24 глибинними шарами. Вхідний рівень мережі отримує зображення розміром 448×448 пікселів. Вихідний рівень мережі містить 84 значення. Кожне значення відповідає ймовірності того, що певний об'єкт є на зображенні.

Цифрові IP-камери використовують у сучасних системах відеоспостереження. Підключення IP-камери до вже локальної мережі потребує мінімальних витрат на встановлення.

Розглянемо характеристики, які необхідно враховувати під час вибору комп'ютерних технологій для цифрової системи відеоспостереження.

Кількість фізичних портів – перша характеристика, за якою можна підключити інші пристрої. Цей параметр може визначати максимальну кількість IP-камер. Для домашньої системи відеоспостереження часто використовують комутатор із чотирма портами, для професійних систем – обладнання з портами 8-16-24 [28].

Пропускна здатність – друга характеристика, що враховує пропускну здатність кожного порту. Найпоширеніші значення – 100 Мбіт/с і 1 Гбіт/с. Сумарна пропускна здатність комутатора може бути

нижчою за сумарне значення всіх портів. Третя характеристика – це швидкість передавання даних, яка обмежує можливість отримання і передавання інформації. Четвертою характеристикою PoE є функція, яка дає змогу жити інші пристрої через кабель, що передає дані. П'ята характеристика – це протоколи управління. Комутатори PoE поділяють на керовані та некеровані. Керовані комутатори – це пристрої, які підтримують низку протоколів (функцій) керування мережею та передавання даних. Щоб побудувати прості та невеликі IP-системи відеоспостереження, які фізично ізольовані від мереж, у яких передаються інші важливі дані (дані телеметрії, банківські та фінансові дані, відеоконференції тощо), можна відмовитися від використання некерованих комутаторів PoE.

### Формулювання цілі статті

Основні цілі цього дослідження:

- Розробити та навчити модель Yolo для виявлення рухомих об'єктів танків.
- Оцінити продуктивність моделі за допомогою стандартних метрик, таких як точність, матриці помилок, точність класифікації та чутливості.
- Дослідити стійкість моделі проти різних типів.
- Розробити програмне забезпечення для відстеження об'єктів у відеопотоці.

Розробити систему швидкої ідентифікації бойових машин на основі нових методів фільтрації з використанням методів глибинного навчання. Проаналізувати вибір математичного апарату для виокремлення контурів рухомих об'єктів у режимах сильного зашумлення. Час, витрачений на ідентифікацію бойової машини, має зменшитись завдяки виявленню рухомих об'єктів в умовах швидкозмінної інформації.

### Виклад основного матеріалу

Розпізнавання об'єктів відбувається (рис. 6) у результаті завантаження відеопослідовності в нейронну мережу та вибору колірної моделі для розпізнавання і налаштування вхідних даних. Паралельно у відеопослідовності здійснюється віфільтрування шумів. Результат – відеопослідовність із розміченими даними, розпізнаними об'єктами.

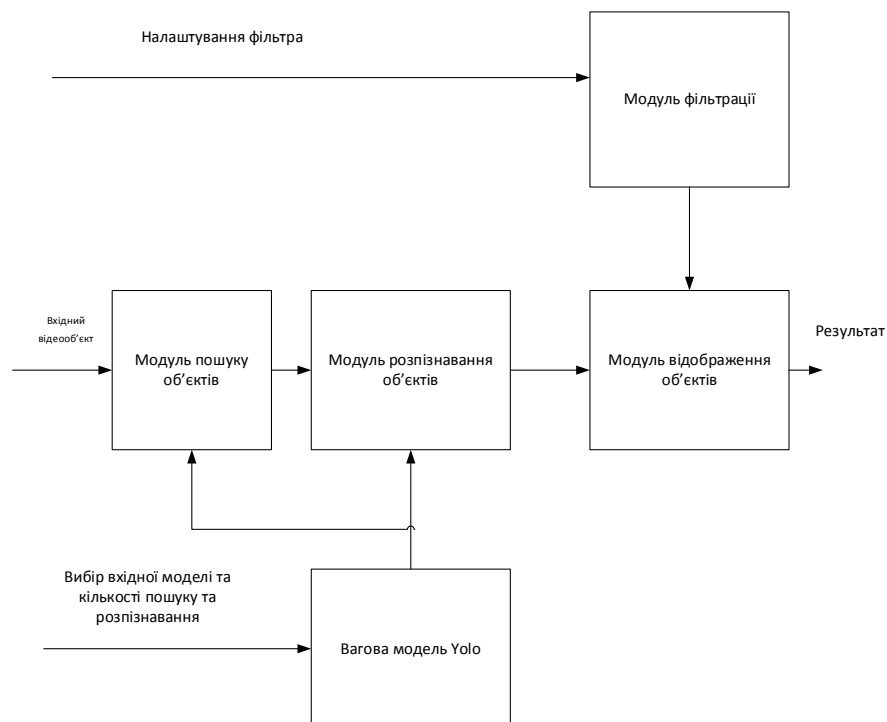


Рис. 6. Метод розпізнавання та виявлення об'єктів у мережі Yolo

## 1. Встановимо відстеження об'єктів у відеопотоці та перевіримо швидкість виявлення об'єктів.

Для цього напишемо код:

```
pip install ultralytics
```

*And then import it into the code:*

```
from ultralytics import YOLO
```

Now everything is ready to create a neural network model:

```
model = YOLO("yolov8m.pt")
```

Модель для розпізнавання відеопослідовності створено та навчено за допомогою PyTorch та експортовано як файли \*.pt. Вперше запустивши цей код, завантажимо файл yolov8m.pt із сервера Ultralytics. Будуємо об'єкт моделі. Тренуємо цю модель, виявляємо об'єкти та експортуємо у відео-послідовність.

Для цього застосовуємо методи:

*train({dataset descriptor file path})* – для навчання моделі на наборі даних зображення;

*predict({image})* – для прогнозування вказаного зображення, наприклад, для виявлення обмежувальних рамок усіх об'єктів, які модель може знайти на зображенні;

*export({format})* – для експорту моделі з формату PyTorch за замовчуванням у вказаний формат.

Усі моделі виявлення об'єктів YOLOv8 уже попередньо навчені на наборі даних COCO, який містить моделі 80 різних типів.

Метод прогнозування приймає багато різних типів вхідних даних, урахувавши шлях до одного зображення, масив шляхів зображення, об'єкт Image з бібліотеки PIL Python тощо.

Коли вхідні дані пройдуть через модель, вона повертає масив результатів для кожного вхідного зображення [29]. Оскільки ми надали лише одне зображення, воно повертає масив з одним елементом, який можна видобути так:

```
result = results[0]
```

Результат містить виявлені об'єкти і зручні властивості для роботи з ними. Найважливішим є масив коробок з інформацією про виявлені обмежувальні рамки на зображенні.

Можна визначити, скільки об'єктів знайдено, запустивши функцію len:

```
len(result.boxes)
```

Після запуску отримано "2", що означає виявлення двох ящиків: один для мобільного телефону, інший для людини. Візьмемо перший об'єкт:

```
box = result.boxes[0]
```

Об'єкт box містить властивості обмежувальної рамки, зокрема:

*xxyy* – координати коробки у вигляді масиву [x1,y1,x2,y2];

*cls* – ідентифікатор типу об'єкта;

*conf* – рівень достовірності моделі щодо цього об'єкта.

Виведемо інформацію про об'єкт:

```
print("Object type:", box.cls)
print("Coordinates:", box.xyxy)
print("Probability:", box.conf)
```

Щодо першого об'єкта отримаємо інформацію:

```
Object type: tensor([16.])
Coordinates: tensor([[261.1901,
Probability: tensor([0.9528])
```

Оскільки YOLOv8 містить моделі PyTorch, виходи з моделей PyTorch кодуються як масив об'єктів PyTorch Tensor, тому потрібно видобути перший елемент із кожного з цих масивів:

```
print("Object type:", box.cls[0])
print("Coordinates:", box.xyxy[0])
print("Probability:", box.conf[0])
```

```
Object type: tensor(16.)
Coordinates: tensor([261.1901,
Probability: tensor(0.9528)
```

Тепер бачимо дані як об'єкти Tensor. Щоб отримати фактичні значення з Tensor, потрібно використовувати метод `.tolist()` для тензорів із масивом усередині та метод `.item()` для тензорів зі скалярними значеннями.

Завантажимо дані у відповідні змінні:

```
coords = box.xyxy[0].tolist()
class_id = box.cls[0].item()
conf = box.conf[0].item()
print("Object type:", class_id)
print("Coordinates:", coords)
print("Probability:", conf)
```

```
Object type: 16.0
Coordinates: [261.1900634765625,
Probability: 0.9528293609619141
```

Отримуємо реальні дані, координати яких можна округлити, ймовірність також можна округлити до двох знаків після коми. Усі об'єкти, які може виявити нейронна мережа, мають цифрові ідентифікатори. У випадку попередньо підготовленої моделі YOLOv8 існує 80 типів функцій із ідентифікаторами від 0 до 79. Класи функцій COCO є відкритими. Крім того, об'єкт результату YOLOv8 містить зручну властивість імен для отримання цих класів.

#### Дані для запропонованої моделі

Вебзастосунок складається з двох основних файлів:

Flaskapp.py – файл, який відповідає за сам проєкт, його вигляд та структуру.

YOLO\_Video.py – файл, який відповідає за алгоритм YOLO, а саме за реалізацію розпізнавання об'єктів у відеопотоці.

Реалізація файлу Flaskapp.py:

*Налаштування Flask додатка:*

Створюємо вебдодаток за допомогою класу Flask.

Встановлюємо конфігураційні параметри, такі як секретний ключ та тека для завантаження файлів.

Визначаємо клас UploadFileForm за допомогою Flask-WTF для обробки завантажень файлів.

*Функції обробки відео:*

Визначаємо функції `generate_frames` та `generate_frames_web` для генерації кадрів на основі виведення з детекції YOLO.

Ці функції використовують функцію `video_detection` з файлу `YOLO_Video.py` для виконання детекції об'єктів на кадрах відео.

Визначаємо маршрути для домашньої сторінки (`/` та `/home`), сторінки вебкамери (`/webcam`) та сторінки завантаження відео (`/FrontPage`).



Маршрути /video та /webapp відповідають за трансляцію кадрів відео з результатами детекції об'єктів. Маршрути webcam та front відтворюють HTML-шаблони для сторінок вебкамери та завантаження відео. Клас UploadFileForm використовуємо для опрацювання завантажень відеофайлів. Додаток запускаємо на розробницькому сервері, якщо скрипт виконується безпосередньо.

### Реалізація файлу YOLO\_Video.py

Функція video\_detection приймає відеошлях як вхідні дані та виконує виявлення об'єкта за допомогою моделі YOLO.

Моделю YOLO від Ultralytics завантажується з указанного файлу контрольної точки (yolov8n.pt).

Визначаються кадри виявлених об'єктів на кожному кадрі, і повертаються оброблені кадри.

Імена класів, що відповідають виявленим об'єктам, визначені в списку classNames.

*Знімання та опрацювання відео:*

OpenCV використовується для захоплення відеокadrів із зазначеного шляху.

У кожному кадрі намальовано виявлені об'єкти разом із мітками класів і рівнями достовірності.

Опрацьовані кадри повертаються для потокового передавання в браузері користувача.

*Загальний хід роботи:*

Користувач завантажує відеофайл через інтерфейс.

Файл зберігається, а його шлях зберігається в сеансі Flask.

Детектор об'єктів викликається з отриманої відеопослідовності.

Опрацьовані відеокадри передаються для перегляду в реальному часі через браузер користувача.

Цей проєкт використовує Flask для вебпрограми та інтегрує YOLO для опрацювання відео в реальному часі. Файл YOLO\_Video.py ізолює функції, пов'язані з YOLO, роблячи його модульним і придатним для повторного використання.

Моделю основана на попередньо навченому OSFA та створена на основі PyTorch. Розмір тренувального зображення досяг  $256 \times 128$ . Розмір пакета з 64 випадково вибраних даних потім подавали в мережу. Під час тестування тестові зображення також змінюються до  $256 \times 128$ . Моделю навчено на 100 епохах. Значення  $\alpha_1$ ,  $\alpha_2$  та швидкості навчання такі самі, як встановлені OSFA.  $\alpha_1$ ,  $\alpha_2$  та швидкість навчання встановлені на 1, 0,0007,  $3,5 \times 10^{-5}$  відповідно. У SAM чотири горизонтальні частини. Усі експерименти виконано з апаратним середовищем Intel(R) Core(TM) i7-11800H 11-го покоління на 2,30 ГГц і NVIDIA GeForce RTX 3060. Результати точності розпізнавання подано у таблиці. Розпізнавання здійснено на зображеннях  $512 \times 512$ ,  $640 \times 640$ ,  $1024 \times 1024$  px. Досягнуто точності розпізнавання 96 %, результати показано на рис. 7–14.

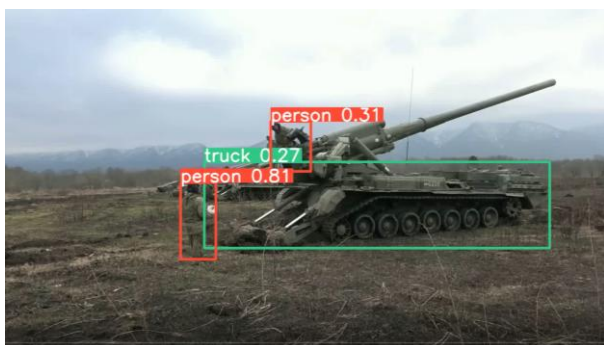


Рис. 7. Результати розпізнавання та ідентифікації двох осіб та вантажівки у відеоряді YOLO

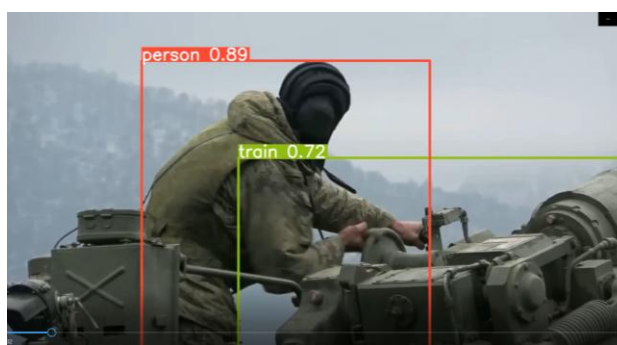


Рис. 8. Результати розпізнавання та ідентифікації особи та відеоряд YOLO

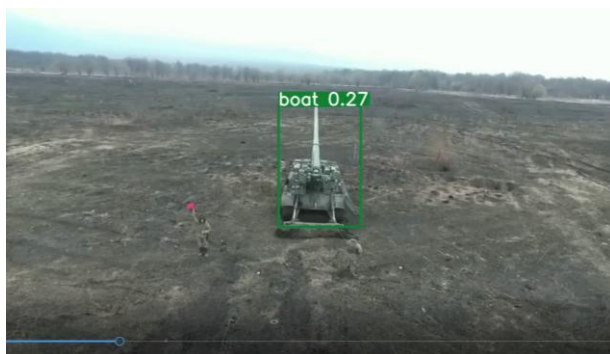


Рис. 9. Результати розпізнавання та ідентифікації об'єктів відеоряду YOLO



Рис. 10. Результати розпізнавання людини та танка у відеоряду YOLO

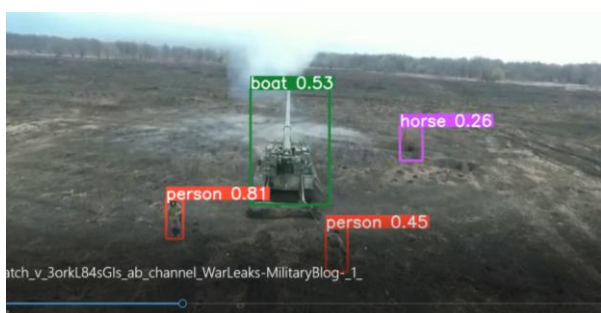


Рис. 11. Результати розпізнавання двох людей, танк у відеоряду YOLO



Рис. 12. Результати розпізнавання та ідентифікації об'єктів відеоряду YOLO



Рис. 13. Результати розпізнавання та ідентифікації трьох об'єктів відеоряду YOLO



Рис. 14. Результати розпізнавання та ідентифікації трьох об'єктів відеоряду YOLO

### Результати навчання, отримані за допомогою технік машинного навчання

Optimizer	Precision	Recall	map@0.5	map@0.95	F1 score
512×512					
Adam W	0,340	0,873	0,352	0,248	0,488
Adam	0,835	0,842	0,811	0,489	0,839
SGD	0,965	0,950	0,977	0,200	0,950
640×640					
Adam W	0,930	0,810	0,878	0,681	0,865
Adam	0,821	0,765	0,788	0,433	0,792
<b>SGD</b>	<b>0,969</b>	<b>0,960</b>	<b>0,982</b>	<b>0,833</b>	<b>0,964</b>

Продовження табл.

Optimizer	Precision	Recall	map@0.5	map@0.95	F1 score
1024×1024					
Adam W	0,921	0,722	0,808	0,622	0,810
Adam	0,820	0,764	0,787	0,432	0,791
SGD	0,963	0,967	0,972	0,825	0,965

### Висновки

Проаналізовано алгоритм розпізнавання об'єктів YOLOv8 та його відмінності від інших алгоритмів машинного навчання. Створено, проаналізовано та протестовано вебзастосунок для розпізнавання об'єктів у відеопотоці. Подано детальний огляд інструментів розробки для веб-програми із використанням YOLOv8. Досліджено середовище програмування PyCharm, фреймворк Flask та його переваги порівняно з іншими фреймворками. Розглянуто також Ultralytics, яка допомагає в розробленні та тестуванні вебдодатків для потокового відео.

Досліджено характеристики нейромережі та її можливості. Серед відомих методів комп'ютерного бачення відокремлення об'єктів – методи нарощування областей, методи, основані на кластеризації, виділення контурів, методи з використанням гістограми. Побудовано систему швидкої ідентифікації бойових машин із використанням методів глибокого навчання.

Досягнуто показників точності розпізнавання 96 %.

### Список літератури

1. Лю Дж., Се Г., Ван Дж., Лі С., Ван Ч., Чжен Ф. та Цзінь Ю. (2024). Глибоке промислове виявлення аномалій зображення: опитування. *Дослідження машинного інтелекту*, 21 (1), 104–135. <https://doi.org/10.1007/s11633-023-1459-z>
2. Крюгер-Маре, Е. (2024). Субтитри для засвоєння мови: стеження за очима як предиктор розподілу уваги в освіті. *Міжнародний журнал мовних досліджень*, 18(2). DOI: 10.1007/s11633-023-1459-z
3. Лі П., Чжан Ю., Юань Л., Сяо Х., Лін Б. та Сю Х. (2024). Ефективна довго-коротка тимчасова мережа уваги для неконтрольованої сегментації відеооб'єктів. *Розпізнавання образів*, 146, 110078. DOI: 10.48550/arXiv.2309.11707
4. Ladonia, M. S. Дослідження впливу значення порогу Non-Maximal Suppression на здатність YOLO до розпізнавання об'єктів на зображеннях низької якості. *Problems of Informatization and Management*, 2(74), 68–73. <https://doi.org/10.18372/2073-4751.74.17884>
5. Weber E., Vedaldi A., Bischof H., Brox T., Frahm J. M. (2020). Detecting Natural Disasters, Damage, and Incidents in the Wild. *Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science*, Vol. 12364, 331–350.
6. Laroca, R., Severo, E., Zanlorensi, L. A., Oliveira, L. S., Gonçalves, G. R., Schwartz, W. R., & Menotti, D. (2018, July). A robust real-time automatic license plate recognition based on the YOLO detector. In *2018 international joint conference on neural networks (ijcnn)* (pp. 1–10). IEEE.
7. Shinde, S., Kothari, A., & Gupta, V. (2018). YOLO based human action recognition and localization. *Procedia computer science*, 133, 831–838.
8. Shinde, S., Kothari, A., & Gupta, V. (2018). YOLO based human action recognition and localization. *Procedia computer science*, 133, 831–838.
9. Du, J. (2018, April). Understanding of object detection based on CNN family and YOLO. In *Journal of Physics: Conference Series* (Vol. 1004, p. 012029). IOP Publishing.
10. Li, Y., Zhao, Z., Luo, Y., & Qiu, Z. (2020). Real-time pattern-recognition of GPR images with YOLO v3 implemented by tensorflow. *Sensors*, 20(22), 6476. <https://doi.org/10.3390/s20226476>
11. Chen, H., He, Z., Shi, B., & Zhong, T. (2019). Research on recognition method of electrical components based on YOLO V3. *IEEE Access*, 7, 157818–157829.

12. Кавамура, Е., Каннан, К., Ломбертс, Т., Степанян, В., Дольф, К., & Іпполіто, К. А. (2024). Наземний відстежувач бачення для вдосконаленої повітряної мобільності та міської повітряної мобільності. На форумі AIAA SciTech 2024 (С. 2010). DOI: 10.2514/6.2024-2010
13. Ін, З., Ні, Ю., Лі, Л., Ван, Т., Ву, Дж., Лі, З., і Тан, Д. (2024). Чисельне моделювання та експериментальне дослідження двофазного стокового вихору та його коливальних характеристик рідина-тверде тіло. *Журнал Zhejiang University-SCIENCE A*, 25(1), 47–62. DOI: 10.1631/jzus.A2200014
14. Вольф Т., Фрідович-Кейл Д. та Джонс Б. А. (2024). Взаємне інформаційне планування траєкторії для стеження за цисюнними космічними об'єктами з використанням послідовної конвексифікації. На форумі AIAA SCITECH 2024 (С. 0626). doi.org/10.2514/6.2024-0626
15. Чжан Х., Чжен Д., Чжан Ю., Цао Дж., Лін В. та Лінг В. К. (2024). Оцінка якості для представлень, синтезованих DIBR, на основі вейвлет-перетворення та подібності величини градієнта. *Транзакції IEEE щодо мультимедіа*. DOI: 10.1109/TMM.2024.3356029.
16. Qin, Q., & Chen, Y. (2024). Огляд сегментації судин сітківки для аналізу зображення очного дна. *Інженерне застосування штучного інтелекту*, 128, 107454. <https://doi.org/10.1016/j.engappai.2023.107454>
17. Li, M., Cui, Q., Wang, X., Zhang, Y., & Xiang, Y. Ftpе-Вс: Швидке шифрування зображень зі збереженням ескізів за допомогою Block-Churning. Доступно за: SSRN 4698446 .doi.org/10.2139/ssrn.4698446
18. Клейтон-Чабб, Д., Кемп, В. В., Маджид, А., Любель, Дж. С., Вудс, Р. Л., Тран, К., ... і Робертс, С. К. (2024). Стеатотична хвороба печінки, пов'язана з метаболічною дисфункцією, у людей похилого віку пов'язана зі слабкістю та соціальною несприятливістю. *Liver International*, 44 (1), 39–51. <https://doi.org/10.1111/liv.15725>
19. Міра, Е. С., Сапрі, А. М. С., Алехані, Р. Ф., Джамбі, Б. С., Башир, Т., Ель-Кенаві, Е. С. М., і Сабер, М. (2024). Рання діагностика раку ротової порожнини за допомогою обробки зображень і штучного інтелекту. *Fusion: Practice and Applications*, 14(1), 293–308.
20. Цяо, Л., Лю, К., Сюе, Ю., Тан, В., і Салехнія, Т. (2024). Багаторівневий пороговий метод сегментації зображення з використанням гібридної арифметичної оптимізації та алгоритмів оптимізатора Harris Hawks. *Експертні системи з додатками*, 241, 122316. <https://doi.org/10.1016/j.eswa.2023.122316>
21. Ван, З., Ден, Ю., Чжан, Ю., Тан, Х., Чжоу, П., Лі, П., ... і Чжан, М. (2024). Осередковані волокнистими сироватковими білками однорідні емульсійні гелі з м'якою текстурою для людей похилого віку: підвищення біодоступності куркуміну. *Харчова хімія*, 437, 137850. <https://doi.org/10.1016/j.foodchem.2023.137850>
22. Гао Дж. та Хуан Ю. (2024). FP-Net: мережа частотного сприйняття з змагальним навчанням для локалізації маніпуляції зображеннями. *Мультимедійні засоби та програми*, 1–19. <https://doi.org/10.1007/s11042-023-17914-1>
23. Су, Ю., Тан, В., Донг, Ю., Сю, В., Хуан, П., Чжан, Дж., і Чжан, Д. (2024). Покращення виявлення прихованих об'єктів на активних міліметрових зображеннях за допомогою вейвлет-перетворення. *Обробка сигналів*, 216, 109303. <https://doi.org/10.1016/j.sigpro.2023.109303>
24. Бхандарі, Дж., і Руссо, Д. (2024). Гарантії глобальної оптимальності для методів градієнта політики. *Дослідження операцій*. <https://doi.org/10.1287/opre.2021.0014>
25. Цянь К. та Дуань Х. К. (2024). Платформа для оптичного підрахунку личинок креветок з використанням замаскованих k-середніх і фільтра бічного вікна. *Прикладна оптика*, 63(6), A7–A15. <https://doi.org/10.1364/AO.502868>
26. Лі С., Кім Дж., Бе П., Лі С. та Кім Х. (2024). Надійний метод аналізу зображення на основі гістограми інтенсивності для флуоресцентного імунологічного аналізу на основі кульок. *BioChip Journal*, 1–9. <https://doi.org/10.1007/s13206-023-00137-9>
27. Редмон, Дж., Діввала, С., Гіршик, Р., і Фархаді, А. (2016). Ви дивитесь лише один раз: уніфіковане виявлення об'єктів у реальному часі. У матеріалах конференції IEEE з комп'ютерного зору та розпізнавання образів. С. 779–788.
28. Чжан Т., Чоудгері А., Бахл П., Джеймсон К. та Банерджі С. (2015, вересень). Проектування та впровадження бездротової системи відеоспостереження. У матеріалах 21-ї щорічної міжнародної конференції з мобільних комп'ютерів і мереж. С. 426–438. <https://doi.org/10.1145/2789168.2790123>

29. Назаркевич М., Логойда М., Троян О., Возний Ю., Шпак З. (2019, вересень). Фільтр ateb-gabog для зняття відбитків пальців. На конференції з інформатики та інформаційних технологій. С. 247–255. Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-030-33695-0\\_18](https://doi.org/10.1007/978-3-030-33695-0_18)

### References

1. Liu, J., Xie, G., Wang, J., Li, S., Wang, C., Zheng, F., & Jin, Y. (2024). Deep industrial image anomaly detection: A survey. *Machine Intelligence Research*, 21(1), 104–135. <https://doi.org/10.1007/s11633-023-1459-z>
2. Kruger-Marais, E. (2024). Subtitling for language acquisition: Eye tracking as predictor of attention allocation in education. *International Journal of Language Studies*, 18(2). DOI: 10.1007/s11633-023-1459-z
3. Li, P., Zhang, Y., Yuan, L., Xiao, H., Lin, B., & Xu, X. (2024). Efficient long-short temporal attention network for unsupervised video object segmentation. *Pattern Recognition*, 146, 110078. DOI: 10.48550/arXiv.2309.11707
4. Ladonia, M. S. Дослідження впливу значення порогу Non-Maximal Suppression на здатність YOLO до розпізнавання об'єктів на зображеннях низької якості. *Problems of Informatization and Management*, 2(74), 68–73. <https://doi.org/10.18372/2073-4751.74.17884>
5. Weber E., Vedaldi A., Bischof H., Brox T., Frahm J. M. (2020). Detecting Natural Disasters, Damage, and Incidents in the Wild. Computer Vision – ECCV 2020. ECCV 2020. *Lecture Notes in Computer Science*. Vol. 12364, 331–350.
6. Laroca, R., Severo, E., Zanlorensi, L. A., Oliveira, L. S., Gonçalves, G. R., Schwartz, W. R., & Menotti, D. (2018, July). A robust real-time automatic license plate recognition based on the YOLO detector. In *2018 international joint conference on neural networks (ijcnn)*, 1–10. IEEE.
7. Shinde, S., Kothari, A., & Gupta, V. (2018). YOLO based human action recognition and localization. *Procedia computer science*, 133, 831–838.
8. Shinde, S., Kothari, A., & Gupta, V. (2018). YOLO based human action recognition and localization. *Procedia computer science*, 133, 831–838.
9. Du, J. (2018, April). Understanding of object detection based on CNN family and YOLO. In *Journal of Physics: Conference Series* (Vol. 1004, p. 012029). IOP Publishing.
10. Li, Y., Zhao, Z., Luo, Y., & Qiu, Z. (2020). Real-time pattern-recognition of GPR images with YOLO v3 implemented by tensorflow. *Sensors*, 20(22), 6476. <https://doi.org/10.3390/s20226476>
11. Chen, H., He, Z., Shi, B., & Zhong, T. (2019). Research on recognition method of electrical components based on YOLO V3. *IEEE Access*, 7, 157818–157829.
12. Kawamura, E., Kannan, K., Lombaerts, T., Stepanyan, V., Dolph, C., & Ippolito, C. A. (2024). Ground-Based Vision Tracker for Advanced Air Mobility and Urban Air Mobility. In *AIAA SciTech 2024 Forum* (p. 2010). DOI: 10.2514/6.2024-2010
13. Yin, Z., Ni, Y., Li, L., Wang, T., Wu, J., Li, Z., & Tan, D. (2024). Numerical modeling and experimental investigation of a two-phase sink vortex and its fluid-solid vibration characteristics. *Journal of Zhejiang University-SCIENCE A*, 25(1), 47–62. DOI: 10.1631/jzus.A2200014
14. Wolf, T., Fridovich-Keil, D., & Jones, B. A. (2024). Mutual Information-Based Trajectory Planning for Cislunar Space Object Tracking using Successive Convexification. In *AIAA SCITECH 2024 Forum* (p. 0626). [doi.org/10.2514/6.2024-0626](https://doi.org/10.2514/6.2024-0626)
15. Zhang, H., Zheng, D., Zhang, Y., Cao, J., Lin, W., & Ling, W. K. (2024). Quality Assessment for DIBR-synthesized Views based on Wavelet Transform and Gradient Magnitude Similarity. *IEEE Transactions on Multimedia*. DOI: 10.1109/TMM.2024.3356029.
16. Qin, Q., & Chen, Y. (2024). A review of retinal vessel segmentation for fundus image analysis. *Engineering Applications of Artificial Intelligence*, 128, 107454. <https://doi.org/10.1016/j.engappai.2023.107454>
17. Li, M., Cui, Q., Wang, X., Zhang, Y., & Xiang, Y. Ftpe-Bc: Fast Thumbnail-Preserving Image Encryption Using Block-Churning. Available at SSRN 4698446. [doi.org/10.2139/ssrn.4698446](https://doi.org/10.2139/ssrn.4698446)
18. Clayton-Chubb, D., Kemp, W. W., Majeed, A., Lubel, J. S., Woods, R. L., Tran, C., ... & Roberts, S. K. (2024). Metabolic dysfunction-associated steatotic liver disease in older adults is associated with frailty and social disadvantage. *Liver International*, 44(1), 39–51. <https://doi.org/10.1111/liv.15725>

19. Mira, E. S., Sapri, A. M. S., Aljehani, R. F., Jambi, B. S., Bashir, T., El-Kenawy, E. S. M., & Saber, M. (2024). Early Diagnosis of Oral Cancer Using Image Processing and Artificial Intelligence. *Fusion: Practice and Applications*, 14(1), 293–308.
20. Qiao, L., Liu, K., Xue, Y., Tang, W., & Salehnia, T. (2024). A multi-level thresholding image segmentation method using hybrid Arithmetic Optimization and Harris Hawks Optimizer algorithms. *Expert Systems with Applications*, 241, 122316. <https://doi.org/10.1016/j.eswa.2023.122316>
21. Wang, Z., Deng, Y., Zhang, Y., Tang, X., Zhou, P., Li, P., ... & Zhang, M. (2024). Fibrous whey protein mediated homogeneous and soft-textured emulsion gels for elderly: Enhancement of bioaccessibility for curcumin. *Food Chemistry*, 437, 137850. <https://doi.org/10.1016/j.foodchem.2023.137850>
22. Gao, J., & Huang, Y. (2024). FP-Net: frequency-perception network with adversarial training for image manipulation localization. *Multimedia Tools and Applications*, 1–19. <https://doi.org/10.1007/s11042-023-17914-1>
23. Su, Y., Tan, W., Dong, Y., Xu, W., Huang, P., Zhang, J., & Zhang, D. (2024). Enhancing concealed object detection in Active Millimeter Wave Images using wavelet transform. *Signal Processing*, 216, 109303. <https://doi.org/10.1016/j.sigpro.2023.109303>
24. Bhandari, J., & Russo, D. (2024). Global optimality guarantees for policy gradient methods. *Operations Research*. <https://doi.org/10.1287/opre.2021.0014>
25. Qian, K., & Duan, H. C. (2024). Optical counting platform of shrimp larvae using masked k-means and a side window filter. *Applied Optics*, 63(6), A7–A15. <https://doi.org/10.1364/AO.502868>
26. Lee, S., Kim, J., Bae, P., Lee, S., & Kim, H. (2024). Intensity Histogram-Based Reliable Image Analysis Method for Bead-Based Fluorescence Immunoassay. *BioChip Journal*, 1–9. <https://doi.org/10.1007/s13206-023-00137-9>
27. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, 779–788.
28. Zhang, T., Chowdhery, A., Bahl, P., Jamieson, K., & Banerjee, S. (2015, September). The design and implementation of a wireless video surveillance system. In Proceedings of the 21st Annual International Conference on Mobile Computing and Networking, 426–438. <https://doi.org/10.1145/2789168.2790123>
29. Nazarkevych, M., Logoyda, M., Troyan, O., Vozniy, Y., & Shpak, Z. (2019, September). The ateb-gabor filter for fingerprinting. In Conference on Computer Science and Information Technologies, 247–255. Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-030-33695-0\\_18](https://doi.org/10.1007/978-3-030-33695-0_18)
30. Boyko, N., Kuba, M., Mochurad, L., & Montenegro, S. (2019). Fractal Distribution of Medical Data in Neural Network. In IDDM, 307–318.

## METHOD OF IDENTIFICATION OF COMBAT VEHICLES BASED ON YOLO

Mariia Nazarkevych<sup>1</sup>, Nazar Oleksiv<sup>2</sup>

Lviv Polytechnic National University,

Information Systems and Networks Department, Lviv, Ukraine

<sup>1</sup>E-mail: [mariia.a.nazarkevych@lpnu.ua](mailto:mariia.a.nazarkevych@lpnu.ua), ORCID 0000-0002-6528-9867

<sup>2</sup>E-mail: [nazar.oleksiv.mnsa.2020@lpnu.ua](mailto:nazar.oleksiv.mnsa.2020@lpnu.ua), ORCID 0000-0001-7821-3522

© Nazarkevych M., Oleksiv N., 2024

**A method for recognizing contours of objects in a video data stream is proposed. Data will be uploaded using a video camera in real time and object recognition will be performed. We will use the YOLO network – a method of identifying and recognizing objects in real time. Recognized objects will be recorded in a video sequence showing the contours of the objects. The approach**

proposed in the project reasonably synthesizes methods of artificial intelligence, theories of computer vision on the one hand, and pattern recognition on the other; it makes it possible to obtain control influences and mathematical functions for decision-making at every moment of time with the possibility of analyzing the influence of external factors and forecasting the flow of processes, and refers to the fundamental problems of mathematical modeling of real processes. The installation of the neural network is shown in detail. The characteristics of the neural network are shown and its capabilities are substantiated. Approaches to computer vision for object extraction are shown. Well-known methods are methods of expanding areas, methods based on clustering, contour selection, and methods using a histogram. The work envisages building a system for rapid identification of combat vehicles based on the latest image filtering methods developed using deep learning methods. The time spent on identifying the machine will be 10–20 % shorter, thanks to the developed new information technology for detecting objects in conditions of rapidly changing information.

**Key words:** YOLO; artificial intelligence; image recognition; machine learning; identification; classifier.