

Магістерська кваліфікаційна робота

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ІДЕНТИФІКАЦІЇ ПІШОХОДІВ У ВІДЕОПОТОЦІ НА ОСНОВІ ЗГОРТКОВИХ НЕЙРОМЕРЕЖ

Виконав: ст. гр. 2КН-19м Радченко Д. Ю.

Керівник: к.т.н., доц. Колесницький О.К.

МЕТА ТА ЗАВДАННЯ ДОСЛІДЖЕННЯ

Метою дослідження магістерської кваліфікаційної роботи є підвищення достовірності ідентифікації пішоходів у відеопотоці програмними засобами за рахунок застосування згорткових нейронних мереж.

Для досягнення поставленої мети необхідно розв'язати такі завдання:

- провести аналіз проблеми розв'язання задачі ідентифікації пішоходів у відеопотоці;
- розглянути існуючі методи вирішення задачі ідентифікації пішоходів у відеопотоці та обрати й обґрунтувати вибір методу, який задовольняє мету даної магістерської кваліфікаційної роботи;
- розробити інформаційну технологію ідентифікації пішоходів у відеопотоці;
- розробити структуру та алгоритм роботи програмного засобу;
- виконати програмну реалізацію запропонованої інформаційної технології;
- провести тестування програмного продукту та виконати аналіз отриманих результатів..

ОБ'ЄКТ, ПРЕДМЕТ ТА МЕТОДИ ДОСЛІДЖЕННЯ

Об'єкт дослідження – процес ідентифікації пішоходів у відеопотоці з використанням згорткових нейронних мереж.

Предмет дослідження – інформаційна технологія та програмні засоби ідентифікації пішоходів у відеопотоці з використанням згорткових нейронних мереж та достовірність їх роботи.

Методи дослідження

У роботі використані наступні методи наукових досліджень:

- системного аналізу,
- розпізнавання образів,
- теорії штучних нейронних мереж для реалізації інформаційної технології ідентифікації пішоходів у відеопотоці,
- методи математичної статистики для розробки процесу ідентифікації пішоходів у відеопотоці та обрахунків результатів експериментів із програмним засобом,
- об'єктно-орієнтованого програмування.

НАУКОВА НОВИЗНА ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

1. Набула подальшого розвитку інформаційна технологія ідентифікації пішоходів у відеопотоці, яка відрізняється використанням згорткової нейронної мережі, що дозволило підвищити достовірність ідентифікації пішоходів у відеопотоці.
2. Вдосконалено процес навчання згорткової нейронної мережі, який відрізняється введенням аугментації тренувальних даних та зміною розміру вікна передбачення об'єктів, що дозволило підвищити достовірність розпізнавання нейронної мережі.

ПРАКТИЧНЕ ЗНАЧЕННЯ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

1. розроблено алгоритм роботи програмного забезпечення ідентифікації пішоходів у відеопотоці на основі згорткової штучної нейронної мережі;
2. розроблено програмні засоби для ідентифікації пішоходів у відеопотоці на основі згорткової штучної нейронної мережі;

Аналіз предметної області ідентифікації пішоходів у відеопотоці

Основними методами розв'язання задачі ідентифікації пішоходів у відеопотоці є:

- ❑ Метод локальних бінарних шаблонів;
- ❑ Гістограма орієнтованих градієнтів;
- ❑ Нейронні мережі.

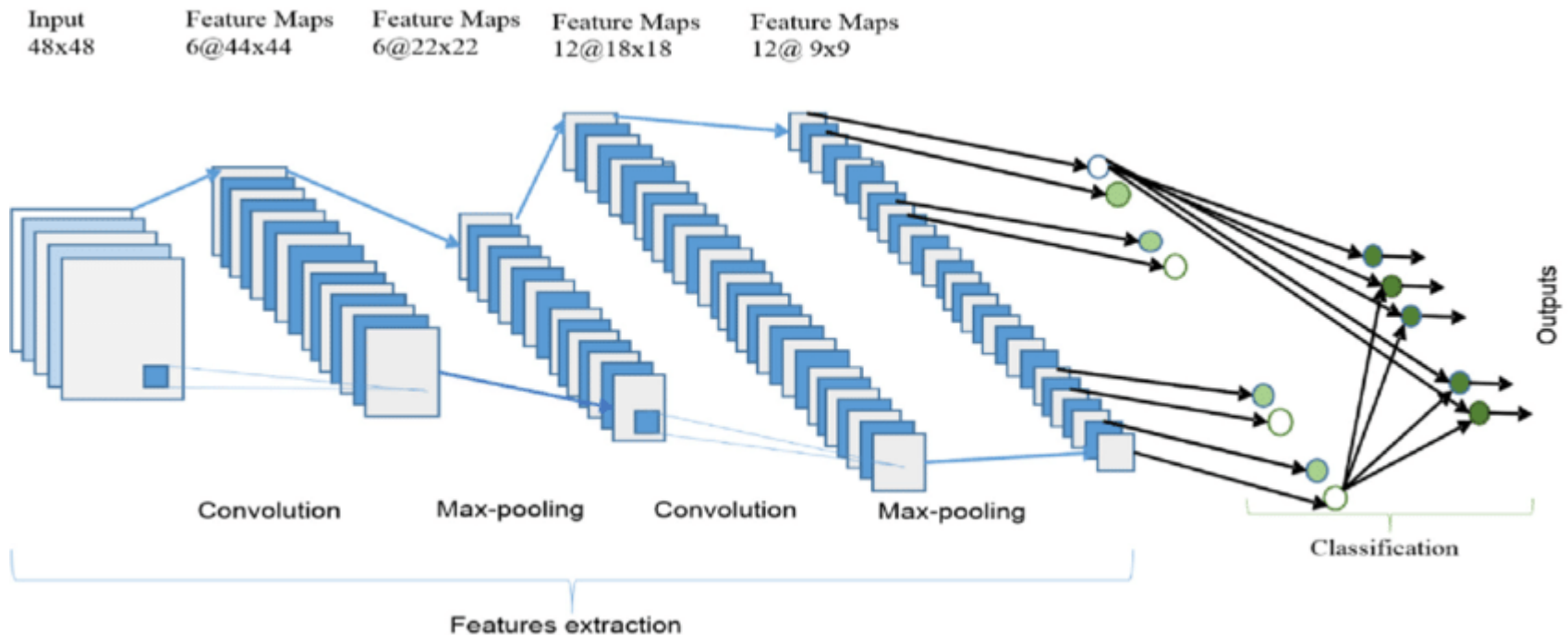
Було обрано метод на основі згорткових нейронних мереж

Вибір і обґрунтування аналогу

Програма RetinaNet:
побудована на використанні
пірамідальної нейронної
мережі, яка працює із
попередньо отриманими
ознаками зображення



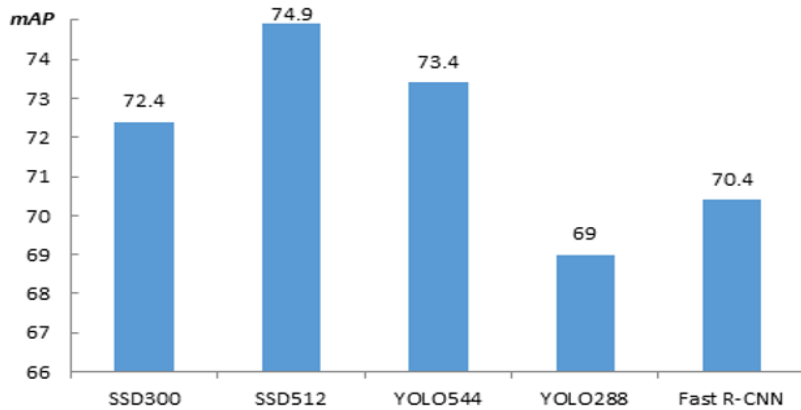
Обґрунтування вибору згорткової нейронної мережі для розв'язання задачі ідентифікації пішоходів у відеопотоці



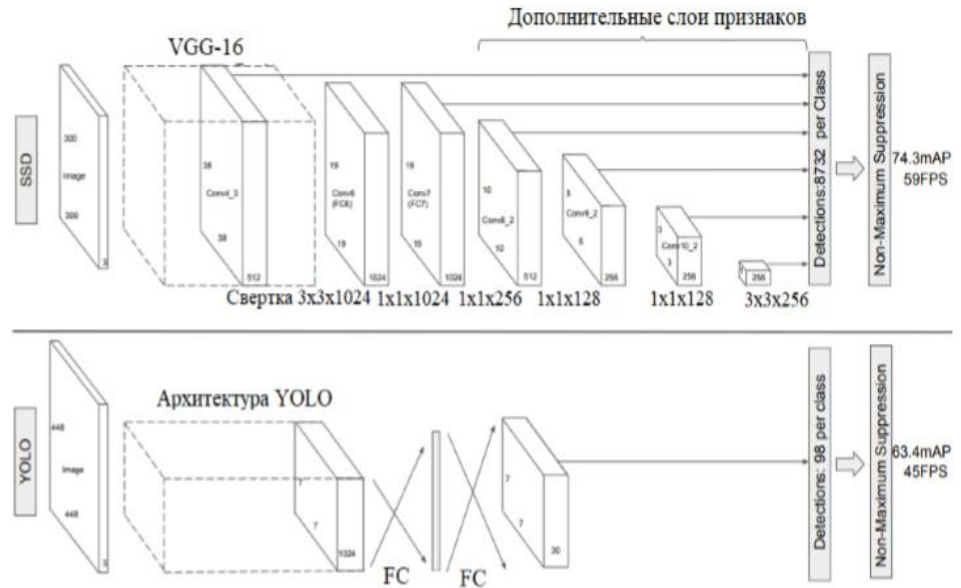
Ідея даних моделей мотивована дослідженнями про зорову кору головного мозку.

Обробка зображень в мозку влаштована як глибока нейронна мережа. Це наближення згорткових нейронних мереж до роботи зорової кори головного мозку дозволяє їм розпізнавати об'єкти на більш високому рівні абстракції з високою точністю

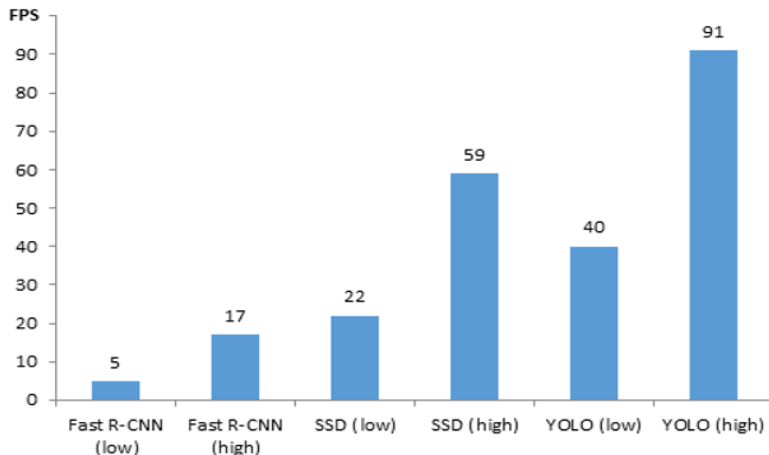
Обґрунтування вибору моделі згорткової нейронної мережі



Порівняння точності різних моделей

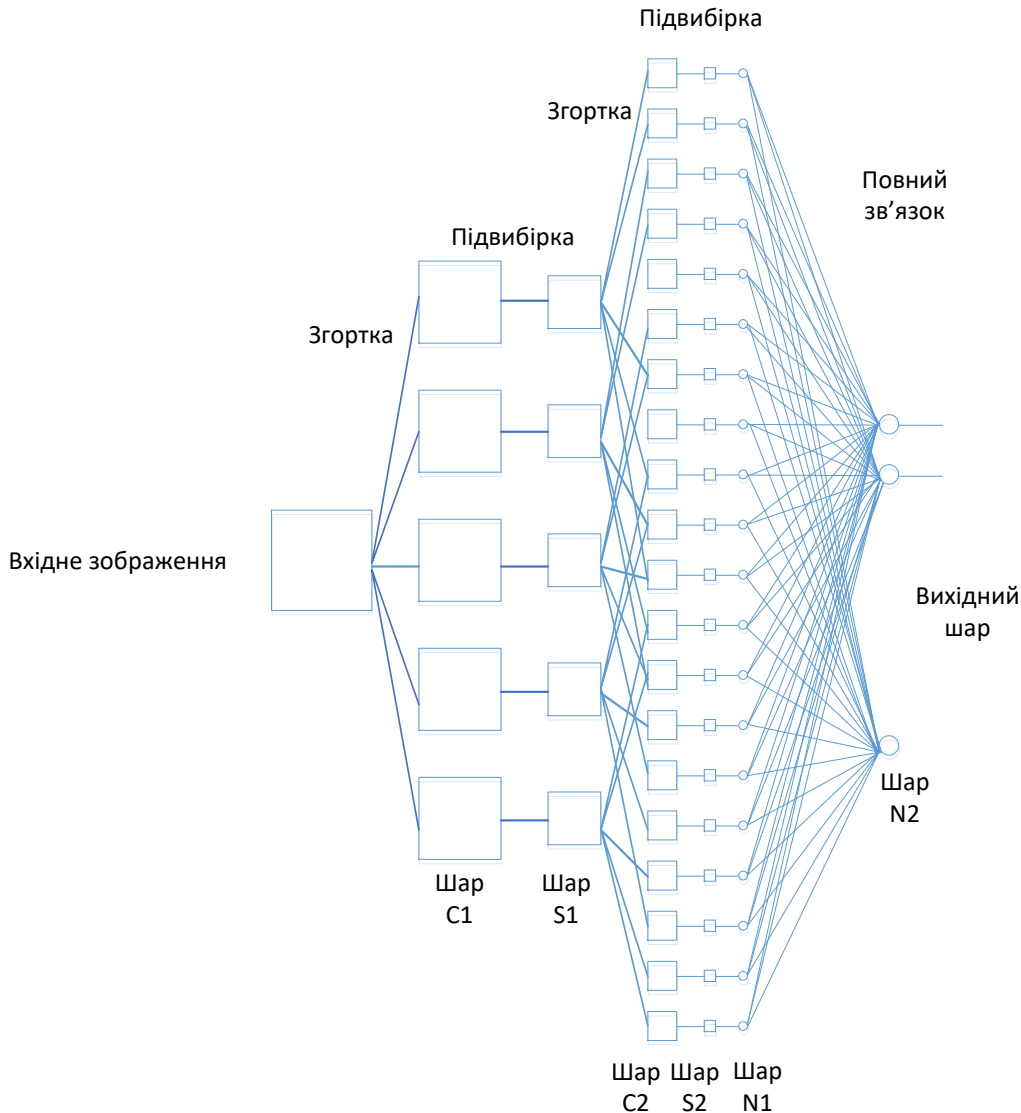


Найбільшу продуктивність забезпечує модель YOLO, але вона має точність нижче в порівнянні з іншими моделями. Високу точність має модель R-CNN і її модифікації, але вона має дуже низьку продуктивність. Модель SSD має оптимальне співвідношення точності і продуктивності, тому ця модель взята за основу для розробки програми ідентифікації пішоходів на відеозображенні



Порівняння швидкодії різних моделей

АРХІТЕКТУРА ТА МАТЕМАТИЧНА МОДЕЛЬ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ



Функціонування нейрона згорткової площини описується виразом [10]:

$$y_k^{(i,j)} = f \left(\sum_{s=1}^K \sum_{t=1}^K (w_{k,s,t} x^{((i-1)+s, (j-1)+t)}) + b_k \right),$$

де $y_k^{(i,j)}$ - вихід нейрона з координатами (i, j) k -ої площині згорткового шару, b_k - нейрони зміщення k -ої площини, K - розмір рецептивного поля нейрона, $w_{k,s,t}$ - синаптична вага k -ої площини, для зв'язку з s -им, t -им значенням рецептивного поля, x - виходи нейронів попереднього шару.

Скільки карт містить згортковий шар C1, стільки буде здійснено згортку вхідного зображення. Розмір згорткової площини визначається наступними виразами:

$$w_c = w_{in} - K + 1,$$

$$h_c = h_{in} - K + 1,$$

де w_c і h_c - відповідно ширина і висота згорткової площини, w_{in} і h_{in} - відповідно ширина і висота площини, що сканується, K - розмір рецептивного поля.

Згортка можлива не тільки з одиничним кроком, при цьому:

$$w_c = (w_{in} - K) / step + 1,$$

$$h_c = (h_{in} - K) / step + 1,$$

де $step$ - крок згортки.

Згортка з кроком, відмінним від одиничного, можлива тільки при певних співвідношеннях параметрів w_{in} , h_{in} , K і $step$. Крок не повинен перевищувати розмір рецептивного поля K і w_c і h_c повинні бути цілими числами. Така згортка може зменшити загальну кількість обчислень, зберігши при цьому прийнятний рівень сканування зображення.

Функціонування нейрона площини підвибірки описується виразом [15]:

$$y_k^{(i,j)} = f \left(\frac{1}{K^2} w_k \sum_{s=1}^K \sum_{t=1}^K (x^{((i-1)+K+s, (j-1)+K+t)}) + b_k \right),$$

Навчання згорткових нейронних мереж

Основне завдання в навчанні нейронних мереж - мінімізація заданої функції помилки [14]. Зазвичай, оптимізується апостеріорна імовірність.

$$p(\theta|D) = p(\theta) \prod_{d \in D} p(d|\theta),$$

де функція правдоподібності $p(\theta)$ - помилка на навчальній вибірці, апіорне розподіл $p(d|\theta)$ — регуляризація.

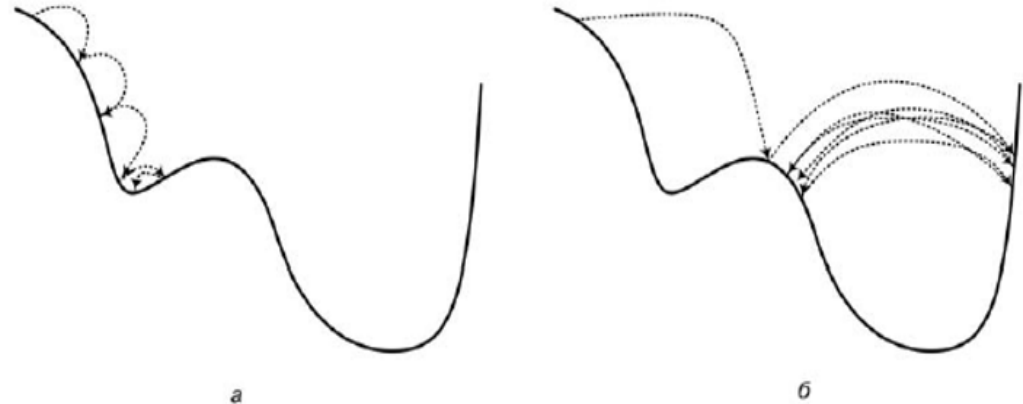
Завдання оптимізації - по заданій функції знайти аргументи, в яких ця функція мінімізується. Функція помилки в нейронних мережах має багато локальних екстремумів. Для того, щоб знайти оптимальний екстремум, використовується евристичний метод оптимізації - градієнтний спуск. Представивши поверхню функції помилки, завдання оптимізації зводиться до обчислення градієнта. Якщо функцію, що визначає поверхню, позначити через

$$E(\boldsymbol{\theta}) = E(\theta_1, \theta_2, \dots, \theta_n),$$

де $(\theta_1, \theta_2, \dots, \theta_n)$ — параметри функції, то її градієнт ∇E — це вектор похідних функції декількох змінних по кожній з компонент

Для вирішення цих проблем використовується **стохастичний градієнтний спуск** - підрахунок помилки і оновлення ваг відбувається не після проходження всієї тренувальної множини, а після кожного прикладу:

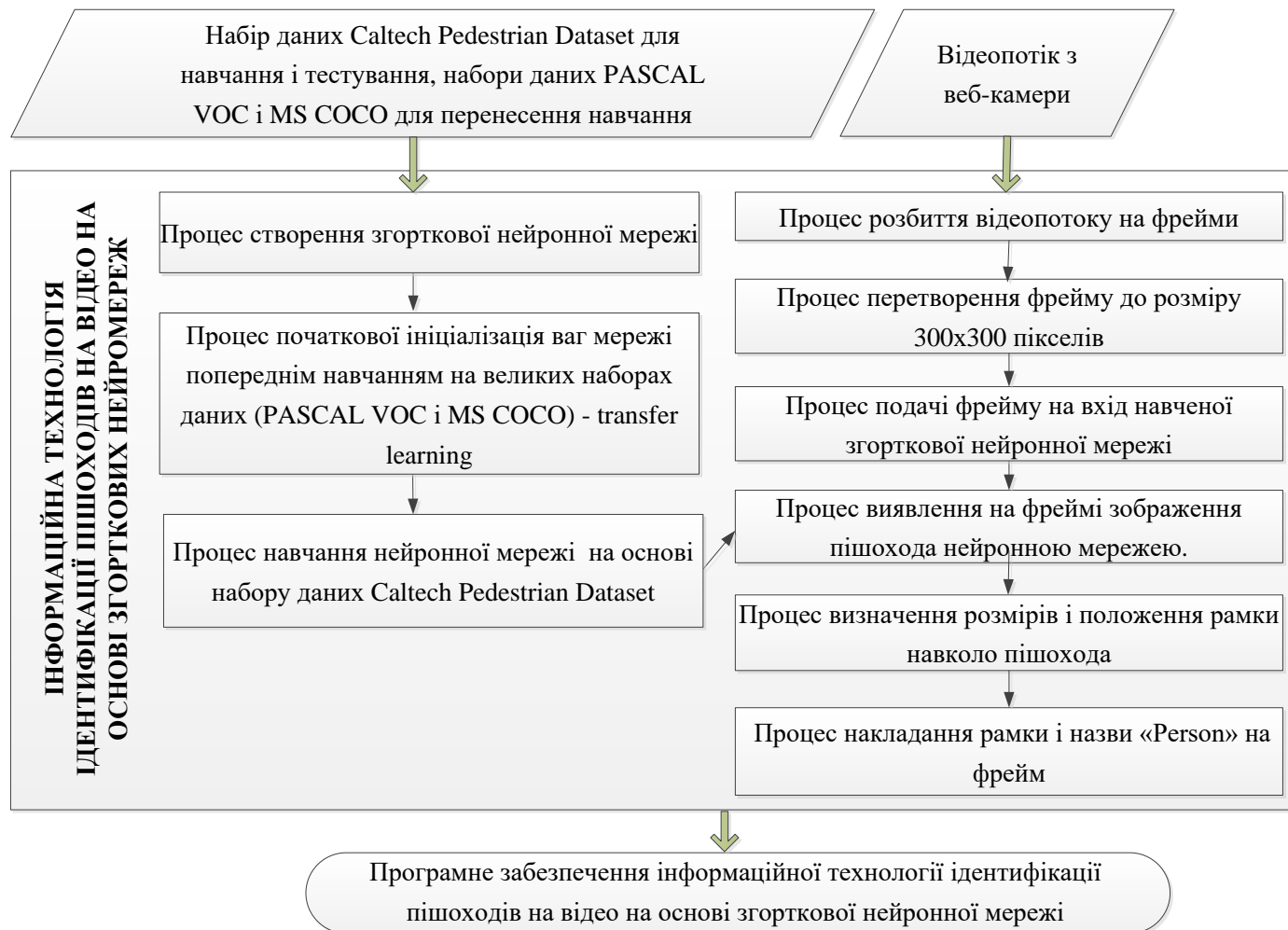
$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} - \eta \nabla E(f(\mathbf{x}_t, \boldsymbol{\theta}_{t-1}), y_t)$$



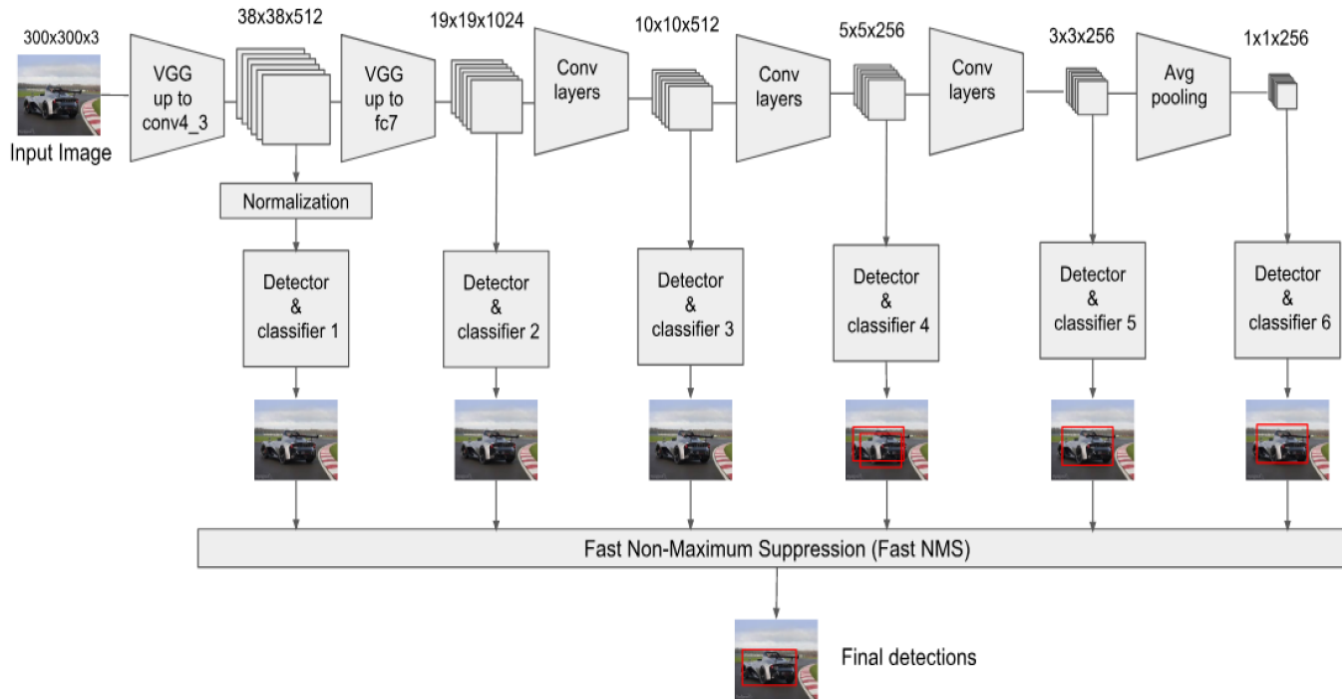
Проблеми зі швидкістю градієнтного спуску:

- а) — занадто маленькі кроки,
- б) — занадто великі кроки.

Структура інформаційної технології ідентифікації пішоходів у відеопотоці на основі згорткової нейронної мережі



Архітектура згорткової НМ SSD (Single Shot Detector)



Для оптимізації моделі використані:

- збільшення (аугментація) тренувальних даних;
- зміна параметрів вихідної мережі, таких, як розмір вікна передбачення об'єктів;
- додаткова регуляризація ваг моделі (регуляризація – це скорочення ваг (weight decay), так як вона призводить до зменшення їх абсолютних значень)

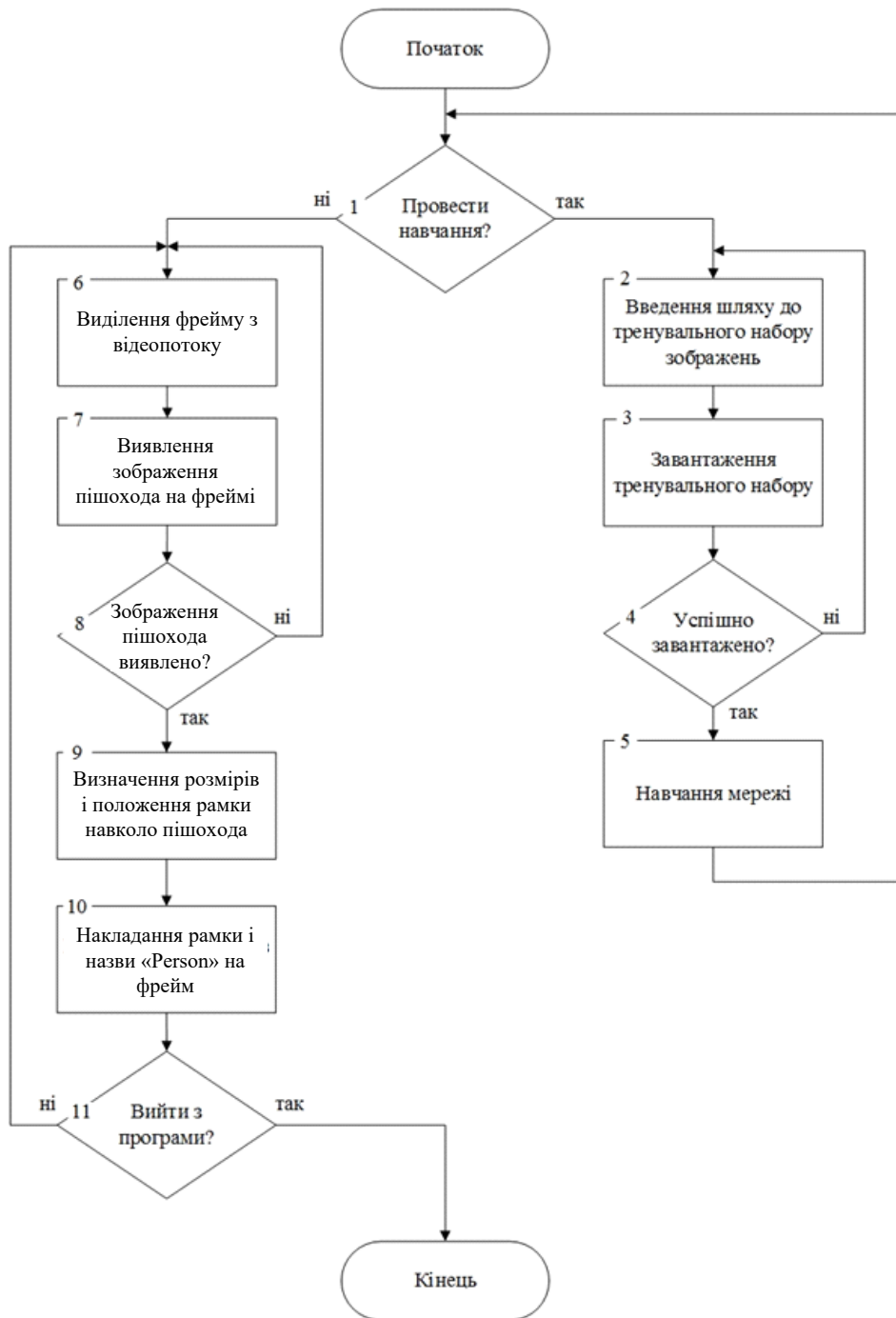
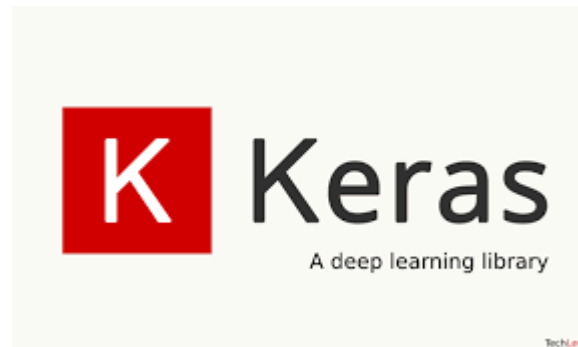
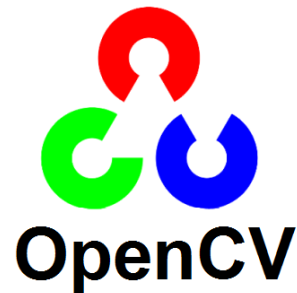


Схема алгоритму роботи головного компонента програми ідентифікації пішоходів у відеопотоці

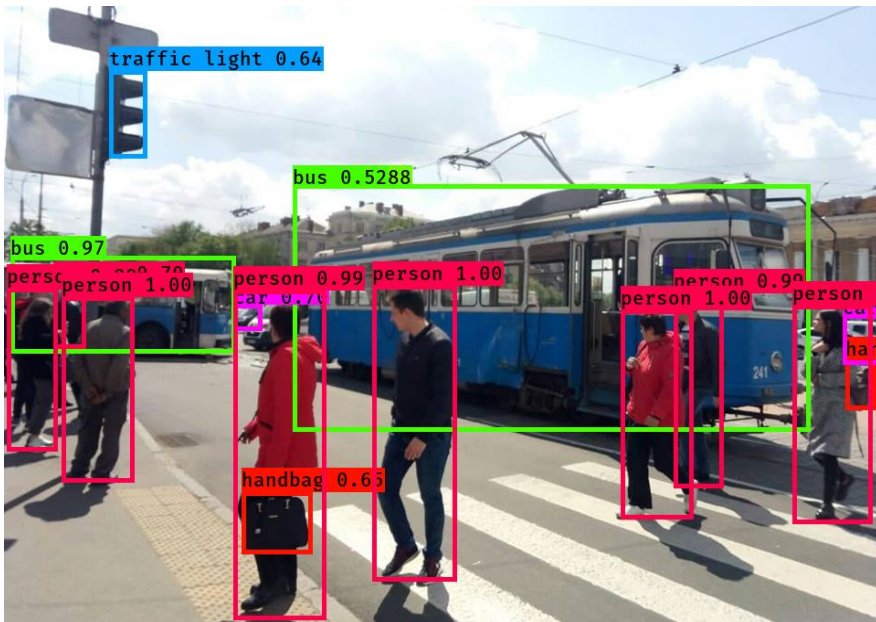
Обґрунтування вибору мови та засобів програмування

Було обґрунтовано вибір мови програмування Python, середовища програмування Visual Studio Code, бібліотеки комп'ютерного зору OpenCV, бібліотеки глибокого навчання Keras та бібліотеки векторно-матричних обчислень TensorFlow

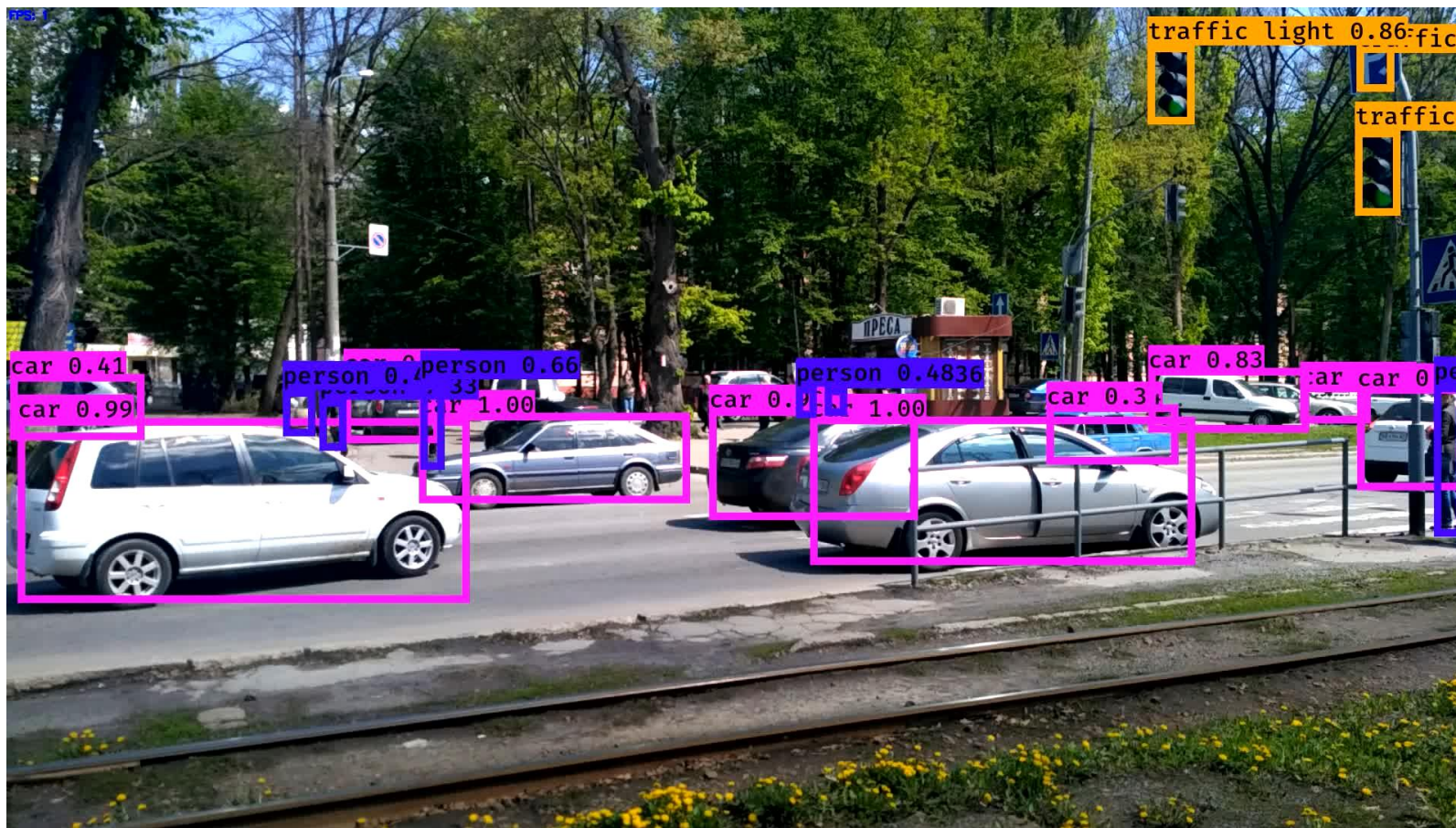




Результати роботи програми ідентифікації пішоходів у відеопотоці на основі згорткової нейронної мережі



Демонстрація роботи ідентифікації пішоходів у відеопотоці на основі згорткової нейронної мережі



Тестування та аналіз результатів роботи програми ідентифікації пішоходів у відеопотоці на основі згорткової нейронної мережі

Порівняння результатів роботи розробленої програми та програми-аналога

Програмний засіб	Кількість відеозображень у тестовій вибірці	Кількість пішоходів на відеозображеннях	Кількість правильно ідентифікованих пішоходів	Достовірність ідентифікації пішоходів, %
Програма RetinaNet	20	130	88	67,7
Розроблена програма	20	130	109	83,8

Із табл. видно, що розроблена програма має вищу достовірність ідентифікації пішоходів (83,8%), ніж аналогічна програма (67,7%), а значить достовірність ідентифікації пішоходів покращена на 16,1%, тобто мета роботи досягнута.

ЕКОНОМІЧНА ЧАСТИНА

Було виконано оцінювання комерційного потенціалу розробки інформаційної технології ідентифікації пішоходів на відео на основі згорткових нейромереж. Було визначено, що рівень потенціалу комерційного розвитку вищий за середній, а значить розроблена інформаційна система випереджає аналогічні інформаційні системи за показниками та має широкі перспективи. Товар має кращі функціональні показники, тому є досить конкурентоспроможним на ринку. Загальна вартість розробки становить 96669,93 грн. Абсолютна ефективність вкладених інвестицій досягла 878 807,88 грн., що свідчить про отримання вигоди від залучення коштів у розробку. Річна ефективність інвестицій у наукову розробку становить 116%, що вище ніж найнижча бар'єрна ставка дисконтування в 25%. Термін окупності інвестицій у проект становить 0,86 років, що ще більше демонструє доцільність фінансування проекту.. .

АПРОБАЦІЯ РЕЗУЛЬТАТІВ РОБОТИ ТА ПУБЛІКАЦІЇ

Апробація результатів роботи.

Результати досліджень апробовані на XLIX Науково-технічній конференції підрозділів Вінницького національного технічного університету (2020)», Вінниця, 2020 [1]

Публікації.

За результатами досліджень опубліковано тези доповіді на науково-технічній конференції [1]

ВИСНОВОК

В результаті виконання МКР розроблено інформаційну технологію та програмне забезпечення для ідентифікації пішоходів у відеопотоці на основі згорткової нейронної мережі. Програмне забезпечення створено мовою програмування Python з використанням бібліотек OpenCV, Keras та TensorFlow. Програма має вищу достовірність ідентифікації пішоходів, ніж аналогічна програма на 16,1%. Таким чином, мета роботи досягнута – достовірність ідентифікації пішоходів у відеопотоці підвищена.

ДЯКУЮ ЗА УВАГУ!