

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя
(повне найменування вищого навчального закладу)
Факультет комп'ютерно-інформаційних систем і програмної інженерії
(назва факультету)
Кафедра комп'ютерних систем та мереж
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістра

(освітній ступінь)

на тему: **Методи і засоби розпізнавання контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження**

Виконав: студент (ка) 6 курсу, групи СІМ-61
спеціальності 123 «Комп'ютерна інженерія»
(шифр і назва спеціальності)

	(підпис)	Щербаков О.О. (прізвище та ініціали)
Керівник	(підпис)	Яцишин В.В. (прізвище та ініціали)
Нормоконтроль	(підпис)	Тиш Є.В. (прізвище та ініціали)
Завідувач кафедри	(підпис)	Осухівська Г.М. (прізвище та ініціали)
Рецензент	(підпис)	Бойко І.В. (прізвище та ініціали)

Тернопіль
2021

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

(повне найменування вищого навчального закладу)

Факультет комп'ютерно-інформаційних систем і програмної інженерії

Кафедра комп'ютерних систем та мереж

ЗАТВЕРДЖУЮ

Завідувач кафедри Осухівська Г.М.

«_____» _____ 2021 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня магістр
(назва освітнього ступеня)

за спеціальністю 123 «Комп'ютерна інженерія»
(шифр і назва спеціальності)

студенту Щербакову Олегу Олексійовичу
(прізвище, ім'я, по-батькові)

1. Тема проекту (роботи) Методи і засоби розпізнавання контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження

Керівник проекту (роботи) Яцишин Василь Володимирович, к.т.н., доцент
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «28» жовтня 2021 року № 4/7-916

2. Термін подання студентом завершеної роботи _____

3. Вихідні дані до роботи Типи систем відеоспостереження, методи комп'ютерного зору, мова програмування Python, архітектури глибоких нейронних мереж

4. Зміст роботи (перелік питань, які потрібно розробити)

Вступ. 1. Аналіз підходів і принципів організації інтелектуалізованих систем відеоспостереження
2. Проектування архітектури системи відеоспостереження з інтеграцією інтелектуальної
складової виявлення контурів рухомих об'єктів 3. Програмування інтелектуальної складової для
системи відеоспостереження 4. Охорона праці та безпека в надзвичайних ситуаціях. Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

1. Актуальність і мета дослідження. 2. Задачі дослідження, об'єкт і предмет, наукова новизна і практична цінність дослідження. 3. Застосування технологій виявлення контурів об'єктів у різних сферах діяльності. 4. Варіанти організації систем відеоспостереження.

5. Сучасні архітектури нейронних мереж при розпізнаванні об'єктів та визначенні їх контурів

6. Архітектура нейронної мережі для визначення контурів рухомих об'єктів на базі Mask R-CNN.

7. Результати виявлення контурів рухомих об'єктів 8. Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
<i>Охорона праці та безпека в надзвичайних ситуаціях</i>	<i>Осухівська Г.М.</i>		
	<i>Стадник І.Я.</i>		

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	<i>Аналіз підходів і принципів організації інтелектуалізованих систем відеоспостереження</i>	<i>28.10.2021-11.11.2021</i>	<i>виконано</i>
2.	<i>Проектування архітектури системи відеоспостереження з інтеграцією інтелектуальної складової виявлення контурів рухомих об'єктів</i>	<i>15.11.2021 – 25.11.2021</i>	<i>виконано</i>
3.	<i>Програмування інтелектуальної складової для системи відеоспостереження</i>	<i>26.11.2021 – 05.12.2021</i>	<i>виконано</i>
4.	<i>Охорона праці та безпека в надзвичайних ситуаціях</i>	<i>05.12.2021 – 10.12.2021</i>	<i>виконано</i>
5.	<i>Оформлення пояснювальної записки</i>	<i>10.12.2021-12.12.2021</i>	<i>виконано</i>
6.	<i>Оформлення графічного матеріалу</i>	<i>12.12.2021-14.12.2021</i>	<i>виконано</i>
7.	<i>Попередній захист кваліфікаційної роботи магістра</i>	<i>15.12.2021</i>	<i>виконано</i>
8.	<i>Захист кваліфікаційної роботи магістра</i>		

Студент

(підпис)*Щербаков О.О.*_____
(прізвище та ініціали)

Керівник проекту (роботи)

(підпис)*Яцишин В.В.*_____
(прізвище та ініціали)

АНОТАЦІЯ

Тема кваліфікаційної роботи: “ Методи і засоби розпізнавання контурів рухомих об’єктів в інтелектуалізованих системах відеоспостереження ” // Кваліфікаційна робота // Щербаков Олег Олексійович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп’ютерно-інформаційних систем та програмної інженерії, група СІм-61 // Тернопіль, 2021 // с. – 86, рис. – 32, табл. – 8, аркушів А1 – 8, додат. – 1, бібліогр. – 31.

Ключові слова: метод, засіб, розпізнавання, об’єкт, інтелект, контур.

Метою роботи є дослідження методів і засобів розпізнавання контурів рухомих об’єктів в інтелектуалізованих системах відеоспостереження.

У кваліфікаційній роботі запропоновано модель архітектурного рішення щодо реалізації інтелектуалізованої системи відеоспостереження з функцією виявлення контурів рухомих об’єктів, яка використовує апаратний модуль ESP32-CAM та відеокамеру OV2640 і дає можливість з високою точністю на основі підходу сегментації екземплярів накладати та візуалізувати маски на кадрах відеопотоку.

Запропоновано оптимізацію структури нейронної мережі R-CNN на основі побудови піраміди функцій у мережі з одномасштабного входу, що дає змогу скоротити витрати часу на навчання мережі та є окремою гілкою при виявленні контурів рухомих об’єктів у системах відеоспостереження.

ABSTRACT

The theme of the thesis: " Methods and tools for the moving objects' contours recognition in intelligent video surveillance systems " /Master thesis / Shcherbakov Oleh Oleksiyovych/ Ternopil Ivan Pul'uj National Technical University, Faculty of Computer Information Systems and software engineering, group CIm -61 // Ternopil, 2021// p. - 86, fig. – 32, table. – 8, Sheets A1 – 8, Add – 1, Ref. – 31.

Keywords: method, tool, recognition, object, intelligence, mask.

The aim of the work is to study the methods and means of recognizing the contours of moving objects in intelligent video surveillance systems.

The thesis proposes a model of architectural solution for the implementation of an intelligent video surveillance system with the function of contouring moving objects, which uses the hardware module ESP32-CAM and OV2640 camcorder and allows with high accuracy based on the segmentation approach to overlay and visualize video masks on the frame.

It is proposed to optimize the structure of the R-CNN neural network based on the construction of a pyramid of functions in the network from a single-scale input, which reduces the time spent on network training and is a separate branch in detecting contours of moving objects in video surveillance systems.

ЗМІСТ

ПЕРЕЛІК ОСНОВНИХ УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ І СКОРОЧЕНЬ .	8
ВСТУП	9
РОЗДІЛ 1 АНАЛІЗ ПІДХОДІВ І ПРИНЦИПІВ ОРГАНІЗАЦІЇ ІНТЕЛЕКТУАЛІЗОВАНИХ СИСТЕМ ВІДЕОСПОСТЕРЕЖЕННЯ	13
1.1. Аналіз особливостей функціонування і сфер застосування систем відеоспостереження	13
1.2. Аналіз сучасних прикладних систем розпізнавання образів.....	18
1.3. Аналіз підходів до виявлення контурів об'єктів на зображеннях	20
1.4. Висновки до розділу	24
РОЗДІЛ 2 ПРОЕКТУВАННЯ АРХІТЕКТУРИ СИСТЕМИ ВІДЕОСПОСТЕРЕЖЕННЯ З ІНТЕГРАЦІЄЮ ІНТЕЛЕКТУАЛЬНОЇ СКЛАДОВОЇ ВІЯВЛЕННЯ КОНТУРІВ РУХОМИХ ОБ'ЄКТІВ.....	26
2.1. Проектування архітектури інтелектуалізованої системи відеоспостереження	26
2.2. Методи виявлення контурів рухомих об'єктів та їх сегментації.....	31
2.2.1. Порогова сегментація.	31
2.2.2. Сегментація на основі границь.....	33
2.2.3. Сегментація на основі областей зображення	34
2.2.4. Сегментація вододілу	35
2.2.5. Алгоритми сегментації на основі кластеризації	35
2.2.6. Нейронні мережі при сегментації.....	36
2.3. Оптимізація архітектури нейронної мережі Mask R-CNN для виявлення контурів рухомих об'єктів.....	37
2.4. Архітектура оптимізованої нейронної мережі	41
2.5. Застосування принципу Transfer Learning при визначенні контурів рухомих об'єктів	43
2.6. Висновки до розділу	45

РОЗДІЛ 3 ПРОГРАМУВАННЯ ІНТЕЛЕКТУАЛЬНОЇ СКЛАДОВОЇ ДЛЯ СИСТЕМИ ВІДЕОСПОСТЕРЕЖЕННЯ.....	46
3.1. Обґрунтування фреймворку для реалізації моделі виявлення контурів рухомих об'єктів	46
3.2. Алгоритм реалізації моделі для виявлення контурів об'єктів на статичних зображеннях.....	52
3.3. Програмна реалізація алгоритму виявлення контурів об'єктів та сегментації об'єктів	54
3.4. Процедура виявлення контурів рухомих об'єктів.....	61
3.5. Висновки до розділу	66
РОЗДІЛ 4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ	68
4.1. Охорона праці.....	68
4.2. Фактори виробничого середовища і їх вплив на життєдіяльність промислово-виробничого персоналу	72
ВИСНОВКИ.....	76
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	78
ДОДАТОК А ТЕЗИ КОНФЕРЕНЦІЙ	81

ПЕРЕЛІК ОСНОВНИХ УМОВНИХ ПОЗНАЧЕНЬ,
СИМВОЛІВ І СКОРОЧЕНЬ

БД	База Даних
ДН	Дистанційне навчання
КС	Комп'ютерна система
ПЗ	Програмне забезпечення
ЗНМ	Згорткова нерйонна мережа
РНМ	Рекурентна нейронна мережа
CNN	Convolutional Neural Network
GPU	Graphics Proccesing Unit
RGB	Red Green Blue
RNN	Recurrent Neural Network
UXGA	Ultra Extended Graphics Array

ВСТУП

Актуальність теми. Сучасна діджиталізація життя суспільства та різних сфер діяльності людини позитивно впливає на ефективність виробничих процесів та якість надання послуг. Значний внесок у технологічний розвиток різних галузей виробництва та сфери послуг забезпечують наукові здобутки та автоматизовані рішення із застосуванням інформаційно-комунікаційних технологій. Особливої уваги заслуговує галузь штучного інтелекту, що характеризується потужним проривом у створенні нових інтелектуальних методів і засобів, орієнтованих на зменшення впливу людського фактору при прийнятті рішень та виконанні бізнес-процесів. Це зумовлено накопиченням величезної кількості даних, розвитком технічної інфраструктури для їх зберігання та опрацювання із застосуванням хмарних технологій, наявністю технологічних інструментів аналізу великих даних.

Однією із сфер впровадження елементів штучного інтелекту, зокрема області комп'ютерного зору, є створення ефективних інтелектуалізованих систем відеоспостереження. Такі системи дозволяють підвищити ефективність та функціональність комплексів охоронної діяльності, забезпечити автоматизацію процесів у сфері громадського порядку, підвищити оптимальність управління транспортними потоками, надати можливість розпізнавання об'єктів у середовищах, які є агресивними або небезпечними для життя людей та ін.

Архітектурно, традиційні системи відеоспостереження, представляють собою сукупність програмно-технічних засобів об'єднаних в єдине інформаційне середовище з наявністю пристроїв відображення і трансляції відеопотоку, нагляд за яким виконує людина.

Проте, поява IoT та розвиток технологій комп'ютерного зору, дали можливість змінити та підвищити ефективність таких систем шляхом імплементації елементів комп'ютерного зору для розпізнавання статичних та рухомих об'єктів. Окрім цього, впровадження сервісу сповіщень щодо виявлення підозрілих чи нетипових об'єктів і поведінки підвищують реакцію спостерігача в режимі реального часу.

Важливим при проектуванні та імплементації інтелектуалізованих систем відеоспостереження є застосування методів і засобів виявлення контурів об'єктів, особливо рухомих, що потребує додаткового дослідження методів і засобів їх сегментації. На сьогодні інтенсивно досліджуються науковцями (М. Рахметова, Л. Рабинера, R. Szeliski, C. Steger, M. Ulrich, C. Wiedemann) та експериментально використовуються технології сегментації візуальних об'єктів з використанням різних підходів. Не зважаючи на значні досягнення у цій галузі, все ж потребують подальшого дослідження та розвитку методи і засоби формування контурів об'єктів. Актуальність такої задачі зумовлена необхідністю забезпечення балансу між використанням апаратних ресурсів та точності розпізнавання і формування контурів об'єктів у системах інтелектуалізованого відеоспостереження.

Мета і задачі дослідження. Метою роботи є дослідження методів і засобів розпізнавання контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження.

Для досягнення мети роботи поставлено і розв'язано наступні **задачі**:

- аналіз наукових публікацій та прикладних застосувань при виявленні контурів об'єктів, одержаних з відеопотоку;
- дослідження технологій проектування і впровадження систем відеоспостереження;
- оптимізація методів і засобів сегментації рухомих об'єктів у відеопотоці;
- розробка алгоритму сегментації рухомих об'єктів при аналізі відеоконтенту;
- програмна реалізація оптимізованого методу виявлення контурів об'єктів у системах відеоспостереження.

Об'єкт дослідження: процес виявлення контурів рухомих об'єктів методами штучного інтелекту.

Предмет дослідження: моделі, методи і засоби сегментації зображень.

Методи дослідження: При розв'язанні задач дипломного проектування використано наступні методи: аналіз та обґрунтування – при аналізі наукових джерел і публікацій щодо проектування систем відеоспостереження та сегментації

зображень; методи машинного навчання і теорія моделювання – при оптимізації методів виявлення контурів рухомих об’єктів; програмування – при написанні програмного коду формування контурів об’єктів на зображеннях; апробація – при визначенні ефективності методу сегментації зображень.

Наукова новизна отриманих результатів. Наукова новизна результатів дослідження полягає в наступному:

– уперше запропоновано модель архітектурного рішення щодо реалізації інтелектуалізованої системи відеоспостереження з функцією виявлення контурів рухомих об’єктів, яка використовує апаратний модуль ESP32-CAM та відеокамеру OV2640 і дає можливість з високою точністю на основі підходу сегментації екземплярів накладати та візуалізувати маски на кадрах відеопотоку.

– набули подальшого розвитку рішення щодо оптимізації структури нейронної мережі R-CNN на основі побудови піраміди функцій у мережі з одномасштабного входу, що дає змогу скоротити витрати часу на навчання мережі та є окремою гілкою при виявленні контурів рухомих об’єктів у системах відеоспостереження.

Практичне значення одержаних результатів. Практичне значення одержаних результатів полягає у програмній реалізації методу сегментації об’єктів на зображеннях, одержаних у системах відеоспостереження.

Публікації. Результати кваліфікаційної роботи апробовані на X міжнародній науково - технічній конференції молодих учених і студентів «Актуальні задачі сучасних технологій» (24-25 листопада 2021 р.) Тернопільського національного технічного університету імені Івана Пулюя та на IX науково-технічній конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2021 року) як тези конференцій.

1. Яцишин В.В., Щербаков О.О., Лова М.Р. Аналіз баз даних зображень у галузі комп’ютерного зору. Матеріали X міжнародної науково - технічної конференції молодих учених і студентів «Актуальні задачі сучасних технологій»

(24-25 листопада 2021 р.) Тернопільського національного технічного університету імені Івана Пулюя. Тернопіль: ТНТУ. 2021. С. 144.

2. Жаровський Р.О., Лова М.Р., Щербаков О.О. Застосування індексу структурної подібності зображень при їх аналізі. Матеріали ІХ науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2021 року). Тернопіль: ТНТУ. 2021. С. 114.

Структура роботи. Кваліфікаційна робота містить розрахунково-пояснювальну записку та графічний матеріал. До складу записки входить вступу, 4 розділи, загальні висновки, список використаних джерел і додатки. Обсяг роботи: розрахунково-пояснювальна записка – 86 арк. формату А4, графічна частина – 8 аркушів формату А1.

РОЗДІЛ 1

АНАЛІЗ ПІДХОДІВ І ПРИНЦИПІВ ОРГАНІЗАЦІЇ ІНТЕЛЕКТУАЛІЗОВАНИХ СИСТЕМ ВІДЕОСПОСТЕРЕЖЕННЯ

1.1. Аналіз особливостей функціонування і сфер застосування систем відеоспостереження

Однією із важливих сфер застосування систем відеоспостереження в комплексі із системами біометричної ідентифікації особи є область правоохоронної діяльності, що стосується моніторингу поведінки людей у громадських місцях і попередження злочинів.

Представники правоохоронних органів мають можливість застосовувати комп'ютерні системи для встановлення особи на основі аналізу як візуального образу людини, так і її характерних біометричних даних. Лівову частку щодо впровадження інтелектуальних систем, пов'язаних з імплементацією алгоритмів комп'ютерного зору, займають автоматизовані системи запобігання злочинним діям і боротьби з тероризмом.

До переваг таких систем належать висока ефективність і точність розпізнавання обличчя, відбитків пальця і сітківки ока, що допомагають запобігати різним видам шахрайства, а також довести причетність до злочинів затриманих осіб. На рис. 1.1 показано приклад практичного використання технологій комп'ютерного зору, щодо розпізнавання обличчя та відбитків пальців.

Ефективність застосування методів розпізнавання обличчя доведено на практиці також і у сфері митного контролю та перетину кордону. Один із сценаріїв їхнього використання передбачає порівняння зображень обличчя одержаного з системи відеоспостереження та цифрового фото із закордонного паспорту. Прикладом впровадженої системи розпізнавання образів щодо виявлення та аналізу обличчя людей є система PARAFE, що встановлена в одному з аеропортів Франції у 2017 році.

Така реалізація інтелектуальної системи відеоспостереження дозволяє пришвидшити потік людей при перетині кордону, замінивши систему біометричної ідентифікації за відбитком пальця.



Рис. 1.1. Результат розпізнавання об'єктів системою біометричного аналізу

Особливості біометричних характеристик обличчя може використовуватися при перевірці документів щодо керування транспортними засобами, а також при підтвердженні особи. Однак даний процес перебуває під суворим контролем відповідних органів у Європі.

Система відеоспостереження з імплементованим інтелектуальним модулем дала змогу представникам FBI виявити особу, що була причетна до організації терактів у Брюсселі, які відбулися у 2016 році, а процес назвали встановлення людини в капелюсі. У 2017 році терориста було виявлено у натовпі на стадіоні при проведенні фіналу Ліги Чемпіонів з футболу представниками поліції Уельсу.

Застосування камер на безпілотних літальних апаратах та програмного забезпечення розпізнавання обличчя дозволяє забезпечити моніторинг ситуації при проведенні різних масових заходів. На сьогодні вже реалізовано БПЛА, які можуть переносити об'єкти камери вагою до 10 кг і при цьому забезпечувати ідентифікацію особи з відстані до 800 м та висоти до 100 м.

До переваг застосування інтелектуалізованих систем відеоспостереження належить здатність підвищити точність і ефективність розпізнавання обличчя при забезпеченні заходів щодо безпеки у громадських місцях.

До конкретних прикладів використання інтелектуалізованих систем відеоспостереження сфери громадської безпеки належить:

- можливість знаходження дітей, які загубились та дорослих із втратою пам'яті.
- здатність встановлення фактів та виявлення осіб, які піддаються незаконній експлуатації;
- можливість встановлення та моніторингу дій осіб, що скоїли злочини;
- забезпечення формування доказової бази при проведенні розслідувань та їх пришвидженні.

Цінність інтелектуалізованих систем відеоспостереження та методів виявлення і сегментації об'єктів на зображеннях полягає в тому, що, наприклад, при наданні батьками зразка фото із зниклою дитиною, система можна знайти на відео та виконати оповіщення щодо наявності подібного зображення.

За допомогою технологій комп'ютерного зору, правоохоронні органи у своїй діяльності можуть підвищити якість пошуку на відеопослідовностях для встановлення локації, у якій перебуває шуканий об'єкт. Практичним прикладом ефективного застосування інтелектуалізованих систем відеонагляду є виявлення місцезнаходження близько трьох тисяч зниклих дітей в Індії протягом чотирьох днів.

Інтелектуальні складові систем аналізу відеоконтенту повинні володіти функцією оповіщення, що працює у режимі реального часу і формує так званий сигнал тривоги у випадку виявлення заданого об'єкту у відеопотоці. Аналогічним чином ці системи можна використовувати під час пошуку людей, які втратили пам'ять або є дезорієнтованими у просторі і часі (наприклад, за наявності таких хвороб як епілепсія, хвороба Альцгеймера, амнезія і т.п.)

Окрім наведених вище випадків щодо доцільності використання інтелектуальних систем відеоспостереження, вони можуть ефективно застосовуватися у прикордонних пунктах. Схема організації «розумної системи» при проходженні контролю на кордоні різних країн показано на рис. 1.2.

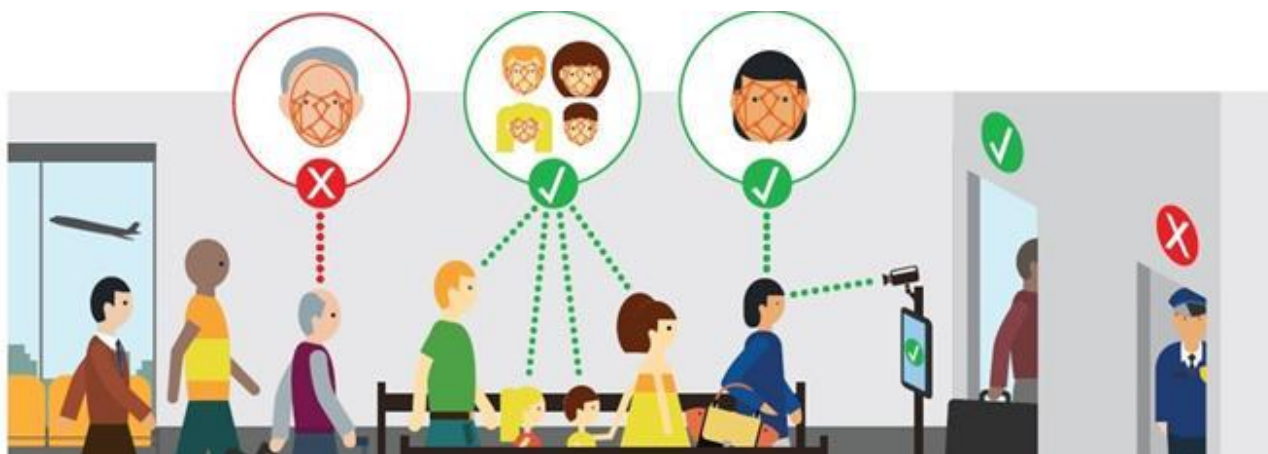


Рис. 1.2. Організація схеми перевірки осіб на прикордонних пунктах

Застосування методів глибокого машинного навчання та реалізація функцій розпізнавання об'єктів і їх контурів відіграють важливу роль у сфері охорони праці, зокрема при впровадженні технологій дистанційного консультування пацієнтів за допомогою засобів відеоконференцій (рис. 1.3.).



Рисунок 1.3 – Застосування технологій аналізу обличчя у медичній сфері

На основі технологій комп'ютерного зору в галузі охорони здоров'я вже сьогодні можна проводити наступні заходи щодо діагностики та спостереження за пацієнтом:

- більш точно аналізувати результати лікування і вживання ліків пацієнтами;
- діагностувати різні генетичні захворювання, наприклад синдром ДіГеорджа, із значенням показника точності на рівні близько 97% ;
- керувати і підтримувати процедури управління больовими відчуттями.

Характерною особливістю використання інтелектуалізованих систем опрацювання відео у банківській сфері є здатність ідентифікації особи за її обличчям. Даний факт підтверджується тим, що починаючи з 2019 року приблизно дві третини усіх банківських рахунків у США були відкритими через мережу Інтернет з використанням системи розпізнавання справжності обличчя, одержаного з відеопотоку. Світова пандемія COVID-19 пришвидшила розвиток процесів дистанційної взаємодії з клієнтами у зв'язку з карантинними обмеженнями.

Окрім цього, масова експлуатація мобільних девайсів створює передумови для компаній щодо зміщення акценту та впровадження зручних у використанні та багатофункціональних сервісів передачі відео з інтегрованою можливістю виявлення об'єктів, зокрема контурів обличчя. Для забезпечення достовірності виявлених об'єктів та уникнення шахрайства, пов'язаного із статичним зображенням, реалізовано алгоритми штучного інтелекту для встановлення їх справжності шляхом аналізу зміни емоцій, руху очей, повертання голови у різні сторони. Такий алгоритм перевірки та ідентифікації особи реалізовано у мобільному додатку Дія і дає можливість згенерувати унікальний цифровий підпис та використовувати ідентифікаційні дані особи.

Для забезпечення зручності і безпеки своїх клієнтів різні фінансові організації та установи інвестують великі кошти у створення та підтримку систем дистанційного прийому громадян з використанням цифрових технологій, мережі Інтернет та мобільних каналів зв'язку.

1.2. Аналіз сучасних прикладних систем розпізнавання образів

Серед найбільш високо технологічних компаній, зокрема Amazon, Google, Microsoft, Facebook та Apple великої популярності та актуальності набули проекти аналізу об'єктів на зображеннях, виявлення і представлення їхніх контурів і масок. Вони поділились своїми досягненнями і результатами з науковцями та фахівцями у галузі Data Science щодо використання накопичених і структурованих даних, алгоритмів та архітектур нейромереж, що дає змогу оптимізувати та впроваджувати різні сценарії при автоматизації бізнес-процесів конкретних підприємств та організацій.

Значним досягненням у сфері комп'ютерного зору є розроблений у 2014 році алгоритм «GaussianFace», що забезпечує точність розпізнавання об'єктів у відеопотоці на рівні 98,52%, в той час, як людина забезпечує точність розпізнавання образів на рівні 97,53%. Проте, практичне застосування цього алгоритму стримується необхідністю використання потужних програмно-апаратних ресурсів, зокрема, великого об'єму дискового простору і оперативної пам'яті, а також спеціалізованих процесорних ядер.

Великим досягненням компанії Facebook у 2014 став сервіс DeepFace. Він дає можливість виконувати порівняння зображень обличчя і визначати приналежність його власнику з точністю до 97,25%. Враховуючи те, що людина розпізнає обличчя з точністю до 97,53%, то такий алгоритм є дуже ефективним, а похибка, в порівнянні з людиною, становить лише 0,28%. [7].

На противагу Facebook, компанія Google реалізувала програмний сервіс FaceNet у 2015 році. Даний сервіс орієнтований на розпізнавання об'єктів живої природи, а навчання нейромережі виконано на загальнодоступному наборі даних «Labeled Face Wild – LFW». Точність, яка була досягнута при використанні FaceNet становить 99,63% [8]. На основі запропонованої архітектури штучної нейронної мережі та нового алгоритму опрацювання даних, компанії Google вдалося встановити відповідність проаналізованого обличчя і його власника практично з ідеальною точністю. Технологія FaceNet інтегрована у сервіс Google Photos і

застосовується при впорядкуванні зображень з можливістю автоматичного встановлення маркера з бази даних зображень людей.

Важливість та актуальність результатів, одержаних у сервісі FaceNet підтверджується появою його відкритої версії під назвою OpenFace.

При проведенні дослідження представниками Массачусетського університету, які використовували здобутки у сфері комп'ютерного зору компаній IBM та Microsoft, було виявлено ряд неточностей і дефектів при визначенні темношкірих жінок. При цьому значно менша кількість помилок була при ідентифікації світлих чоловіків. У 2018 році Microsoft офіційно повідомив про виправлення неточностей при класифікації осіб за зображенням їхнього обличчя.

Ще одним технологічним гігантом Amazon було розроблено та впроваджено службу розпізнавання обличчя Rekognition, яка викорисовує програмно-апаратні ресурси хмарних сервісів та орієнтована на застосування правоохоронними органами. За допомогою цього сервісу можна одночасно виконувати розпізнавання до 100 осіб, які знаходяться в межах однієї територіальної області та порівнювати їх з існуючими записами зображень, які містяться у базі даних. Кількість записів у базі даних становить понад десять мільйонів зображень облич.

На основі програмного комплексу LFIS, призначеного для розпізнавання зображень, запропоновано рішення Thales, яке продемонструвало точність ідентифікації людини на рівні 99,44% з продуктивністю менш ніж 5 с., а параметр реального виявлення продавця становив 98% і також був обчислений менше, чим за 5 с. Враховуючи, що середнє значення показника точності розпізнавання становив 66%, то ефективність запропонованого рішення значно підвищала його ефективність [9].

1.3. Аналіз підходів до виявлення контурів об'єктів на зображеннях

Сегментація зображень належить до галузі цифрового опрацювання зображень, що забезпечує декомпозицію зображення і виділення контурів об'єктів на різних його частинах у відповідності до визначених особливостей та властивостей. Основна ціль, яку переслідує сегментація зображень полягає у спрощенні представлення зображення при його аналізі. Під час проведення процедури сегментації виконується поділ зображення за різними областями, які мають подібні атрибути. Частина, на які поділяється зображення, називають об'єктами.

Сегментація є першим кроком в процесі аналізу зображення, без якого реалізація технік комп'ютерного зору була б практично неможливою. Застосування методів поділу графічних об'єктів, виявлених у відеопотоці, орієнтовано на групування специфічних пікселів зображення, призначення їм міток та класифікації інших пікселів у відповідності до цих маркерів.

У машинному навчанні можна використовувати мітки, згенеровані у результаті сегментації, для використання алгоритмів навчання з вчителем або без нього. Це дозволяє вирішувати багато проблем бізнесу. На прикладі краще зрозуміти, як працює сегментація зображень (рис. 1.3).



Рис. 1.3. Приклад зображення із стільцем

Як видно з рис. 1.3, стілець розміщений посеред дороги. Використовуючи сегментацію зображення, можна відокремити стілець від зображення. Крім того, існує можливість застосування численних методів сегментації зображень для одержання різних результатів. Наприклад, якщо необхідно використовувати зображення з кількома стільцями, то тоді доведеться використовувати семантичну сегментацію. З іншого боку, якщо потрібно визначити кожен стілець, наявний на зображенні (рис. 1.4), то тоді доведеться використовувати сегментацію на рівні екземплярів.



Рис. 1.4. Приклад зображення при сегментації екземплярів

Сегментація зображень є фундаментальним аспектом в області комп'ютерного зору і має багато застосувань у різних галузях. Сфери найбільш необхідного використання підходу сегментації зображень наведені нижче.

1. Розпізнавання обличчя.

Технологія розпізнавання облич, наявна у смартфонах iPhone і реалізує функції передових систем безпеки, що побудовані на використанні сегментації зображень для ідентифікації власника. Смартфон повинен мати можливість

ідентифікувати унікальні риси обличчя, щоб будь-яка інша особа, крім власника, не могла отримати доступ телефону або системи.

2. Ідентифікація номерних знаків

Багато світлофорів і камер використовують ідентифікацію номерних знаків для стягнення штрафів і допомоги при обшуку. Технологія ідентифікації номерних знаків дозволяє системі дорожнього руху розпізнавати автомобіль та отримувати інформацію про його власника. Така система використовує сегментацію зображення, щоб відокремити номерну табличку та її інформацію від решти об'єктів, присутніх у його полі зору. Ця технологія значно спростила процес накладення штрафів для урядових департаментів.

3. Пошук на основі зображень

Google та інші пошукові системи, які пропонують засоби пошуку на основі зображень, використовують методи сегментації зображень, щоб ідентифікувати об'єкти, присутні на зображенні-запиті, і порівняти їхні результати з відповідними зображеннями, які вони знаходять у результаті пошуку.

4. Медична візуалізація

У медичній галузі сегментація використовується для того, щоб знаходити та ідентифікувати ракові клітини, вимірювати об'єм тканин, запускати віртуальне моделювання хірургічних операцій та здійснювати навігацію в межах операції. Сегментація зображень має багато застосувань у медичній сфері. Це допомагає визначити уражені ділянки та спланувати лікування для них.

Окрім цих додатків, сегментація зображень використовується у виробництві, сільському господарстві, безпеці та багатьох інших секторах. Оскільки технології комп'ютерного зору стають все більш розвиненими, використання методів сегментації зображень буде відповідно зростати.

Наприклад, деякі виробники почали використовувати методи сегментації зображень, щоб знайти неякісні продукти. Тут алгоритм фіксує лише необхідні компоненти з зображення об'єкта та класифікує їх як несправні чи оптимальні. Ця система знижує ризик людських помилок і робить процес тестування більш ефективним для організації.

Сегментація зображень – це дуже широка галузь, що володіє потужним інструментом реалізації цього процесу. Підходи щодо сегментації зображень можна класифікувати за параметрами наступним чином:

1. Класифікація, заснована підходах до реалізації

У найпростішому сенсі сегментація зображення – це ідентифікація об'єкта. Алгоритм не може класифікувати різні компоненти, не ідентифікувавши спочатку об'єкт. Від простих до складних реалізацій усі сегменти зображень працюють на основі ідентифікації об'єкта. Отже, можна проводити класифікацію методів сегментації зображень на основі того, як алгоритми ідентифікують об'єкти, тобто яким чином групують подібні пікселі та відокремлюють їх від різнорідних. Існує два підходи до виконання цього завдання:

Підхід на основі виявлення подібності

Цей метод передбачає виявлення схожих пікселів на зображенні відповідно до вибраного порогу, об'єднання зон, розширення та зменшення області. Кластеризація та подібні алгоритми машинного навчання використовують цей метод для детектування невідомих функцій або ознак. Алгоритми класифікації дотримуються цього підходу для виявлення ознак і розділення сегментів зображення відповідно до них.

Підхід на основі границь (виявлення розривів)

Підхід на основі границь є протилежним до підходу на основі порогів при ідентифікації об'єктів. На відміну від алгоритму визначення подібності, коли шукають пікселі зі схожими характеристиками, підхід на основі границь виконує пошук пікселів, які відрізняються один від одного. Виявлення точок, виявлення краю, виявлення лінії та подібні алгоритми дотримуються цього методу, коли вони виявляють край різнорідних пікселів і відповідно відокремлюють їх від решти зображення.

2. Класифікація на основі техніки сегментації

Існує два глобальних підходи, що реалізують різні техніки сегментації і які включають в себе різні методи сегментації зображень. Їхнє застосування визначається у відповідності до типу зображення, яке необхідно опрацювати та

проаналізувати, і від результатів, які очікуються отримати. На основі цих критеріїв можна розділити алгоритми сегментації зображень на такі категорії:

Алгоритми на основі аналізу структури

Особливістю цих алгоритмів є наявність структурних елементів зображення. Сюди входять пікселі, розподіли, гістограми, щільність пікселів, розподіл кольору та інша відповідна інформація. Потім необхідно володіти інформацією про структурні дані щодо області зображення, яку потрібно відокремити від нього. На основі цих даних алгоритм може виявити об'єкт, а клас алгоритмів дотримується підходу аналізу області зображення.

Клас стохастичних технік

Алгоритми класу стохастичних технік вимагають інформації про дискретні значення пікселів зображення, а не про структуру необхідної ділянки зображення. Завдяки цьому вони не вимагають багато інформації для сегментації зображень і є корисними, коли доводиться працювати з кількома зображеннями. Алгоритми машинного навчання, такі як кластеризація K-середніх і алгоритми ANN, належать до цієї категорії.

Гібридні техніки.

Як можна здогадатися з назви, ці алгоритми використовують як стохастичні, так і структурні методи. Це означає, що вони використовують структурну інформацію необхідної області та інформацію про дискретні пікселі всього зображення для виконання сегментації зображення.

1.4. Висновки до розділу

У даному розділі одержано основні наукові і практичні результати:

1. Проведено аналіз особливостей функціонування і сфер застосування систем відеоспостереження, виявлено їхні основні переваги та недоліки, що дало можливість обґрунтувати необхідність впровадження інтелектуальних складових, що дозволить більш якісно проводити моніторинг у конкретній прикладній області

та при розв'язанні поставлених задач, шляхом зменшення впливу людського фактору.

2. Проаналізовано характеристики прикладних систем відеоспостереження з інтегрованими інтелектуальними модулями, які забезпечують виконання різних задач, зокрема ідентифікації особи у відеопотоці, і встановлено, що їхня реалізація потребує значних фінансових і програмно-апаратних засобів, а актуальним завданням є оптимізація існуючих моделей і методів розпізнавання контурів рухомих об'єктів (сегментація об'єктів).

3. Проведено аналіз підходів до ідентифікації об'єктів на зображеннях, у результаті якого встановлено, що найбільш ефективними є підходи на основі сегментації, які дають змогу визначати як їх семантично близькі області, так і забезпечувати сегментацію на рівні екземплярів.

РОЗДІЛ 2

ПРОЕКТУВАННЯ АРХІТЕКТУРИ СИСТЕМИ ВІДЕОСПОСТЕРЕЖЕННЯ З ІНТЕГРАЦІЄЮ ІНТЕЛЕКТУАЛЬНОЇ СКЛАДОВОЇ ВИЯВЛЕННЯ КОНТУРІВ РУХОМИХ ОБ'ЄКТІВ

2.1. Проектування архітектури інтелектуалізованої системи відеоспостереження

В загальному випадку, під системою відеоспостереження розуміють сукупність програмних та апаратних засобів, а також каналів комунікації, які в комплексі здатні одержувати візуальну інформацію про ситуацію і стан об'єкта у реальному часі. Такі системи дозволяють збирати та опрацьовувати відеопотік з подальшим його перетворенням, накопиченням і зберіганням, а при потребі й архівуванням. Це дає змогу проводити інтелектуальний аналіз зображень та ідентифікувати об'єкти у визначений час та у визначеному місці.

Традиційні системи відеонагляду, зазвичай, виконують функції спостереження за територією, де зберігаються цінні матеріальні речі та об'єкти, зокрема, це стосується автомобільних паркінгів, внутрішнього нагляду у гіпер- та супермаркетах, складських приміщеннях і т.д. Такі системи показують високу ефективність та продуктивність в процесі моніторингу переміщення рухомих об'єктів, для прикладу, на залізничних та автовокзалах, на промислових підприємствах.

Системи відеоспостереження допомагають також у прийнятті управлінських рішень, коли необхідний контроль за технологією виробництва певних товарів і послуг.

Для обґрунтування вибору системи відеонагляду, перш за все, необхідно визначити її структурні компоненти і зв'язки між ними, а також визначити сферу застосування та функціональність системи. Якісними і фінансовими критеріями вибору обладнання для реалізації системи відеоспостереження виступають:

- для задачі ідентифікації особи (за умови відомих про неї ознак) доцільно використовувати HD камери нижчого цінового діапазону, що разом із приладами опрацювання відеопотоку («Digital Video Recorder - DVR»), забезпечують можливість оператору системи розпізнавати людину у русі на основі аналізу прямих і додаткових її властивостей. До таких ознак можна віднести зріст особи, форма одягу, головний убір, тип ходи, інші габаритні розміри або особливі ознаки;
- для задачі розпізнавання рухомих об'єктів та людей, інформація про ознаки яких відсутні, варто використовувати відеокамери, роздільна здатність яких коливається в межах 1,3 МПк – 2,0 МПк або мережевих камер з підтримкою IP протоколу, обладнання для обробки відео у форматі HD або NVR відео-реєстраторів. Це зумовлено необхідністю забезпечення високих якісних показників відео (чіткість, контрастність) для розпізнавання ознак невідомих об'єктів;
- для задачі детального виявлення об'єктів – застосовуються IP-камери захоплення відео з роздільною здатністю вище 3 МПк в комплексі з системою автоматичної реєстрації або сервера інтелектуального розпізнавання образів. Такі системи дають змогу розрізняти, наприклад, номерні знаки автомобілів при їхньому русі, написи на окремих об'єктах та ін. Зовнішній вигляд IP-камери наведемо на рис. 2.1.



Рис. 2.1. Зовнішній вигляд IP-відеокамери

Основними факторами, які впливають на якість та ефективність систем відеоспостереження при їх проектуванні та впровадженні є чітко та однозначно сформульовані вимоги до функціонального призначення, які повинна забезпечувати система. Це допомагає фахівцям, а особливо інтеграторам відеосистем, узгодити потреби зацікавленої сторони і виконати якісний підбір технічного забезпечення з урахуванням вимог територіального розташування, а також фізичної та логічної топологій середовища обміну відеоконтентом.

Варто відмітити, що сучасні системи відеоспостереження, на відміну від телебачення (замкнута система), можуть взаємодіяти із сервісами, які не входять у структуру комплексу відеонагляду і є зовнішніми по відношенню до них (рис. 2.2).

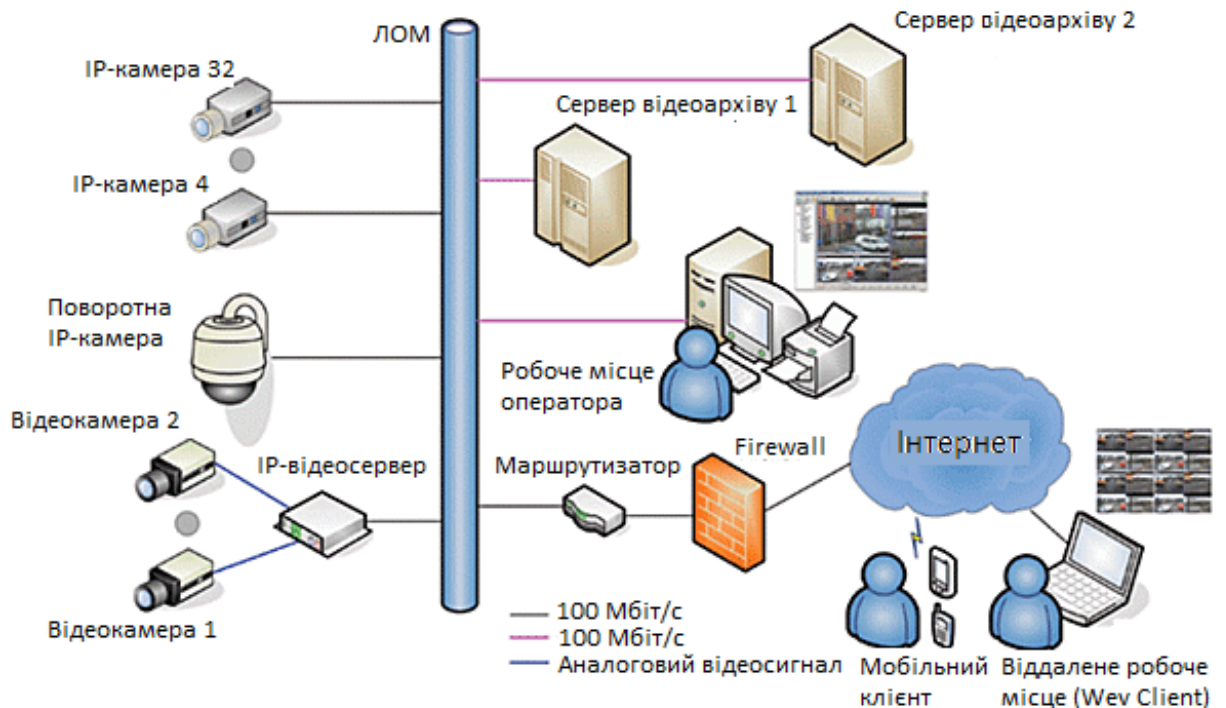


Рис. 2.2. Взаємодія систем відеоспостереження із зовнішніми сервісами

До основних функцій сучасних систем відеонагляду належать:

- здатність потокової трансляції через мережу Інтернет;
- здатність спостереження та розпізнавання об'єктів нагляду;
- можливість зберігання та накопичення відео даних;

- забезпечення можливості опрацювання відео за допомогою інтегрованих алгоритмів штучного інтелекту;
- надання одночасного і паралельного доступу до відеопотоку в режимі реального часу для багатьох користувачів;
- здатність забезпечувати функції авторизованого доступу до ресурсів системи відеоспостереження;
- надають можливість прийняття рішень при ізольованому автоматичному режимі.

Прикладом традиційної, класичної архітектури локальної системи відеоспостереження, що володіє інтерфейсом доступу до мережі Інтернет є структура, показано на рис. 2.3.

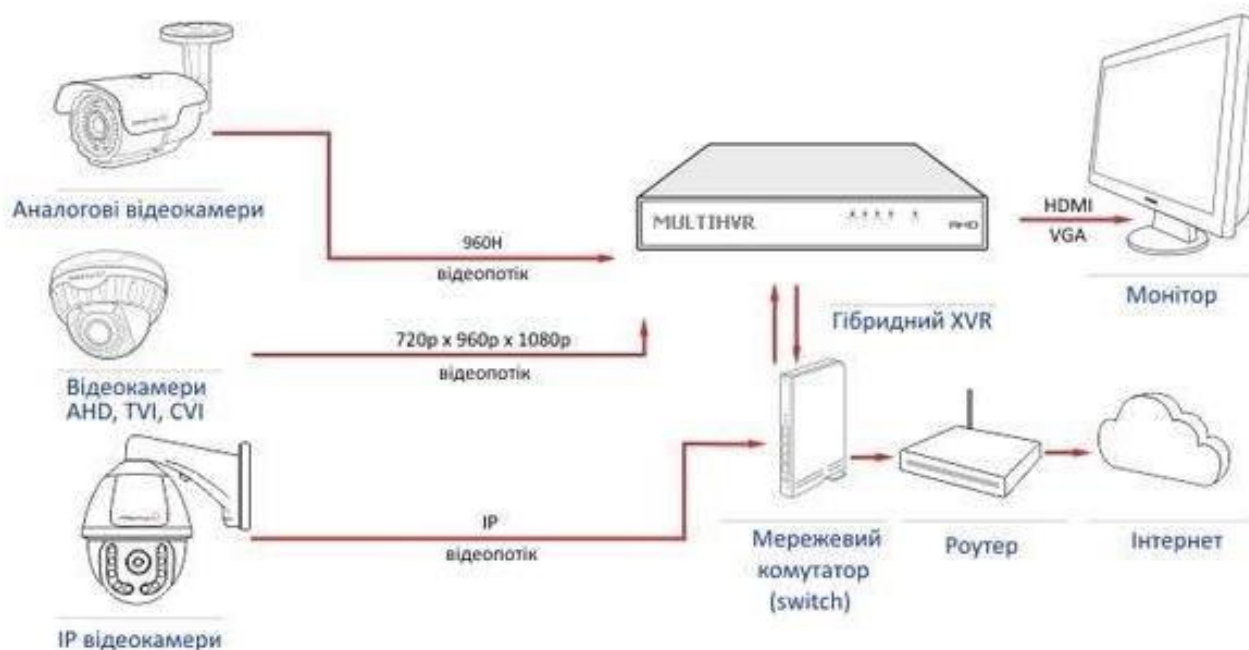


Рис. 2.3. Структура локальної системи відеоспостереження

При проектуванні сучасних систем відеонагляду потрібно враховувати особливості застосування та властивості їх структурних компонентів, зокрема, цифрових відеокамер, NVR та серверів, а також особливостей комплексу прикладного програмного забезпечення.

Щодо апаратних пристроїв відтворення відеоінформації, то в даному випадку можна застосовувати будь-яке обладнання з екраном – починаючи від смартфонів закінчуючи мультимедійними дошками.

Ефективним засобом для накопичення одержаної інформації та проведення її аналізу виступають хмарні сховища і сервіси. Однак варто відмітити поки що доволі високу вартість їхнього використання.

Модернізація встановлених раніше систем відеонагляду повинна супроводжуватися заміною аналогових камери на цифрові з підтримкою HD та FHD, а середовище передачі повинно бути реалізовано мінімум за допомогою коаксіального кабелю. Це забезпечує значне підвищення якості зображення при трансляції відео від 2 до 8 МП, що значно краще, ніж у випадку 0,3 Мп аналогових камер.

У даному дослідженні пропонується проект архітектури комп'ютерної системи відеонагляду, який реалізується із застосуванням модуля управління ESP32-CAM і 2 Мп відеокамери. На рис. 2.4 показано архітектуру на концептуальному рівні.

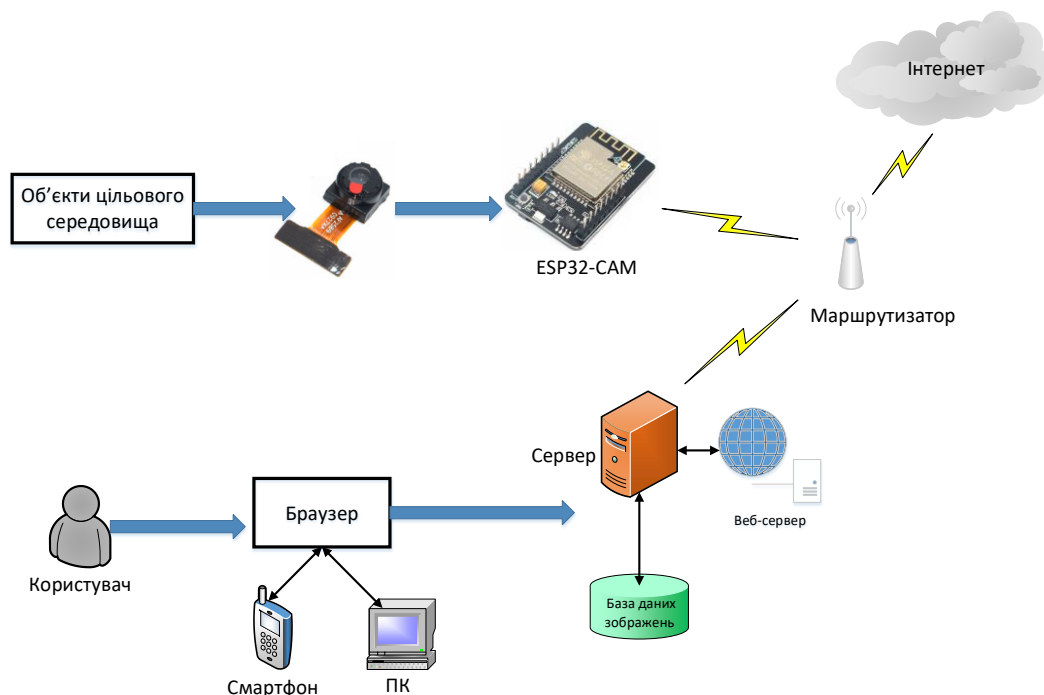


Рис. 2.4. Структура інтелектуалізованої системи відеоспостереження з виявленням контурів рухомих об'єктів

Основними структурними компонентами системи, показаної на рис. 2.4 є:

- відеокамера – використовується для захоплення зображення;
- модуль ESP32-CAM – забезпечує опрацювання і передачу інформації, а також має зв'язок з сервером та мережею Інтернет;
- сервер – призначений для зберігання відеоінформації на якому розгорнуто веб-сервер та базу даних зображень;
- пристрої відображення відеоінформації (смартфон, персональний комп'ютер) – забезпечують доступ до модуля через браузер, який має доступ до web-сервера.

Така архітектура дає змогу одержати доступ до відеоконтенту в режимі реального часу та забезпечити функцію розпізнавання контурів об'єктів. Наступний крок полягає в налаштуванні параметрів апаратного забезпечення та аналізу їх технічних характеристик.

2.2. Методи виявлення контурів рухомих об'єктів та їх сегментації

Проаналізувавши підходи до сегментації зображень потрібно більш детально зупинитися на методах сегментації, зокрема: порогова сегментація, сегментація на основі границь, сегментація на основі областей, сегментація «вододілу», алгоритми сегментації на основі кластеризації, нейронні мережі.

2.2.1. Порогова сегментація.

Найпростішим методом сегментації при опрацюванні зображень є пороговий метод. Він ділить пікселі в зображенні, порівнюючи інтенсивність пікселя з заданим значенням (порогом). Це корисно, коли необхідний об'єкт має більшу інтенсивність, ніж фон (непотрібні частини). В даному випадку можна вважати порогове значення (T) константою, але воно працюватиме лише в тому випадку, якщо зображення має дуже мало шумів (непотрібна інформація та дані).

Порогове значення може бути постійним або динамічним відповідно до вимог. Метод порогового значення перетворює зображення у відтінках сірого в

двійкове, розділяючи його на два сегменти (обов'язкові та не обов'язкові). Відповідно до різних порогових значень можна класифікувати порогову сегментацію на такі категорії:

Просте порогове значення.

У цьому методі виконується заміна пікселів зображення білими або чорними. Тепер, якщо інтенсивність пікселя в певній позиції менша за порогове значення, то виконується його заміна на піксель чорного кольору. З іншого боку, якщо він вищий за поріг, то його слід замінити на білий. Це просте встановлення порогів і особливо підходить для початківців у сегментації зображень.

Бінаризація.

При використанні простого порогового значення обирається постійне його значення, яке застосовується для сегментації зображення. Однак як визначити, що вибране значення було правильним? Хоча суть простого методу для цього полягає в тому, щоб перевірити різні значення та вибрати одне, він не найефективніший. Для приклад візьмемо зображення з гистограмою, яка має два піки: один для переднього плану, а другий для заднього. Використовуючи бінаризацію, можна взяти приблизне значення середини цих піків як порогове значення. При бінаризації виконується обчислення порогового значення з гистограми, якщо зображення є бімодальним. Цей процес досить популярний для сканування документів, розпізнавання образів та видалення непотрібних кольорів із файлу. Однак він має багато обмежень. Його не можна використовувати для зображень, які не є бімодальними (зображення, гистограми яких мають кілька піків (більше, ніж 2)).

Адаптивне порогове значення.

Наявність одного постійного порогового значення може бути неефективним підходом для різних зображень. Кожне зображення має різний фон та умови, які впливають на їхні властивості. Таким чином, замість використання одного постійного порогового значення для сегментації на всьому зображенні, можна визначити змінну, у якій будуть зберігатись багато порогових значень. При застосуванні такого алгоритму застосовуються різні порогові значення для різних

частин зображення. Цей метод добре працює з зображеннями з різними умовами освітлення. У такому випадку буде використовуватися алгоритм, який сегментує зображення на менші ділянки та обчислює порогове значення для кожного з них.

2.2.2. Сегментація на основі границь

Сегментація на основі границь є однією з найпопулярніших реалізацій сегментації при опрацюванні зображень. Він фокусується на визначенні країв різних об'єктів на зображенні. Це важливий крок, оскільки він допомагає знайти особливості різних об'єктів, присутніх на зображенні, оскільки краї містять багато корисної інформації. Метод виявлення границь набув широкої популярності, оскільки допомагає видалити небажану та непотрібну інформацію із зображення. Це значно зменшує його розмір та полегшує аналіз. Алгоритми, що використовуються в сегментації на основі аналізу границь, визначають краї зображення відповідно до відмінностей у текстурі, контрасті, рівні сірого, кольорі, насиченості та інших властивостях. Покращити якість результатів можна, з'єднавши всі краї в ланцюжки, які точніше відповідають межах зображення. Існує багато доступних методів сегментації на основі країв. Їх можна розділити на дві категорії:

Виявлення краю на основі пошуку

Методи виявлення країв на основі пошуку зосереджені на обчисленні міри «міцності» краю та пошуку локальних спрямованих максимумів величини градієнта за допомогою обчисленої оцінки локальної орієнтації краю.

Виявлення країв на основі нуля

Методи виявлення країв на основі перетину нуля шукають нульові перетини в похідному виразі, отриманому із зображення. Як правило, доведеться попередньо опрацювати зображення, щоб видалити небажаний шум і полегшити виявлення границь. Roberts Cross, Canny, Deriche Prewitt є одними з найпопулярніших операторів виявлення країв і полегшують виявлення розривів. При застосуванні методу виявлення об'єктів на основі границь основна мета – отримати мінімум часткової сегментації, щоб забезпечити групування усіх локальних ребер у

двійкове представлення. У створеному двійковому зображенні крайові ланцюжки повинні відповідати існуючим компонентам відповідного зображення.

2.2.3. Сегментація на основі областей зображення

Алгоритми сегментації на основі областей поділяють зображення на ділянки з подібними ознаками. Ці області є лише групою пікселів, і алгоритм знаходить ці групи, спочатку визначаючи початкову точку, яка може бути невеликою частиною або великою частиною вхідного зображення. Після знаходження початкових точок алгоритм сегментації на основі областей або додає до них більше пікселів, або зменшує їх, щоб об'єднати з іншими початковими точками. На основі цих двох методів можна класифікувати сегментацію на основі областей на такі категорії:

Збільшення області.

У цьому методі процес сегментації починається з невеликого набору пікселів, а потім відбувається ітеративне об'єднання інших у відповідності до конкретних умов подібності. Алгоритм збільшення області вибирає довільний початковий піксель на зображенні, порівнює його з сусідніми і починає збільшувати область, знаходячи збіги з початковою точкою. Коли конкретна область не може розвиватися далі, алгоритм обирає інший початковий піксель, який може не належати жодній з існуючих областей. Одна область може мати занадто багато атрибутів, через що вона займає більшу частину зображення. Щоб уникнути такої помилки, алгоритми нарощування областей розвивають кілька регіонів одночасно. Алгоритми збільшення області використовують для зображень, які мають багато шуму, оскільки він ускладнює пошук країв або використання порогових алгоритмів.

Поділ та злиття областей.

Як випливає з назви, сфокусований метод поділу області та об'єднання виконує дві дії разом – розділяючи та об'єднуючи частини зображення. Спочатку зображення розбивається на області, які мають подібні атрибути, і об'єднує сусідні частини, схожі одна на одну. При розділенні областей алгоритм розглядає все зображення, тоді як при збільшенні області він фокусується на конкретній точці.

Метод поділу та злиття областей дотримується методології «розділяй і владарюй». Він виконує розбиття зображення на різні частини, а потім узгоджує їх відповідно до заданих умов. Інша назва алгоритмів, які виконують це завдання, — алгоритми розділеного злиття.

2.2.4. Сегментація вододілу

При обробці зображень шляхом застосування методики «вододілу» передбачає перетворення зображення у формат з відтінками сірого. Він відноситься до «геологічного вододілу» або простого «вододілу». Алгоритм вододілу опрацьовує зображення так, ніби це була топографічна карта. Він розглядає яскравість пікселя як його висоту і знаходить лінії, які проходять уздовж вершини цих ребер. Вододіл має багато технічних визначень і має кілька застосувань. Окрім визначення ребер пікселів, він зосереджується на визначенні западин (протилежних від «хребтів») і заливає їх маркерами, поки вони не зустрінуться з лініями вододілу, що проходять через хребти. Оскільки западини мають багато маркерів, а гребені – ні, зображення розбивається на кілька областей відповідно до «висоти» кожного пікселя. Метод вододілу перетворює кожне зображення на топографічну карту і відображатиме топографію через сірі значення пікселів. Тепер ландшафт з долинами та хребтами, безумовно, мав би тривимірні аспекти. Вододіл розглядатиме тривимірне відображення зображення і відповідно створюватиме області, які називаються «басейнами водозбору». Він має багато застосувань у медичній галузі, наприклад, МРТ, медичне зображення тощо.

2.2.5. Алгоритми сегментації на основі кластеризації

Алгоритми сегментації на основі кластеризації належать до класу алгоритмів без міток цільових змінних і допомагають знайти приховані дані на зображенні, які можуть бути невидимими для звичайного зору. Ці приховані дані містять таку інформацію, як кластери, структури, затінення тощо. Як випливає з назви, алгоритм кластеризації ділить зображення на кластери (непересікаючі групи) пікселів, які мають схожі характеристики. Деякі з популярних алгоритмів кластеризації

включають нечіткі c-means (FCM), k-means та покращені алгоритми k-means. При сегментації зображень в основному використовується алгоритм кластеризації k-means, оскільки він досить простий та ефективний. З іншого боку, алгоритм FCM поміщає пікселі в різні класи відповідно до їх різного ступеня належності. Найважливішими алгоритмами кластеризації для сегментації при опрацюванні зображень є:

Кластеризація на основі K-means

K-means – це простий алгоритм машинного навчання без вчителя. Він класифікує зображення за певною кількістю кластерів і починає процес з поділу простору зображення на k пікселів, які представляють k центроїдів групи. Потім кожен об'єкт відносять до групи на основі відстані між ними та центроїдом. Коли алгоритм призначив усі пікселі всім кластерам, він може переміщати та перепризначати центроїди.

Кластеризація на основі нечітких C-means

За допомогою методу кластеризації нечітких C-means пікселів на зображенні можна їх об'єднати в кілька кластерів. Це означає, що піксель може належати більше, ніж одному кластеру. Однак кожен піксель буде мати різні рівні подібності з кожним кластером. Алгоритм нечітких C-means має функцію оптимізації, яка впливає на точність результатів.

2.2.6. Нейронні мережі при сегментації

Згорткові нейронні мережі досить популярні при розв'язанні задач сегментації зображень, оскільки можуть ідентифікувати та обробляти дані зображення дуже швидко й ефективно. Експерти Facebook AI Research (FAIR) створили архітектуру глибокого навчання під назвою Mask R-CNN, яка може створити піксельну маску для кожного об'єкта, який присутній на зображенні. Це розширена версія архітектури виявлення об'єктів Faster R-CNN.

FAST R-CNN використовує дві частини даних для кожного об'єкта на зображенні, координати рамки та клас об'єкта. З Mask R-CNN одержують додатковий поділ у цьому процесі. Mask R-CNN виводить маску об'єкта після

виконання сегментації. У цьому процесі спочатку передається вхідне зображення до ConvNet, яка генерує карту об'єктів для зображення. Потім система застосовує мережу регіональних ознак (RPN) до карт об'єктів і генерує їх з оцінками об'єктивності. Після цього до виявлених ознак застосовується шар об'єднання, щоб привести їх до одного розміру. На останньому етапі система передає ознаки повнозв'язному шару для класифікації та генерує вихід з обмежувачими контурами для кожного об'єкта.

2.3. Оптимізація архітектури нейронної мережі Mask R-CNN для виявлення контурів рухомих об'єктів

Mask R-CNN концептуально проста архітектура: швидший R-CNN має два виходи для кожного об'єкта-кандидата, мітку класу та зміщення обмежувальної рамки. До цього пропонується додати третю гілку, яка виводить маску об'єкта. Таким чином, Mask R-CNN є природною та інтуїтивно зрозумілою ідеєю. Але вихід додаткової маски відрізняється від виходів класів і блоків, що вимагає вилучення набагато більш тонкого просторового макета об'єкта. Далі представимо ключові елементи Mask R-CNN, включаючи по піксельне вирівнювання, що є основним відсутнім елементом Fast/Faster R-CNN.

Fast R-CNN складається з двох частин. На першому етапі, який називається мережею пропозицій регіону (RPN), пропонуються обмежувальні рамки для об'єктів-кандидатів. Другий етап, який, по суті, є Fast R-CNN, добуває ознаки з використанням RoIPool з кожного поля-кандидата і виконує класифікацію та регресію обмежувальної рамки. Функції, які використовуються на обох етапах, можна спільно використовувати для швидшого результату.

Маска R-CNN використовує ту саму двоетапну процедуру з ідентичним першим етапом (який є RPN). На другому етапі, паралельно з прогнозуванням класу та зміщення блоку, Mask R-CNN також виводить двійкову маску для кожного RoI. Це є відмінною рисою від останніх систем, де класифікація залежить від прогнозування маски.

Запропонований підхід дотримується принципів Fast R-CNN, який паралельно застосовує класифікацію обмеженої рамки та регресію (що, як виявилось, значною мірою спрощує багатоетапний конвеєр оригінального R-CNN).

Формально під час навчання визначається втрата багатозадачності для кожного вибіркового RoI як

$$L = L_{cls} + L_{box} + L_{mask} \quad (2.1)$$

Класифікаційні втрати L_{cls} і втрати L_{box} в обмежувальній рамці ідентичні тим, які визначені у [12]. Гілка маски має розмір Km^2 для кожного RoI, який кодує K двійкових масок з роздільною здатністю $m \times m$, по одній для кожного з K класів.

Для цього застосовується сигмоїд на піксель і визначається L_{mask} як середня двійкову втрата крос-ентропії. Для RoI, пов'язаного з класом базової істини k , L_{mask} визначається лише на k -ій масці (інші вихідні дані маски не сприяють втратам).

Запропоноване визначення L_{mask} дозволяє мережі генерувати маски для кожного класу без змагання між ними, оскільки ставку зроблено на спеціальну гілку класифікації, щоб передбачити мітку класу, яка використовується для вибору маски виведення. Це роз'єднує маску та прогнозування класу, що відрізняється від звичайної практики при застосуванні FCN [30] до семантичної сегментації, яка зазвичай використовує softmax на піксель і багатоміальну втрату перехресної ентропії.

У цьому випадку маски між класами змагаються; у нашому випадку, із сигмоїдною системою на піксель і двійковою втратою, вони цього не роблять. Експериментами далі показано, що ця формула є ключовою для хороших результатів сегментації екземплярів.

Представлення маски: маска кодує просторовий макет вхідного об'єкта. Таким чином, на відміну від міток класів або зміщень блоків, які неминуче згортаються в короткі вихідні вектори повнозв'язними шарами, добування

просторової структури масок може бути вирішено природним шляхом за допомогою відповідності піксель до пікселя, що забезпечується згортками.

Зокрема, передбачимо маску $m \times m$ від кожного RoI за допомогою FCN. Це дозволяє кожному шару у гілці маски підтримувати явний просторовий макет об'єкта $m \times m$, не згортаючи його у векторне представлення, яке не має просторових розмірів.

На відміну від попередніх методів, які вдаються до використання FC-шарів для прогнозування маски, запропоноване повністю згорткове представлення вимагає менше параметрів і є більш точним, як продемонстровано далі в експериментах. Така поведінка піксель-до-пікселя вимагає, щоб функції RoI, які самі є невеликими картами об'єктів, були добре вирівняні, щоб точно зберегти явну просторову відповідність на піксель.

Це спонукало розробити наступний шар RoIAlign, який відіграє ключову роль у прогнозуванні маски.

RoIPool — це стандартна операція для добування невеликої карти об'єктів (наприклад, 7×7) з кожного RoI.

RoIPool спочатку квантує RoI з плаваючим числом до дискретної зернистості карти об'єктів, потім цей квантований RoI розбивається на просторові блоки, які самі квантуються, і, нарешті, значення об'єктів, що покриваються кожною ланкою, агрегуються (зазвичай шляхом максимального об'єднання).

Квантування виконується, наприклад, на безперервній осі координат x шляхом обчислення $[x/16]$, де 16 — це крок карти ознак, а $[\cdot]$ — округлення; так само квантування виконується під час поділу на бункери (наприклад, 7×7).

Ці квантування вносять неузгодженість між RoI та добутими ознаками. Хоча це може не вплинути на класифікацію, яка є надійною для невеликих областей, вона має великий негативний вплив на прогнозування точності у пікселях масок. Щоб вирішити це, пропонується шар RoIAlign, який усуває жорстке квантування RoIPool, належним чином узгоджуючи добуті ознаки з вхідними.

Запропонована зміна проста: вона полягає в уникненні будь-якого квантування границь або біномів RoI (тобто ми використовуємо $x/16$ замість $[x/16]$).

У даному випадку використовується білінійна інтерполяція для обчислення точних значень вхідних об'єктів у чотирьох місцях із регулярними вибірками в кожному контейнері RoI та агрегується результат (використовуючи максимальне або середнє значення), як показано на рис. 2.5.

Слід відмітити, що результати не чутливі до точних місць вибірки або кількості точок, які відбираються, доки не виконується квантування. RoIAlign призводить до значних покращень. На відміну від RoIAlign, RoIWarp не звертає увагу на проблему вирівнювання і реалізується як квантування RoI, як і RoIPool. Таким чином, навіть незважаючи на те, що RoIWarp також використовує біліїну передискретизацію, вона працює на рівні з RoIPool, демонструючи вирішальну роль вирівнювання.

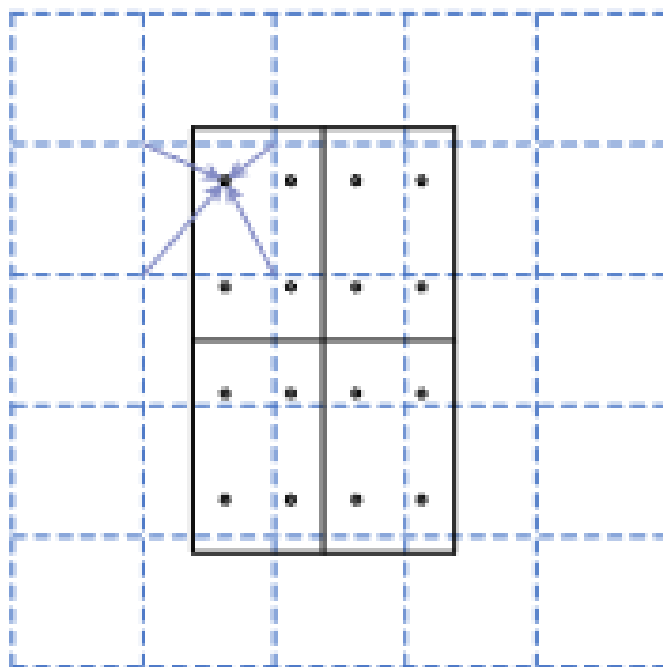


Рис 2.5. Формування RoIAlign

На рис.2.5 пунктирна сітка представляє карту об'єктів, суцільні лінії — RoI (з комірками 2×2 в даному прикладі), а точки – 4 точки вибірки в кожній комірці.

RoIAlign обчислює значення кожної точки вибірки шляхом білінійної інтерполяції з сусідніх точок сітки на карті об'єктів. Жодне квантування не виконується для будь-яких координат, які беруть участь у RoI, його комірках або точках вибірки.

2.4. Архітектура оптимізованої нейронної мережі

Щоб продемонструвати загальність підходу, створюється Mask R-CNN з кількома архітектурами. Для наочності розрізняють:

- архітектуру згорткової магістралі, яка використовується для добування ознак з усього зображення,
- заголовок мережі для розпізнавання обмежувальної рамки (класифікація та регресія) і прогнозування маски, яка застосовується окремо до кожного RoI.

У даному випадку архітектуру магістралі позначено за допомогою номенклатури *network-depth-features*. Далі виконується оцінка мережі ResNet і ResNeXt глибиною 50 або 101 шар. Оригінальна реалізація Faster R-CNN з ResNets добуває ознаки з останнього згорткового шару 4-го етапу, який названо C4.

Ця магістраль із ResNet-50, наприклад, позначається ResNet-50-C4. Також досліджується інший більш ефективний хребет, що називається мережею пірамід функцій (FPN).

FPN використовує архітектуру «зверху вниз» із бічними з'єднаннями для побудови піраміди функцій у мережі з одномасштабного входу. Faster R-CNN з основою FPN добуває функції RoI з різних рівнів піраміди функцій відповідно до їх масштабу, але в іншому підхід схожий на звичайний ResNet.

Використання магістралі ResNet-FPN для добування функцій із Mask R-CNN дає чудовий приріст як точності, так і швидкості.

Для заголовку мережі уважно обираємо архітектури [19], до яких додається повністю згорткова гілка прогнозування маски. Зокрема, розширюються заголовки блоків Faster R-CNN з документів ResNet та FPN. Деталі архітектурного рішення показано на рис. 2.6.

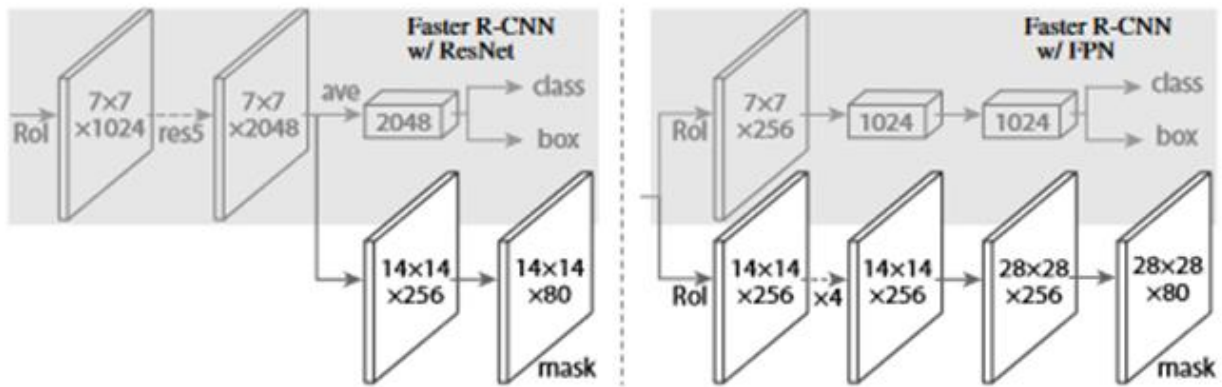


Рис. 2.6. Архітектура заголовку запропонованої нейронної мережі

На рис. 2.6 ліва/права частини показують заголовки для магістралей ResNet C4 і FPN, до яких додано гілку маски.

Цифри позначають просторову роздільну здатність і канали. Стрілки позначають шари conv, deconv або fc, як можна зробити висновок із контексту (conv зберігає просторовий вимір, а deconv збільшує його).

Усі згорткові шари мають розмір 3×3 , за винятком вихідного, який дорівнює 1×1 . Deconv шар має розмір 2×2 із кроком 2, і використовує ReLU у прихованих шарах.

Ліворуч на рис. 2.6 «res5» позначає п'яту стадію ResNet, яку для простоти змінили так, що перший згортковий шар працює на 7×7 RoI з кроком 1, замість 14×14 з кроком 2. Праворуч: « $\times 4$ » означає набір з чотирьох послідовних згорткових шарів.

Головна частина магістралі ResNet-C4 містить 5-й етап ResNet (а саме, 9-шаровий 'res5'), який пов'язаний з використанням обчислень. Для FPN магістраль вже включає res5 і, таким чином, забезпечує більш ефективний заголовок, який використовує менше фільтрів. Зауважимо, що гілки маски мають просту структуру. Більш складні конструкції можуть покращити продуктивність.

2.5. Застосування принципу Transfer Learning при визначенні контурів рухомих об'єктів

Для забезпечення продуктивності застосування нейромереж, орієнтованих на вирішення задач розпізнавання образів та виявлення меж локалізації об'єктів запропоновано скористатись підходом «передачі знань». Це дозволить скоротити часові і фінансові затрати при її навчанні та створенні вхідних фреймів даних.

Підхід Transfer Learning базується на застосуванні визначених експериментальних шляхом архітектур розпізнавання образів. Такі архітектури, зазвичай, проектуються і використовуються потужними компаніями у сфері інформаційних технологій, зокрема, Amazon, Google, IBM та ін. На концептуальному рівні процедура Transfer Learning передбачає передачу знань одного об'єкту іншому (рис. 2.7).



Рис. 2.7. Принцип передачі знань

Підхід Transfer Learning демонструє вищу ефективність реалізації і проектування нейронних мереж, порівнюючи з підходом, що використовує випадкову ініціалізацію значень вікна. Це пов'язано з тим, що оптимізуються параметри завершального повнозв'язного шару, що в свою чергу дозволяє значно скоротити часові ресурси та прискорити розв'язання задач на користувачькому фреймі даних.

По суті Transfer Learning надає можливість використання існуючої архітектури нейронної мережі, що є заздалегідь навченою на великих наборах даних. Однак, крім компонентів архітектури, при використанні попередньо навченої нейронної мережі у користувацькому проєкті, можна застосувати уже визначені синаптичні ваги та забезпечити донавчання на конкретних зображеннях.

При цьому, окрім архітектури можна використати значення синаптичних ваг і/або використати повністю новий набір зображень чи довчити на існуючих даних.

Рис. 2.8 демонструє принцип передачі знань.

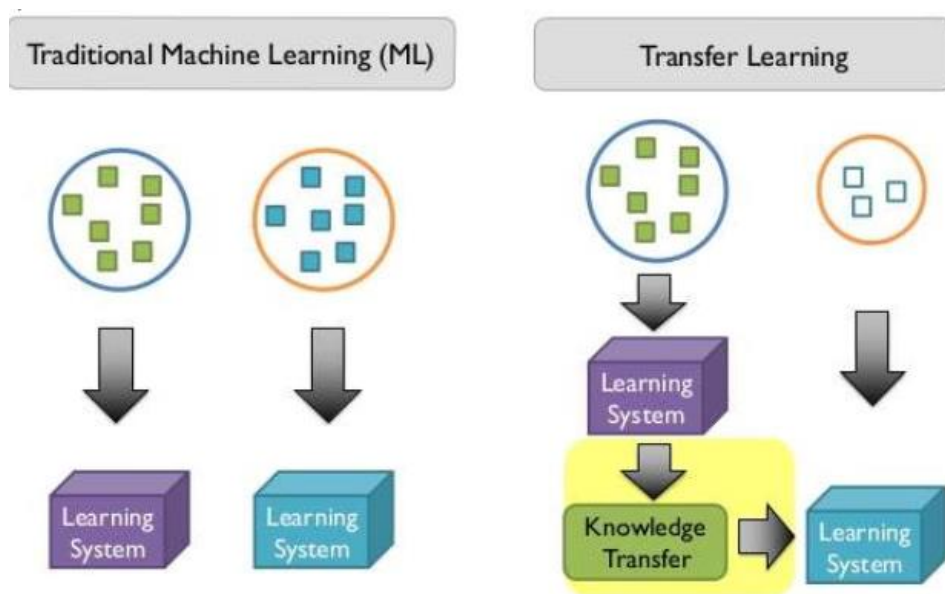


Рис. 2.8. Відмінності підходів традиційного машинного навчання та підходу передачі знань

Тому раціонально при виявленні і встановленні контурів рухомих об'єктів застосовувати підхід Transfer Learning, що дозволить значно скоротити навчання нейронних мереж та забезпечити ефективність запропонованого вище методу.

2.6. Висновки до розділу

У даному розділі одержано наступні наукові та практичні результати:

1. Запропоновано модель архітектурного рішення щодо реалізації інтелектуалізованої системи відеоспостереження з функцією виявлення контурів рухомих об'єктів, яка використовує апаратний модуль ESP32-CAM та відеокамеру OV2640 і дає можливість з високою точністю на основі підходу сегментації екземплярів накладати та візуалізувати маски на кадрах відеопотоку.

2. Проведено детальний аналіз методів сегментації об'єктів, що дало можливість визначити кращі рішення при імплементації системи відеоспостереження з функцією виявлення контурів, які базуються на основі фреймворків Mask R-CNN та Faster R-CNN та підходу Transfer Learning.

3. Запропоновано рішення щодо оптимізації структури нейронної мережі R-CNN на основі побудови піраміди функцій у мережі з одномасштабного входу, що дає змогу скоротити витрати часу на навчання мережі та є окремою гілкою при виявленні контурів рухомих об'єктів у системах відеоспостереження.

РОЗДІЛ 3

ПРОГРАМУВАННЯ ІНТЕЛЕКТУАЛЬНОЇ СКЛАДОВОЇ ДЛЯ СИСТЕМИ ВІДЕОСПОСТЕРЕЖЕННЯ

3.1. Обґрунтування фреймворку для реалізації моделі виявлення контурів рухомих об'єктів

Mask R-CNN, по суті, є розширенням Faster R-CNN. Fast R-CNN широко використовується для виявлення об'єктів на зображеннях. Для конкретного зображення він повертає мітку класу та координати рамки для кожного об'єкта, який на ньому виявлено.

Сегментація зображення створює піксельну маску для кожного об'єкта, наявного на зображенні і дає набагато більш детальне його розуміння. Зображення (рис. 3.1), що показано нижче наочно демонструє поняття сегментації зображень.

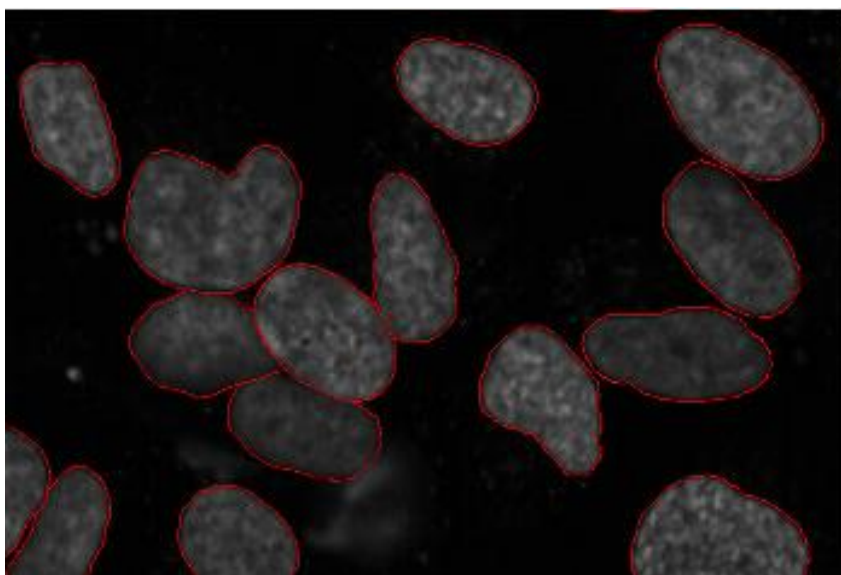


Рис. 3.1. Демонстрація сегментації об'єктів

На рис. 3.1. можна побачити, що кожен об'єкт (який є клітиною на даному зображенні) був сегментований. Це принцип за яким працює сегментація зображень.

Як зазначалось раніше у розділі 2, існує кілька принципів до виявлення контурів об'єктів, що залежить від цілі поставленої задачі. Це означає, що можливе застосування двох типів сегментації зображень:

- семантичну сегментацію;
- сегментація екземплярів.

На рис. 3.2 показано різницю між семантичною та сегментацією екземплярів.

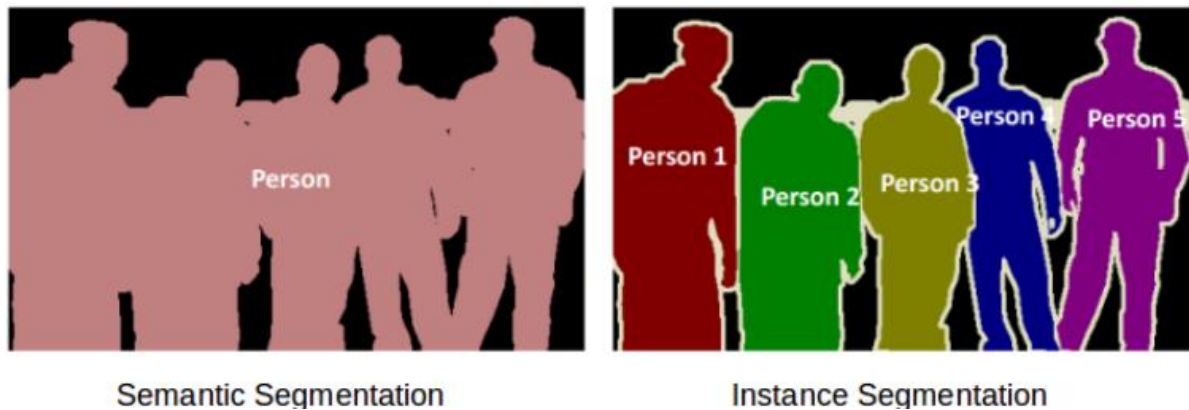


Рис. 3.2. Семантична сегментація та сегментація екземплярів

На рис. 3.2, усі 5 об'єктів на зображенні ліворуч – це люди. Отже, семантична сегментація буде класифікувати всіх людей як одиничний екземпляр. Тепер на зображенні праворуч також є 5 об'єктів (усі вони люди). Але тут по-іншому об'єкти одного класу були призначені як різні екземпляри.

У роботі реалізується сучасний метод сегментації, що використовує модифіковану техніку Mask R-CNN для вирішення проблеми сегментації екземплярів.

Mask R-CNN, по суті, є розширенням Faster R-CNN, що широко використовується для виявлення об'єктів. Для заданого зображення цей метод повертає мітку класу та координати рамки для кожного об'єкта, що присутнє на зображенні. Отже, припустимо, ми маємо зображення, яке показано на рис. 3.3.



Рис. 3.3. Приклад тестового зображення

У результаті застосування моделі Fast R-CNN одержимо результат на зразок того, яке показано на рис. 3.4.

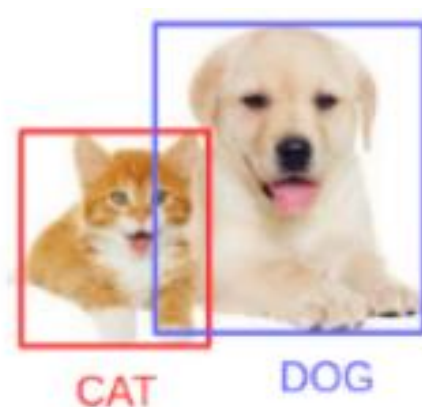


Рис. 3.4. Результат опрацювання зображення Fast R-CNN

Фреймворк Mask R-CNN побудований на основі Faster R-CNN. Отже, для даного зображення Mask R-CNN, на додаток до мітки класу та координат рамки для кожного об'єкта, також буде повертати маску об'єкта.

Принцип функціонування Faster R-CNN:

- Faster R-CNN спочатку використовує ConvNet для добування карт ознак об'єктів із зображень;
- виявлені карти об'єктів передаються через мережу регіональних пропозицій («Region Proposal Network – RPN»), яка повертає обмежувальні рамки кандидата;

- застосовується шар об'єднання ROI до цих обмежувальних квадратів-кандидатів, щоб привести всіх кандидатів однакового розміру;
- пропозиції передаються на повнозв'язний шар для класифікації та виведення границь рамок об'єктів.

Основна модель (модель «хребта») подібна до ConvNet, що використовується у Faster R-CNN для добування карт ознак із зображення. Пропонується в якості такої моделі використати архітектуру ResNet 101 для добування ознак із зображень у Mask R-CNN.

Отже, перший крок – це аналіз зображень та добування ознак за допомогою архітектури ResNet 101. Ці ознаки виступають як вхідні дані для наступних шарів.

Наступний крок полягає у створення мережі регіональних пропозицій (RPN). Для цього беруться карти ознак, отримані на попередньому кроці, і до них застосовується мережа пропозицій регіону (RPM). Це в основному визначає, чи належить об'єкт до даної області чи ні. На даному кроці одержують ці області або карти ознак, які прогнозує модель щодо наявності якогось об'єкту.

Регіон інтересів (RoI)

Регіони, отримані від RPN, можуть мати різну форму. Отже, потрібно інтегрувати об'єднуючий шар і перетворити та привести всі області до однакової форми. Далі ці регіони пропускаються через повнозв'язні мережі, щоб спрогнозувати мітку класу та обмежувальні рамки. До цього моменту кроки майже схожі на те, як працює Faster R-CNN. Тепер виникає різниця між двома рамками.

На додаток до цього Mask R-CNN також створює маску сегментації. Для цього спочатку обчислюється область інтересів, щоб можна було скоротити час обчислення.

Для всіх прогнозованих областей, обчислюється перетин через об'єднання (IoU) з основними квадратами істини. Показник IoU обчислюється наступним чином:

$$IoU = \frac{\text{Площа перетину}}{\text{Площа об'єднання}} \quad (3.1)$$

Тепер, лише якщо IoU більше або дорівнює 0,5, одразу розглядається це як область інтересу. В іншому випадку можна знехтувати цією конкретною областю. Ця процедура виконується для всіх регіонів, а потім вибирається лише набір областей для яких IoU більше 0,5. Розглянемо зображення, наведене на рис. 3.5.



Рис. 3.5. Приклад виявлення контурів об'єкту

На рис. 3.5 червоні контури є основною рамки істинності для цього зображення. Тепер, припустимо, що одержано 4 регіони від RPN як показано на рис. 3.6.

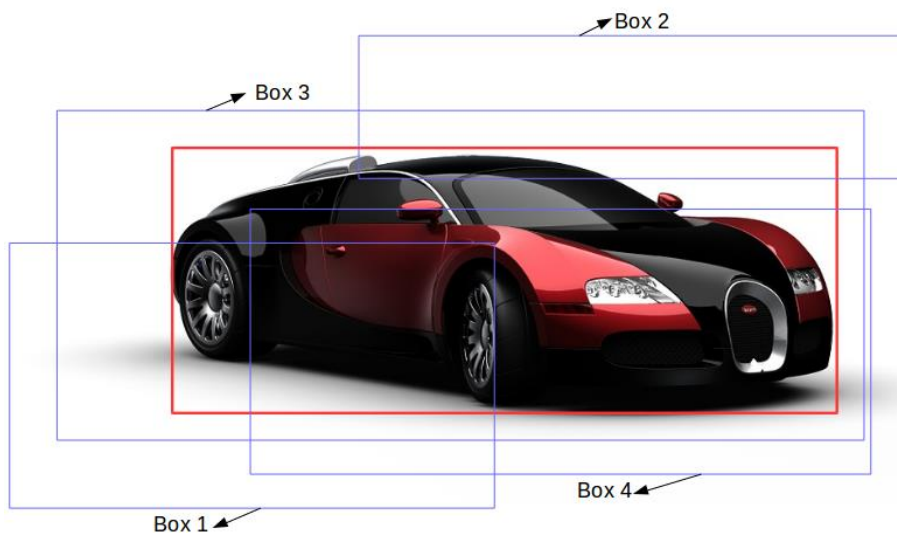


Рис. 3.6. Контури одержані від RPN

На рис. 3.6 IoU контуру 1 і контуру 2, можливо, менше 0,5, тоді як IoU контуру 3 і контуру 4 трохи перевищують значення 0,5. Отже, можна зробити

висновок, що контур 3 і контур 4 є областю інтересу для цього конкретного зображення, тоді як контурами 1 і 2 буде знехтовано. Далі потрібно перейти до сегментації шляхом нанесення маски на об'єкт.

Після того, як отримано RoI на основі значень IoU, можна додати гілку маски до існуючої архітектури. Це дозволить повернути маску сегментації для кожної області, що містить об'єкт. Розмір маски, що повертається, має розмір 28×28 пікселів для кожної області, яка потім може збільшуватись. Розглянемо наступний приклад (рис. 3.7).



Рис. 3.7. Вхідне зображення для формування контурів регіону

Маска при проведенні сегментації об'єктів із застосування Mask R-CNN матиме вигляд, подібний до того, як показано на рис. 3.8.



Рис. 3.8. Застосування маски з Mask R-CNN

На рис. 3.8 модель сегментувала всі об'єкти на зображенні. Це останній крок у Mask R-CNN, що виконує прогнозування маски для всіх об'єктів на зображенні. Варто відзначити, що час навчання для Mask R-CNN досить великий і може займати приблизно 1-2 дні, щоб навчити Mask R-CNN на відомому наборі даних COCO. Тому, доцільно скористатися підходом Transfer Learning, який описаний в останньому підрозділі розділу 2.

При використанні підходу передачі знань пропонується застосувати ваги навченої моделі Mask R-CNN на наборі даних COCO. Алгоритм реалізації моделі Mask R-CNN для виконання сегментація екземплярів на статичних зображеннях розглянуто у наступному підрозділі.

3.2. Алгоритм реалізації моделі для виявлення контурів об'єктів на статичних зображеннях

До основних етапів реалізації Mask R-CNN, які формують послідовність дій щодо сегментації об'єктів, а отже і визначення їх контурів належить 4 основних стадії:

- клонування сховища Mask R-CNN;
- встановлення залежностей бібліотек;
- завантаження попередньо навчених ваг;
- програмна імплементація моделі.

Для виконання першого кроку пропонується використовувати фреймворк Mask R-CNN, створений дослідниками Facebook AI Research (FAIR).

Спочатку виконується клонування репозиторію Mask R-CNN, який містить архітектуру для Mask R-CNN. Для клонування виконується команда:

```
clone git https://github.com/matterport/Mask_RCNN.git
```

Як тільки це буде зроблено, необхідно встановити залежності для Mask R-CNN, що передбачає виконання другого кроку алгоритму.

Перелік усіх залежностей, необхідних для ефективного функціонування Mask R-CNN наведено нижче:

- numpy;
- scipy;
- Pillow;
- cython;
- matplotlib;
- scikit-image;
- tensorflow>=1.3.0;
- keras>=2.0.8;
- opencv-python;
- h5py;
- imgaug;
- IPython.

Наведені вище залежності повинні бути встановленими перед безпосереднім використанням Mask R-CNN.

Третій крок алгоритму полягає у завантаженні попередньо навчених на наборі даних MS COCO синаптичних ваг. Завантаження можна виконати скориставшись посиланням «https://github.com/matterport/Mask_RCNN/releases». Коли ваги збережено на локальній машині у вигляді відповідного файлу, його потрібно перемістити у каталог зразків репозиторію Mask_RCNN, який клоновано на першому кроці алгоритму.

Завершальним етапом алгоритму є використання архітектури Mask R-CNN з коефіцієнтами, одержаними з попередньо навчених нейронних мереж, для прогнозування контурів власних образів. Після виконання цих чотирьох кроків настав час перейти до реалізації розпізнавання контурів об'єктів засобами мови програмування Python.

3.3. Програмна реалізація алгоритму виявлення контурів об'єктів та сегментації об'єктів

Реалізація будь-якого програмного забезпечення засобами мови програмування Python починається з імпорту потрібних бібліотек. У даному випадку, тих, які визначено на другому кроці алгоритму. Лістинг 3.1 демонструє імпортовані бібліотеки для реалізації Mask R-CNN.

Лістинг 3.1. Імпорт необхідних бібліотек

```
import os
import sys
import random
import math
import numpy as np
import skimage.io
import matplotlib
import matplotlib.pyplot as plt
# Root directory of the project
ROOT_DIR = os.path.abspath("../")
import warnings
warnings.filterwarnings("ignore")
# Import Mask RCNN
sys.path.append(ROOT_DIR) # To find local version of the
library
from mrcnn import utils
import mrcnn.model as modellib
from mrcnn import visualize
# Import COCO config
sys.path.append(os.path.join(ROOT_DIR, "samples/coco/")) # To
find local version
import coco
%matplotlib inline
```

Наступний крок полягає у визначенні шляхів для попередньо навчених ваг і зображень, на яких потрібно виконати сегментацію. У лістингу 3.2 реалізовано процедуру звернення до значень попередньо визначених параметрів.

Лістинг 3.2. Процедура звернення до попередньо навчених даних

```
# Directory to save logs and trained model
MODEL_DIR = os.path.join(ROOT_DIR, "logs")

# Local path to trained weights file
COCO_MODEL_PATH = os.path.join('.', "mask_rcnn_coco.h5")

# Download COCO trained weights from Releases if needed
if not os.path.exists(COCO_MODEL_PATH):
    utils.download_trained_weights(COCO_MODEL_PATH)

# Directory of images to run detection on
IMAGE_DIR = os.path.join(ROOT_DIR, "images")
```

Далі необхідно створити клас, який буде використовуватися для визначення моделі Mask R-CNN (лістинг 3.3).

Лістинг 3.3. Оголошення моделі Mask R-CNN

```
class InferenceConfig(coco.CocoConfig):
    # Set batch size to 1 since we'll be running inference on
    # one image at a time. Batch size = GPU_COUNT *
    IMAGES_PER_GPU
    GPU_COUNT = 1
    IMAGES_PER_GPU = 1

config = InferenceConfig()
config.display()
```

У результаті ініціалізації моделі Mask R-CNN, одержано результат, як показано на рис. 3.9.

З рис. 3.9 можна побачити сукупність специфікацій моделі Mask R-CNN, які будуть використовуватися. Отже, основою моделі є ResNet101. Форма маски, яку повертає модель, має розмір 28×28 , оскільки вона навчається на наборі даних COCO.

Всього наявним є 81 клас (включаючи фоновий). Окрім цього, можна спостерігати різні інші статистичні дані, наприклад:

- вхідний розмір;
- кількість використовуваних графічних процесорів;
- етапи перевірки.

```

Configurations:
BACKBONE                resnet101
BACKBONE_STRIDES        [4, 8, 16, 32, 64]
BATCH_SIZE              1
BBOX_STD_DEV            [0.1 0.1 0.2 0.2]
COMPUTE_BACKBONE_SHAPE None
DETECTION_MAX_INSTANCES 100
DETECTION_MIN_CONFIDENCE 0.7
DETECTION_NMS_THRESHOLD 0.3
FPN_CLASSIF_FC_LAYERS_SIZE 1024
GPU_COUNT               1
GRADIENT_CLIP_NORM      5.0
IMAGES_PER_GPU          1
IMAGE_CHANNEL_COUNT     3
IMAGE_MAX_DIM           1024
IMAGE_META_SIZE         93
IMAGE_MIN_DIM           800
IMAGE_MIN_SCALE         0
IMAGE_RESIZE_MODE       square
IMAGE_SHAPE             [1024 1024  3]
LEARNING_MOMENTUM       0.9
LEARNING_RATE           0.001
LOSS_WEIGHTS            {'rpn_class_loss': 1.0, 'rpn_bbox_loss': 1.0, 'mrcnn_class_loss': 1.0, 'mrcnn_bbox_loss': 1.0, 'mrcnn_mask_loss': 1.0}
MASK_POOL_SIZE          14
MASK_SHAPE              [28, 28]
MAX_GT_INSTANCES        100
MEAN_PIXEL              [123.7 116.8 103.9]
MINI_MASK_SHAPE         (56, 56)
NAME                    coco
NUM_CLASSES              81
POOL_SIZE               7
POST_NMS_ROIS_INFERENCE 1000
POST_NMS_ROIS_TRAINING  2000
PRE_NMS_LIMIT           6000
ROI_POSITIVE_RATIO      0.33
RPN_ANCHOR_RATIOS       [0.5, 1, 2]
RPN_ANCHOR_SCALES       [32, 64, 128, 256, 512]
RPN_ANCHOR_STRIDE       1
RPN_BBOX_STD_DEV        [0.1 0.1 0.2 0.2]
RPN_NMS_THRESHOLD       0.7
RPN_TRAIN_ANCHORS_PER_IMAGE 256
STEPS_PER_EPOCH         1000
TOP_DOWN_PYRAMID_SIZE   256
TRAIN_BN                 False
TRAIN_ROIS_PER_IMAGE    200
USE_MINI_MASK           True
USE_RPN_ROIS            True
VALIDATION_STEPS        50
WEIGHT_DECAY            0.0001

```

Рис. 3.9. Властивості Mask R-CNN

Далі потрібно створити власну модель і завантажити попередньо підготовлені ваги, які завантажено на третьому кроці алгоритму. Тут варто переконатися в тому, що попередньо підготовлені ваги знаходяться в тій самій папці, бо в іншому випадку доведеться вказати розташування файлу ваг (лістинг 3.4)

Лістинг 3.4. Завантаження ваг для власної моделі Mask R-CNN

```
# Create model object in inference mode.
model = modellib.MaskRCNN(mode="inference",
model_dir='mask_rcnn_coco.h5', config=config)

# Load weights trained on MS-COCO
model.load_weights('mask_rcnn_coco.h5', by_name=True)
```

Тепер доцільно визначити класи набору даних COCO, які необхідні будуть на етапі прогнозування. Для цього виконується лістинг 3.5.

Лістинг 3.5. Визначення класи об'єктів даних COCO

```
# COCO Class names
class_names = ['BG', 'person', 'bicycle', 'car', 'motorcycle',
'airplane', 'bus', 'train', 'truck', 'boat', 'traffic light',
'fire hydrant', 'stop sign', 'parking meter', 'bench', 'bird',
'cat', 'dog', 'horse', 'sheep', 'cow', 'elephant', 'bear',
'zebra', 'giraffe', 'backpack', 'umbrella', 'handbag', 'tie',
'suitcase', 'frisbee', 'skis', 'snowboard', 'sports ball',
'kite', 'baseball bat', 'baseball glove', 'skateboard',
'surfboard', 'tennis racket', 'bottle', 'wine glass', 'cup',
'fork', 'knife', 'spoon', 'bowl', 'banana', 'apple',
'sandwich', 'orange', 'broccoli', 'carrot', 'hot dog',
'pizza',
'donut', 'cake', 'chair', 'couch', 'potted plant', 'bed',
'dining table', 'toilet', 'tv', 'laptop', 'mouse', 'remote',
'keyboard', 'cell phone', 'microwave', 'oven', 'toaster',
'sink', 'refrigerator', 'book', 'clock', 'vase', 'scissors',
'teddy bear', 'hair drier', 'toothbrush ']
```

Після цього завантажимо зображення і спробуємо побачити, як працює модель. Тут можна використовувати будь-які зображень для перевірки ефективності моделі (лістинг 3.5).

Лістинг 3.5. Завантаження зображень

```
# Load a random image from the images folder
image = skimage.io.imread('sample.jpg')
```

```
# original image
plt.figure(figsize=(12,10))
skimage.io.imshow(image)
```

Завантажене зображення показано на рис. 3.10.



Рис. 3.10. Результат завантаження зображення

Саме із зображенням, що було виділено з відеопотоку (рис. 3.10), буде продовжене дослідження. З рис. 3.10 можна чітко встановити, що на зображенні наявні два автомобілі (один спереду і один ззаду) разом із велосипедом. Далі будемо використовувати модель Mask R-CNN разом із попередньо підготовленими вагами і подивимося, наскільки добре вона сегментує об'єкти на зображенні. Для початку виконаємо прогнозування на основі створеної моделі (лістинг 3.11), а потім побудуємо результати для їх візуалізації.

Лістинг 3.6. Прогнозування контурів об'єктів

```

# Run detection
results = model.detect([image], verbose=1)

# Visualize results
r = results[0]
visualize.display_instances(image, r['rois'], r['masks'],
r['class_ids'], class_names, r['scores'])

```

Результат виконання лістингу 3.6 показаний на рис. 3.11.

```

Processing 1 images
image           shape: (480, 640, 3)   min:  0.00000  max: 255.00000  uint8
molded_images  shape: (1, 1024, 1024, 3) min: -123.70000 max: 151.10000  float64
image metas    shape: (1, 93)         min:  0.00000  max: 1024.00000  float64
anchors        shape: (1, 261888, 4)  min: -0.35390  max:  1.29134    float32

```

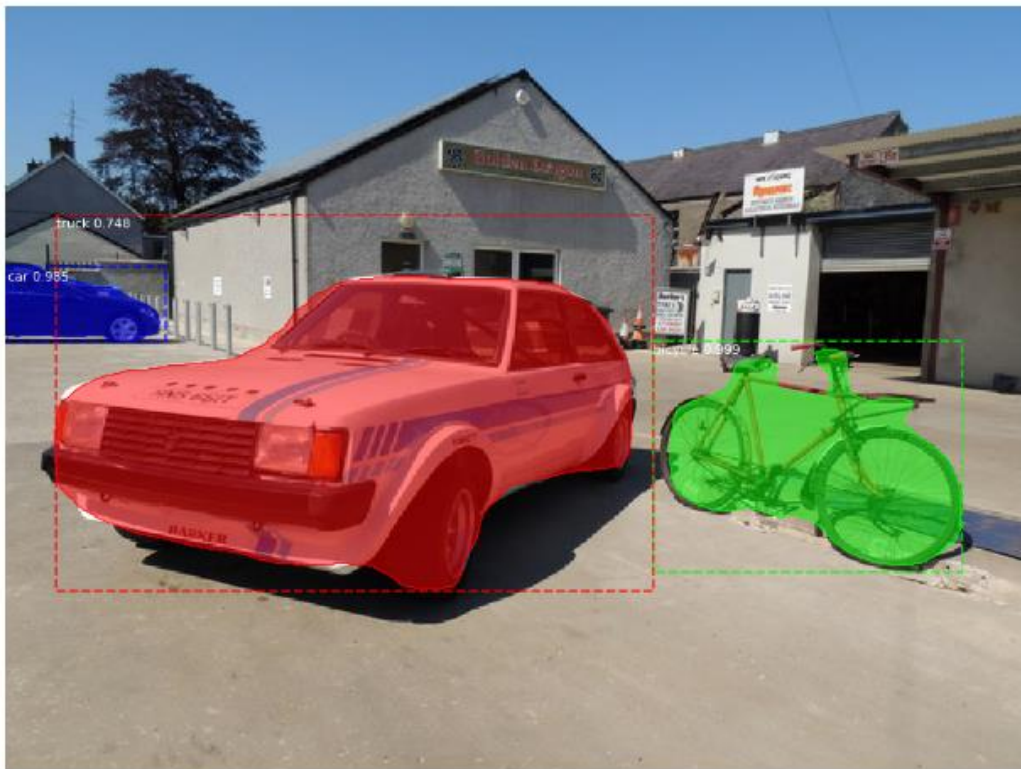


Рис. 3.11. Результати прогнозування моделі

Модель досить добре просегментувала як автомобілі, так і велосипед на зображенні. Також можна розглянути кожну маску або сегментовані об'єкти окремо. Для цього спочатку потрібно всі маски, визначені моделлю, зберегти у

відповідній змінній. Тепер вони мають булеву форму (True і False), тому потрібно перетворити їх у числа (1 і 0). Для цього використовується програмний код, наведений у лістингу 3.7.

Лістинг 3.7. Перетворення булевого представлення у цілочисельне

```
mask = r['masks']
mask = mask.astype(int)
mask.shape
```

У результаті виконання лістингу 3.7 одержимо масив 0 і 1, де 0 означає, що в цьому пікселі немає об'єкта, а 1 – що в цьому пікселі він наявний. Формат результату має вигляд – (480, 640, 3). Варто зауважити, що форма маски подібна до форми вихідного зображення. Однак цифра «3» тут у формі маски не представляє канали. Натомість вона відображає кількість об'єктів, сегментованих моделлю. Оскільки модель ідентифікувала 3 об'єкти на наведеному вище зразку зображення, форма маски є (480, 640, 3).

Якби було 5 об'єктів, ця форма була б наступного вигляду – (480, 640, 5). Тепер одержано оригінальне зображення та набір масок. Для того, щоб відобразити або отримати кожен сегмент із зображення, створено цикл for і виконано множення кожної маски на вихідне зображення (лістинг 3.8).

Лістинг 3.8. Формування контурів кожного екземпляра об'єктів

```
for i in range(mask.shape[2]):
    temp = skimage.io.imread('sample.jpg')
    for j in range(temp.shape[2]):
        temp[:, :, j] = temp[:, :, j] * mask[:, :, i]
plt.figure(figsize=(8, 8))
plt.imshow(temp)
```

Результати виконання лістингу 3.8 показано на рис. 3.12.

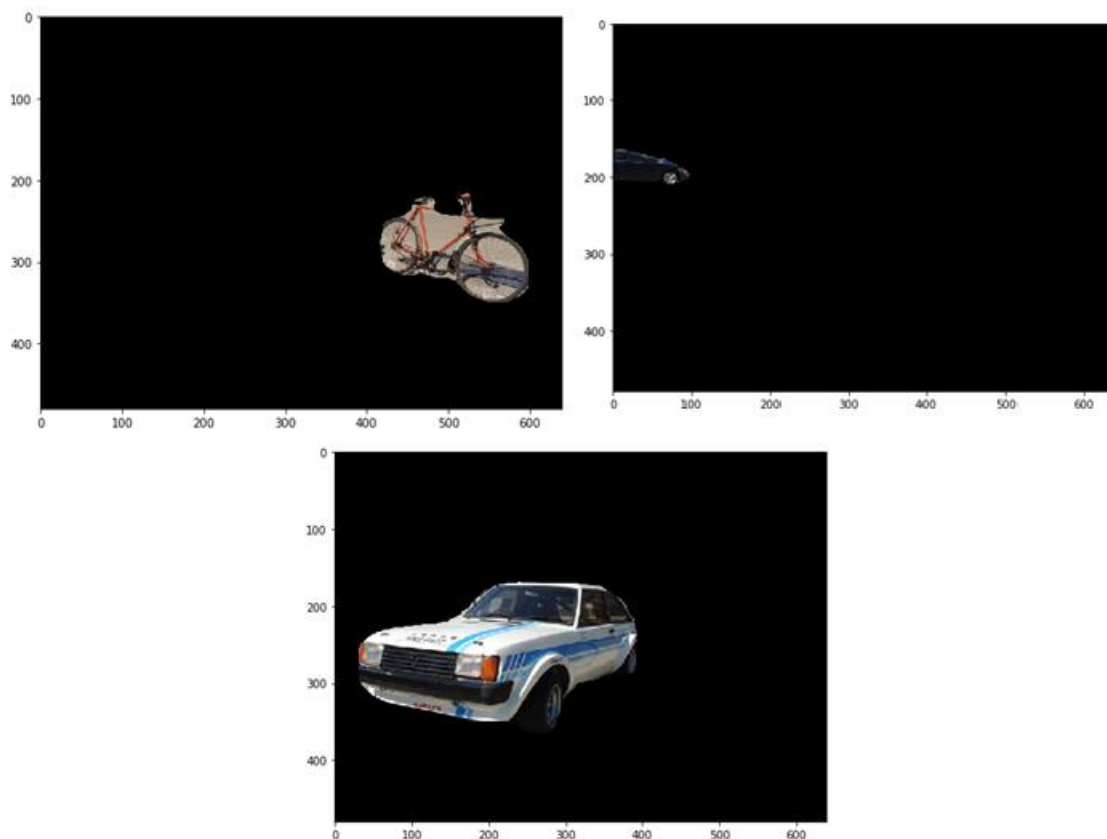


Рис. 3.12. Сегментація об'єктів як окремих екземплярів

Таким чином, реалізовано виявлення контурів об'єктів із застосуванням оптимізованої архітектури та моделі Mask R-CNN, що дало можливість переконатися в її ефективності.

3.4. Процедура виявлення контурів рухомих об'єктів

Раніше було досліджено застосування моделі Mask R-CNN на прикладі групи статичних зображень. Для виявлення об'єктів з відеопотоку, виконуються практично ті самі операції, однак перед тим потрібно опрацювати відео для розбиття його на кадри. Для прикладу візьмемо відео запуску/приземлення Falcon-9 і створимо невелику програму для розпізнавання цього ракетоносія.

Однак перед тим, як виконувати автоматичну ідентифікацію у відеопотоці об'єкта Falcon-9 потрібно спочатку виконати виділення кадрів з відео та сформувати і розмітити відповідні зображення.

Тому спочатку виконаємо фрагмент лістингу для утворення навчального набору зображень. Для цього пропонується з відео виділяти кожен п'ятий кадр (лістинг 3.9).

Лістинг 3.9. Виділення кадрів з відео

```
from detecto import core, visualize
from detecto.utils import split_video
import time
#Start time
startTime = time.time()
"Splitting training video on frames"
split_video('VIDEO_FOR_PREDICT/Falcon_Heavy_Test_Flight.mp4',
            'VIDEO_FOR_TRAINING/',
            step_size=5)
```

У результаті виконання лістингу 3.9, одержано набір зображень. З відео тривалістю 15 секунд виділено 77 кадрів (рис. 3.13)

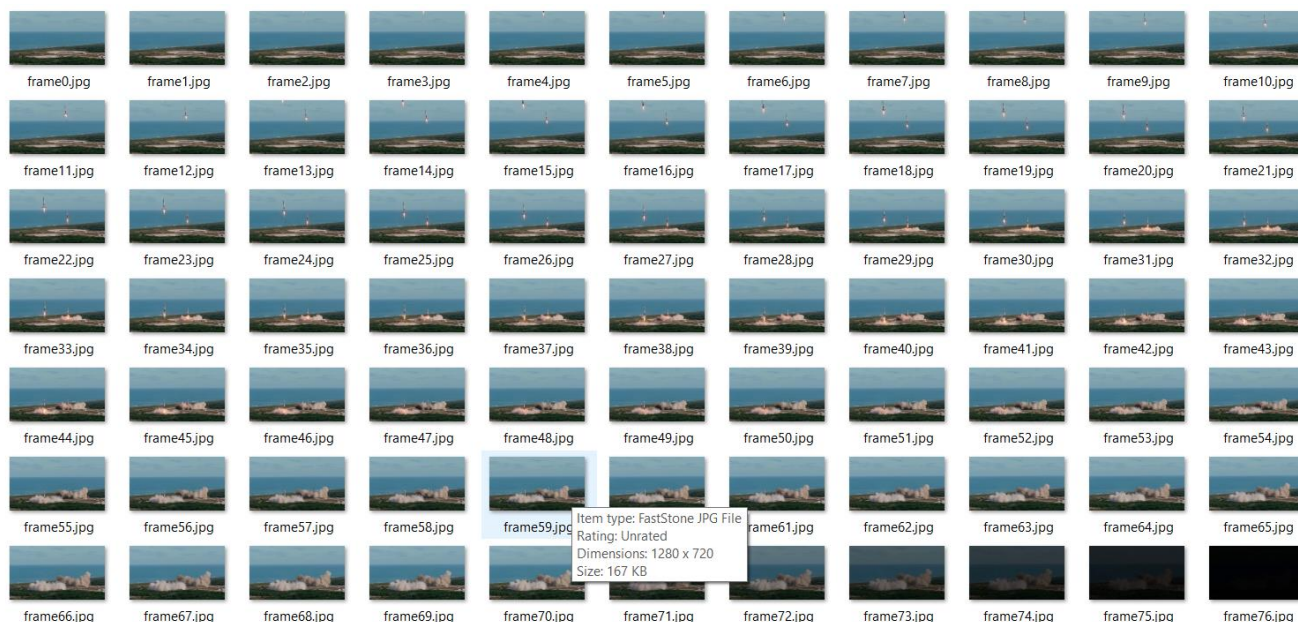


Рис. 3.13. Перелік кадрів з відео

Після того, як одержано зображення з відео, потрібно провести асесорну розмітку об'єктів. Для цього можна використати такий застосунок, як LabelImg (рис. 3.14).

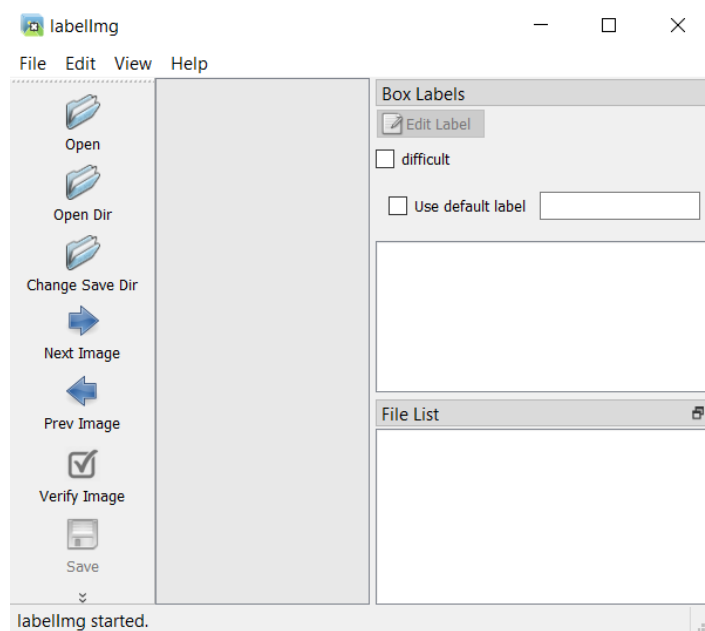


Рис. 3.14. Програма для розмітки зображень

Графічне представлення нульового кадру та приклади розмічених зображень показано відповідно на рис. 3.15 та рис. 3.16-3.17.

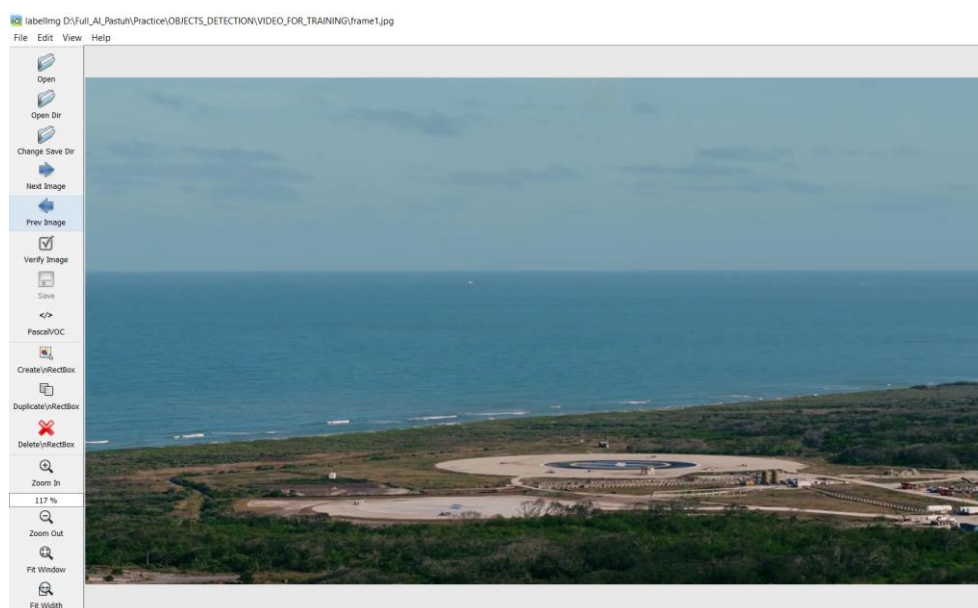


Рис. 3.15. Зображення нульового кадру з відео

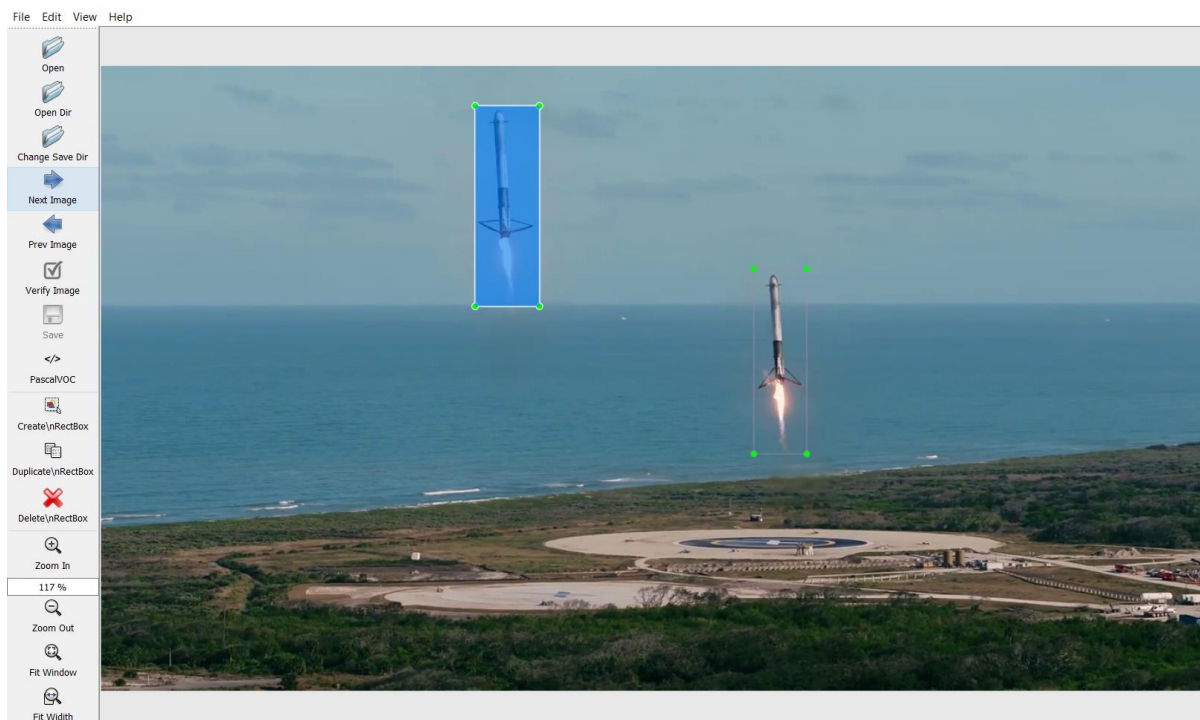


Рис. 3.16. Кадр з ракетноносіями в повітрі



Рис. 3.17. Кадр з мітками ракетноносіїв при приземленні

Програмний код застосування моделі розпізнавання об'єктів на відео приведено у лістингу 3. 10.

Лістинг 3.10. Програмна реалізація моделі виявлення об'єктів у відеопотоці

```

"""Training"""
#Download frames
frames = core.Dataset('VIDEO_FOR_TRAINING/')

#Initialization of model
model = core.Model(['falcon'])
"""
#ATTENTION!!! MACHINE VERY LONG LEARNED, APPROXIMATELY, 12
HOURS!!!
#Fitting model
model.fit(frames)
"Save and load model"
#Save model
model.save('model_weights_for_video.pth')
"""
#Load model
pretrained_model =
core.Model.load('model_weights_for_video.pth', ['falcon'])
"""Visualization"""
visualize.detect_video(pretrained_model,
'VIDEO_FOR_PREDICTION/Falcon_Heavy_Test_Flight.mp4',
'VIDEO_FOR_PREDICTION/output.avi')
"""Program run time"""
#End time
endTime = time.time()
print("#####")
print('Program run time = {0} second'.format(endTime -
startTime))
print("#####")

```

У результаті, одержано відео з накладеним контуром, що детектує Falcon-9, фрагмент якого показано на рис. 3.18.

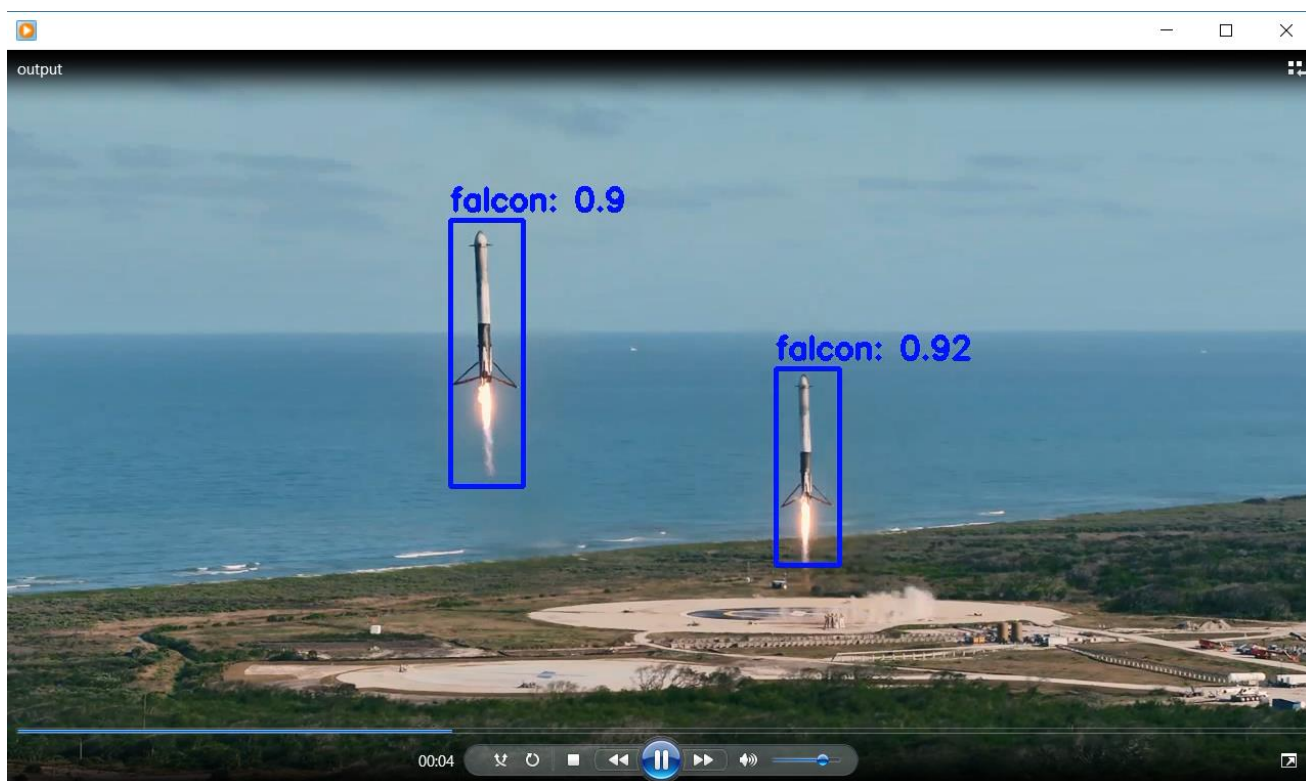


Рис. 3.18. Результат розпізнавання рухомих об'єктів

Таким чином, результат прогнозування контурів та сегментації об'єктів показує хороші результати, а оптимізовано модель Mask R-CNN є ефективною у застосуванні.

3.5. Висновки до розділу

1. Обґрунтовано застосування фреймворку Mask R-CNN та розроблено алгоритм виявлення контурів рухомих об'єктів у системах відеоспостереження, який включає чотири кроки і дозволяє прогнозувати маски при візуалізації результатів виявлення об'єктів .

2. За допомогою мови програмування Python, встановлених бібліотек і відповідних залежностей реалізовано модель оптимізованої Mask R-CNN із

застосуванням підходу Transfer Learning, що дало можливість забезпечити точність виявлення контурів об'єктів на рівні не нижче, ніж 92%.

3. Апробовано запропоновані у кваліфікаційній роботі теоретичні результати шляхом їх застосування до відео тривалістю 15 секунд, що дає змогу точно виявляти контури об'єктів, які належать Falcon-9.

РОЗДІЛ 4

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Охорона праці

У кваліфікаційній роботі магістра проводиться дослідження методів і засобів розпізнавання контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження. Оскільки, такі роботи передбачають використання комп'ютерної техніки, зокрема ПК та периферійних пристроїв, то обов'язком виконавця такого процесу є дотримання оптимальних умов праці з охорони праці і техніки безпеки.

Оскільки, в якості виконавця, зазвичай, виступає колектив працівників деякої фірми, то її керівництво повинно забезпечити безпечні умови праці.

Роботодавець зобов'язаний згідно Закону України «Про охорону праці» стаття 13 «Управління охороною праці та обов'язки роботодавця» створити на робочому місці в кожному структурному підрозділі умови праці відповідно до нормативно-правових актів, а також забезпечити додержання вимог законодавства щодо прав працівників у галузі охорони праці.

Із цією метою роботодавець забезпечує функціонування системи управління охороною праці, а саме:

- створює відповідні служби і призначає посадових осіб, які забезпечують вирішення конкретних питань охорони праці, затверджує інструкції про їхні обов'язки, права та відповідальність за виконання покладених на них функцій, а також контролює їх додержання;
- розробляє за участю сторін колективного договору і реалізує комплексні заходи для досягнення встановлених нормативів та підвищення існуючого рівня охорони праці;
- забезпечує виконання необхідних профілактичних заходів відповідно до обставин, що змінюються;

- впроваджує прогресивні технології, досягнення науки і техніки, засоби механізації та автоматизації виробництва, вимоги ергономіки, позитивний досвід з охорони праці тощо;

- забезпечує належне утримання будівель та споруд, виробничого обладнання та устаткування, моніторинг за їх технічним станом;

- забезпечує усунення причин, що призводять до нещасних випадків, професійних захворювань, та здійснення профілактичних заходів, визначених комісіями за підсумками розслідування цих причин;

- організовує проведення аудиту охорони праці, лабораторних досліджень умов праці, оцінку технічного стану виробничого обладнання та устаткування, атестацій робочих місць на відповідність нормативно-правовим актам з охорони праці в порядку і строки, що визначаються законодавством, та за їх підсумками вживає заходів з усунення небезпечних і шкідливих для здоров'я виробничих факторів;

- розробляє і затверджує положення, інструкції, інші акти з охорони праці, що діють у межах підприємства та встановлюють правила виконання робіт і поведінки працівників на території підприємства, у виробничих приміщеннях, на будівельних майданчиках, робочих місцях відповідно до нормативно-правових актів з охорони праці, забезпечує безоплатно працівників нормативно-правовими актами підприємства з охорони праці;

- здійснює контроль за додержанням працівником технологічних процесів, правил поведінки з машинами, механізмами, устаткуванням та іншими засобами виробництва, використанням засобів колективного та індивідуального захисту, виконанням робіт відповідно до вимог з охорони праці;

- організовує пропаганду безпечних методів праці та співробітництво з працівниками у галузі охорони праці.

Роботодавець несе безпосередню відповідальність за порушення нормативно-правових актів з охорони праці.

Для забезпечення оптимальних умов праці працівників при розпізнаванні контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження, необхідно передбачити відповідність мікроклімату у приміщеннях згідно вимог ДСН 3.3.6.042-99.

Категорія робіт при експлуатації засобу розпізнавання контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження належить до легкої – Іб.

Для того щоб визначити, чи відповідає повітряне середовище певного приміщення встановленим нормам, необхідно кількісно оцінити кожний з його параметрів. Оптимальні показники мікроклімату, які необхідно забезпечити у приміщеннях, де експлуатуються ПК у теплу пору року повинні становити: температура – $+22^{\circ}\text{C}$ - $+24^{\circ}\text{C}$, відносна вологість – 40-60%, швидкість руху повітря 0,1 м/с.

Окрім, забезпечення оптимальних показників мікроклімату, необхідно передбачити ще й оптимальні показники шуму та вібрації на робочих місцях.

Граничні величини шуму на робочих місцях регламентуються ДСН 3.3.6.037 – 99 „Державні санітарні норми виробничого шуму, ультразвуку та інфразвуку”. В ньому закладено принцип встановлення певних параметрів шуму, виходячи з класифікації приміщень та їх використання для трудової діяльності.

Окрім цього, на робочих місцях працівників необхідно забезпечити дотримання вимог НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Основними вимогами, визначеними у цьому нормативному документі є:

- площу та об'єм для одного робочого місця оператора визначають згідно з вимогами ДСанПіН 3.3.2-007-98. Площа має бути не менше 6,0 кв.м, об'єм - не менше 20,0 куб.м.

- заземлені конструкції, що знаходяться в приміщеннях, де розміщені робочі місця операторів (батареї опалення, водопровідні труби, кабелі із заземленим відкритим екраном), мають бути надійно захищені діелектричними щитками або сітками з метою недопущення потрапляння працівника під напругу.

– приміщення, де розміщені робочі місця операторів, крім приміщень, у яких розміщені робочі місця операторів великих ЕОМ загального призначення (сервер), повинні бути оснащені системою автоматичної пожежної сигналізації.

При експлуатації програмного комплексу розпізнавання контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження важливим, з точки зору охорони праці, є забезпечення достатньої величини природного та штучного освітлення. Нормованим параметром природного освітлення є коефіцієнт природного освітлення (КПО). КПО встановлюється в залежності від розряду виконуваних зорових робіт. Робота працівників, що експлуатують систему розпізнавання контурів рухомих об'єктів, відноситься до робіт середньої точності, що передбачає IV розряд зорових робіт з мінімальним розміром об'єктів розрізнення 0,5 – 1,0 мм. При цьому в будівлях через віконні отвори передбачається мінімальне бокове освітлення з $K_{\text{ПО}}=1,5 \%$.

Для штучного освітлення нормованим параметром виступає $E_{\text{мін}}$ – мінімальний рівень освітленості, та $K_{\text{п}}$ – коефіцієнт пульсації світлового потоку. Для забезпечення оптимальних умов праці необхідно передбачити коефіцієнт пульсації світлового потоку на рівні не більшому, ніж 20% відповідно до ДБН В.2.5-28 : 2018. Оскільки, робота щодо автоматизованого розпізнавання контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження відноситься до IV розряду зорових робіт, то мінімальний рівень штучного освітлення, який необхідно передбачити складає 300...500 Лк.

Отже, можна зробити висновок, що дослідження методів і засобів розпізнавання контурів рухомих об'єктів в інтелектуалізованих системах відеоспостереження виконувалось при дотриманні достатніх умов освітленості і при правильній організації робочого місця, тому розробка системи є цілком безпечною.

4.2. Фактори виробничого середовища і їх вплив на життєдіяльність промислово-виробничого персоналу

Під терміном «виробниче середовище» слід розуміти комплекс організаційно-технічних параметрів і архітектурно-будівельних рішень, спрямованих на створення комфортних умов праці персоналу і стабільну роботу устаткування при раціональному поєднанні трудових операцій з роботою машин і пристроїв [25]. Виробниче середовище, не здійснює прямого впливу на виробничі процеси, але значною мірою впливає на продуктивність праці персоналу і водночас повинно знімати його стомлюваність та запобігати захворюваності.

Технологічний процес обробки та зберігання інформації вимагає високої культури виробництва, особливої чистоти повітря виробничого середовища. При цьому запиленість і загазованість негативно впливає на точність і надійність роботи електронного устаткування, а також на якість технічних пристроїв.

На підприємствах інформаційної обробки даних потрібний захист від пилоутворення, надійна звукоізоляція між виробничими приміщеннями, а також оптимальні санітарно-гігієнічні умови (мікроклімат, освітлення, опалювання, вентиляція та ін.).

Джерелами шуму на підприємствах такого типу є самі обчислювальні машини (вбудовані в стійки ЕОМ вентилятори, принтери і так далі), центральна система вентиляції і кондиціонування повітря й інше устаткування.

У виробничих приміщеннях, в яких робота на відеодисплейних терміналах (ВДТ) і персональних комп'ютерах (ПК) є допоміжною, рівні шуму на робочих місцях не повинні перевищувати значень, встановлених для цих видів робіт санітарними нормами допустимих рівнів шуму на робочих місцях.

При виконанні основної роботи на ВДТ і ПК (диспетчерські, операторські, розрахункові кабіни і пости управління, зали обчислювальної техніки та ін.) рівень шуму на робочому місці не повинен перевищувати 50 дБА.

У приміщеннях, де працюють інженерно-технічні працівники, що здійснюють лабораторний, аналітичний або вимірювальний контроль, рівень шуму не повинен перевищувати 60 дБА.

У приміщеннях операторів ПК (без дисплеїв) рівень шуму не повинен перевищувати 65 дБА. На робочих місцях в приміщеннях, де розміщені шумні агрегати обчислювальних машин, рівень шуму згідно СанПиН 2.2.2. 542-96 не повинен перевищувати 75 дБА.

Ще однією групою факторів, які негативно можуть впливати на людину є мікрокліматичні параметри. Мікрокліматичні параметри виробничого середовища - це поєднання температури, відносної вологості і швидкості руху повітря.

Ці параметри значною мірою впливають на функціональну діяльність людини, її самопочуття, здоров'я, а також і на надійність роботи обчислювальної техніки. Причому у виробничих умовах характерна сумарна дія мікрокліматичних параметрів.

Великий вплив на мікроклімат в приміщеннях підприємства здійснюють джерела тепла. До них відноситься: обчислювальне устаткування, прилади освітлення, обслуговуючий персонал, а також сонячна радіація. Причому найбільші сумарні тепловиділення серед приміщень підприємства мають зали з великою кількістю ПК, а в них основним тепловиділяючим устаткуванням є ПК, які дають в середньому до 80% сумарних тепловиділень. Від приладів освітлення тепловиділення складають в середньому 12%, від обслуговуючого персоналу - 1%, від сонячної радіації - 6%. Приплив тепла через непрозорі конструкції, що захищають, складає 1%.

На організм людини і роботу устаткування великий вплив чинить відносна вологість повітря. При вологості повітря до 40% стає крихкою основа магнітної стрічки, підвищується знос магнітних голівок, виходить з ладу ізоляція дротів, а також виникає статична напруга при русі носіїв інформації в ПК.

Для зниження концентрації пилу в приміщеннях з електронним устаткуванням потрібно, щоб обслуговуючий персонал працював в халатах і легкому змінному взутті. Запиленість в приміщеннях з ПК не повинна

перевищувати $0,5 \text{ мг/м}^3$. Забороняється палити, оскільки частки попелу, осідаючи на поверхнях магнітних носіїв, викликають збій в роботі ПК.

Устаткування, рівні шуму якого перевищують нормовані, повинно знаходитися поза приміщенням з ВДТ і ПК. Шум в машинних залах знижують, послабляючи шуми самих джерел і спеціальними архітектурно-будівельними рішеннями.

Додатковими заходами щодо шумопогашення в машинних приміщеннях можуть бути:

- пристрій підвісної стелі, яка служить звукопоглинальним екраном;
- використання звукопоглинальних матеріалів з максимальними коефіцієнтами звукопоглинання в області частот 63-8000 Гц для обробки приміщень;
- зменшення площі скляних обгороджувачів і віконних отворів;
- встановлення особливо шумлячих пристроїв на пружні подушки;
- застосування на робочих місцях звукопоглинаючих екранів;
- використання однотонних завіс з щільної тканини, підвішених в складку на відстані 15-20 см від обгороджувача. Ширина завіси має бути в 2 рази більше ширини вікна.

У виробничих приміщеннях, в яких робота на ВДТ і ПК є допоміжною, температура, відносна вологість і швидкість руху повітря на робочих місцях повинні відповідати діючим санітарним нормам мікроклімату виробничих приміщень.

Для підтримки відповідних мікрокліматичних параметрів на підприємствах використовуються системи опалювання і вентиляції, а також проводиться кондиціонування повітря в приміщеннях.

Система опалювання забезпечує достатньо постійне і рівномірне нагрівання повітря в приміщеннях в холодний період року, а також пожежо- і вибухобезпечність. При цьому коливання температури протягом доби не повинні перевищувати $2-3 \text{ С}^\circ$; у горизонтальному напрямі - 2 С° на кожен метр довжини, у вертикальному – 1 С° на кожен метр висоти приміщення.

Систему опалювання розраховують на відшкодування втрат теплоти через конструкції будівлі, на нагрів проникаючого в приміщення холодного повітря і матеріалів, що надходять ззовні, устаткування.

Для забезпечення встановлених норм мікрокліматичних параметрів і чистоти повітря в приміщеннях з ПК застосовують вентиляцію. Проектування системи вентиляції припускає визначення витрати повітря для вентиляції приміщення і охолодження ПК, складання принципової схеми вентиляції приміщення та аеродинамічного розрахунку повітровідводів, вибір повітрязбірних і повітророзподільних пристроїв.

Мінімальна витрата повітря визначається з розрахунку 50-60 м³/год на одного працівника.

Умови по повітрообміну (за 1 годину) наступні:

- двох-трьохкратний - в приміщенні з ПК, в приміщеннях підготовки технічних носіїв, рахункового контролю, в групах прийому і випуску продукції, службах технічного обслуговування устаткування, в архівах;
- п'ятикратний - в приміщеннях копіювання і оформлення документації;
- півторакратний - в інших приміщеннях.

Кондиціонування повітря потрібне в приміщеннях з ПК, приміщеннях для зберігання носіїв інформації, приміщеннях для розміщення сервісної і периферійної апаратури, приміщеннях прийому і видачі інформації.

Висновок.

Отже, на підприємствах інформаційної обробки даних потрібно впроваджувати заходи захисту від таких факторів впливу виробничого середовища, як пил, шум, електромагнітне випромінювання, дотримуватись і забезпечувати відповідність мікроклімату, освітлення та інших показників діючим санітарно-гігієнічним умовам (опалювання, вентиляція та ін.).

ВИСНОВКИ

Основні наукові та практичні результати полягають в наступному.

1. Проведено аналіз особливостей функціонування і сфер застосування систем відеоспостереження, виявлено їхні основні переваги та недоліки, що дало можливість обґрунтувати необхідність впровадження інтелектуальних складових, що дозволить більш якісно проводити моніторинг у конкретній прикладній області та при розв'язанні поставлених задач, шляхом зменшення впливу людського фактору.

2. Проаналізовано характеристики прикладних систем відеоспостереження з інтегрованими інтелектуальними модулями, які забезпечують виконання різних задач, зокрема ідентифікації особи у відеопотоці, і встановлено, що їхня реалізація потребує значних фінансових і програмно-апаратних засобів, а актуальним завданням є оптимізація існуючих моделей і методів розпізнавання контурів рухомих об'єктів (сегментація об'єктів).

3. Проведено аналіз підходів до ідентифікації об'єктів на зображеннях, у результаті якого встановлено, що найбільш ефективними є підходи на основі сегментації, які дають змогу визначати як їх семантично близькі області, так і забезпечувати сегментацію на рівні екземплярів.

4. Запропоновано модель архітектурного рішення щодо реалізації інтелектуалізованої системи відеоспостереження з функцією виявлення контурів рухомих об'єктів, яка використовує апаратний модуль ESP32-CAM та відеокамеру OV2640 і дає можливість з високою точністю на основі підходу сегментації екземплярів накладати та візуалізувати маски на кадрах відеопотоку.

5. Проведено детальний аналіз методів сегментації об'єктів, що дало можливість визначити кращі рішення при імплементації системи відеоспостереження з функцією виявлення контурів, які базуються на основі фреймворків Mask R-CNN та Faster R-CNN та підходу Transfer Learning.

6. Запропоновано рішення щодо оптимізації структури нейронної мережі R-CNN на основі побудови піраміди функцій у мережі з одномасштабного входу, що дає змогу скоротити витрати часу на навчання мережі та є окремою гілкою при виявленні контурів рухомих об'єктів у системах відеоспостереження.

7. Обґрунтовано застосування фреймворку Mask R-CNN та розроблено алгоритм виявлення контурів рухомих об'єктів у системах відеоспостереження, який включає чотири кроки і дозволяє прогнозувати маски при візуалізації результатів виявлення об'єктів .

8. За допомогою мови програмування Python, встановлених бібліотек і відповідних залежностей реалізовано модель оптимізованої Mask R-CNN із застосуванням підходу Transfer Learning, що дало можливість забезпечити точність виявлення контурів об'єктів на рівні не нижче, ніж 92%.

9. Апробовано запропоновані у кваліфікаційній роботі теоретичні результати шляхом їх застосування до відео тривалістю 15 секунд, що дає змогу точно виявляти контури об'єктів, які належать Falcon-9.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Барабаш Ю. Л. Коллективные статистические решения при распознавании. М. : Радио и связь, 1983. 224 с.
2. Васильев В. И. Распознающие системы : справочник. К. : Наукова думка, 1983. 230 с.
3. Горелик А. Л. Методы распознавания. М. : Высшая школа, 1984. 219 с.
4. Дуда Р. Распознавание образов и анализ сцен : пер. с англ. М. : Мир, 1978. 510 с.
5. Форсайт Д. А. Компьютерное зрение. Современный подход : пер. с англ. М. : Вильямс, 2004. 928 с.
6. Шапиро Л. Компьютерное зрение : пер. с англ. БИНОМ. Лаборатория знаний, 2006. 752 с.
7. Яцишин В.В., Щербаков О.О., Лова М.Р. Аналіз баз даних зображень у галузі комп'ютерного зору. Матеріали Х міжнародної науково - технічної конференції молодих учених і студентів «Актуальні задачі сучасних технологій» (24-25 листопада 2021 р.) Тернопільського національного технічного університету імені Івана Пулюя. Тернопіль: ТНТУ. 2021. С. 144.
8. Жаровський Р.О., Лова М.Р., Щербаков О.О. Застосування індексу структурної подібності зображень при їх аналізі. Матеріали ІХ науково-технічної конференції Тернопільського національного технічного університету імені Івана Пулюя «Інформаційні моделі, системи та технології» (8-9 грудня 2021 року). Тернопіль: ТНТУ. 2021. С. 114.
9. Кузин Л.Т. Расчет и проектирование дискретных систем управления.-М.: ГН ТИМЛ, 2012.- 648 с.
10. Python Tutorial. URL: <https://www.w3schools.com/python/default.asp> (дата звернення 15.10.2021 р.).
11. Pandas documentation. URL: <https://pandas.pydata.org/docs/index.html> (дата звернення 28.10.2021 р.).

12. RASPBERRY PI 3 MODEL B+. URL: <https://www.digikey.com/en/products/detail/raspberry-pi/RASPBERRY-PI-3-MODEL-B/8571724> (дата звернення 05.11.2021 р).

13. Install TensorFlow. URL: <https://www.tensorflow.org/install> (дата звернення 17.04.2021 р.).

14. TensorFlow Datasets: a collection of ready-to-use datasets. URL: <https://www.tensorflow.org/datasets> (дата звернення 17.11.2021 р.)

15. Cuda. URL: <https://opencv.org/platforms/cuda/> (дата звернення 21.11.2021 р.)

16. Петин В. Микрокомпьютеры Raspberry Pi: Практическое руководство. БХВ-Петербург. 2015. 240 с.

17. Макаров С. Arduino Uno и Raspberry Pi 3: от схемотехники к интернету вещей. Litres. 2019 р. 202 с.

18. Тиммонс-Браун М. Робототехника на Raspberry Pi для юных конструкторов и программистов Робототехника на Raspberry Pi для юных конструкторов и программистов. БХВ-Петербург. 2020. 208 с.

19. ESP32-CAM Video Streaming Web Server (works with Home Assistant). URL: <https://randomnerdtutorials.com/esp32-cam-video-streaming-web-server-camera-home-assistant/> (дата звернення 16.04.2021 р.)

20. Лукьяница А. А. Цифровая обработка видеоизображений. М. : Ай-Эс-Эс Пресс, 2009. 518 с.

21. Офіційна сторінка Open CV. URL: <https://opencv.org/about/> (дата звернення 10.05.2021).

22. ESP8266:Примеры/Веб-сервер с помощью ESP8266 и IDE Arduino. URL: http://wikihandbk.com/wiki/ESP8266:%D0%9F%D1%80%D0%B8%D0%BC%D0%B5%D1%80%D1%8B/%D0%92%D0%B5%D0%B1-%D1%81%D0%B5%D1%80%D0%B2%D0%B5%D1%80_%D1%81_%D0%BF%D0%BE%D0%BC%D0%BE%D1%89%D1%8C%D1%8E_ESP8266_%D0%B8_IDE_Arduino (дата звернення 28.04.2021 р.)

23. Python Tutorial. URL: <https://www.w3schools.com/python/default.asp> (дата звернення 18.05.2021 р.).

24. Pandas documentation. URL: <https://pandas.pydata.org/docs/index.html> (дата звернення 26.05.2021 р.).

25. Микитишин А.Г., Митник М.М., Стухляк П.Д., Пасічник В.В. Комп'ютерні мережі. Книга 1. Львів, «Магнолія 2006». 2013. 256 с.

26. Микитишин А.Г., Митник М.М., Стухляк П.Д., Пасічник В.В. Комп'ютерні мережі. Книга 2. Львів, «Магнолія 2006», 2014. 312 с.

27. Микитишин А.Г., Митник М.М., Стухляк П.Д. Телекомунікаційні системи та мережі. Тернопіль: Вид-во ТНТУ імені Івана Пулюя, 2016. 384 с.

28. Микитишин А.Г., Митник М.М., Стухляк П.Д. Комплексна безпека інформаційних мережевих систем: навчальний посібник. Тернопіль: Вид-во ТНТУ імені Івана Пулюя, 2016. 256 с.

29. НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Київ. 2018.

30. Катренко Л.А., Катренко А.В. Охорона праці в галузі комп'ютерингу. Львів: Магнолія-2006. 2012. 544 с.

31. Бедрій Я. Основи охорони праці користувачів персональних комп'ютерів: навчальний посібник для студентів ВНЗ та інженерів-практиків. Навчальна книга-Богдан. 2014. 144 с.

Додаток А
Тези конференцій

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Тернопільський національний технічний університет імені Івана Пулюя (Україна)
Університет імені П'єра і Марії Кюрі (Франція)
Маріборський університет (Словенія)
Технічний університет у Кошице (Словаччина)
Вільнюський технічний університет ім. Гедимінаса (Литва)
Білоруський національний технічний університет (Республіка Білорусь)
Міжнародний університет цивільної авіації (Марокко)
Наукове товариство ім. Т.Шевченка

АКТУАЛЬНІ ЗАДАЧІ СУЧАСНИХ ТЕХНОЛОГІЙ

Збірник тез доповідей Том I

**X Міжнародної науково-практичної
конференції молодих учених та студентів
24-25 листопада 2021 року**



**УКРАЇНА
ТЕРНОПІЛЬ – 2021**

<i>Матеріали X Міжнародної науково-практичної конференції молодих учених та студентів</i>		
<i>«АКТУАЛЬНІ ЗАДАЧІ СУЧАСНИХ ТЕХНОЛОГІЙ» – Тернопіль 24-25 листопада 2021 року</i>		
32.	С.В. Тиш, В.В.Б. Кохан ФОРМУВАННЯ СУСПІЛЬНОЇ ДУМКИ В СОЦІАЛЬНИХ МЕРЕЖ НА ПРИКЛАДІ МЕРЕЖІ TWITTER	127
33.	Р. Трач, Ю. Баляс, Р. Трембач ВДОСКОНАЛЕННЯ СИСТЕМИ ВІБРОКОНТРОЛЮ МЛИНА	129
34.	Г.І.Франчевська ПРОБЛЕМИ ТА ПЕРСПЕКТИВИ РОЗВИТКУ МЕТОДІВ ВИЯВЛЕННЯ СИГНАЛІВ ПЛОДУ НА ФОНІ МАТЕРІ ТА ШУМУ	131
35.	Г.П.Химич, В.В.Демчук ДОСЛІДЖЕННЯ УМОВ РОЗПОВСЮДЖЕННЯ НАЗЕМНОГО ТА СУПУТНИКОВОГО ЗВ'ЯЗКУ ЗА ТЕХНОЛОГІЄЮ 5G	133
36.	Г.П.Химич, І.Є.Яцюк ВПРОВАДЖЕННЯ РОЗУМНИХ ТЕХНОЛОГІЙ ІЗ ШТУЧНИМ ІНТЕЛЕКТОМ ДЛЯ КЕРУВАННЯ АВТОМОБІЛЬНИМ ТА ПІШОХІДНИМ РУХОМ НА ВУЛ. РУСЬКА МІСТА ТЕРНОПОЛЯ	135
37.	О. К. Шкодзінський, М. М. Луцків, І-М. С. Смолій РОЗВИТОК ЗАСОБІВ ВЕРИФІКАЦІЇ ОСОБИ ТА ЇЇ ДІЙ ПРИ КОНТРОЛІ ЗНАНЬ В УМОВАХ ДИСТАНЦІЙНОГО НАВЧАННЯ	138
38.	М.І. Шоцький, В.В. Федина, С.В. Марценко ДОСЛІДЖЕННЯ ПРОЦЕСІВ АВТОМАТИЗАЦІЇ КЕРУВАННЯ МЕРЕЖЕВИМИ ПРИСТРОЯМИ	140
39.	М.І. Шоцький, В.В. Федина ДОСЛІДЖЕННЯ ПРОЦЕСУ ОРГАНІЗАЦІЇ ЗОНОВОЇ БЕЗПЕКИ У КОМП'ЮТЕРНІЙ МЕРЕЖІ	141
40.	А. В. Юхименко, О. В. Чебанюк МЕТОДИКА ПОПЕРЕДЖЕННЯ ВИТОКУ МОВНОЇ ІНФОРМАЦІЇ ЧЕРЕЗ ГІРОСКОП У МОБІЛЬНИХ ПРИСТРОЯХ НА ОС ANDROID	142
41.	В.В. Яцишин, О.О.Щербаков, М.Р.Лова АНАЛІЗ БАЗ ДАНИХ ЗОБРАЖЕНЬ У ГАЛУЗІ КОМП'ЮТЕРНОГО ЗОРУ	144
42.	В.В.Яцишин, В.В.Шуптарський, Д.А.Цісарук АЛГОРИТМИ МАШИННОГО НАВЧАННЯ ДЛЯ СЕГМЕНТАЦІЇ КОРИСТУВАЧІВ У МАРКЕТИНГОВИХ КОМП'ЮТЕРНИХ СИСТЕМ	145
43.	В.В. Яцишин, Х.В. Яворська АНАЛІЗ ОСОБЛИВОСТЕЙ ВІЗУАЛЬНИХ МОВ ПРОГРАМУВАННЯ	146

УДК 004.89

Яцишин В.В. канд. техн. наук, доцент, Щербаков О.О., Лова М.Р.

Тернопільський національний технічний університет імені Івана Пулюя

АНАЛІЗ БАЗ ДАНИХ ЗОБРАЖЕНЬ У ГАЛУЗІ КОМП'ЮТЕРНОГО ЗОРУ**Yatsyshyn V.V. PhD, Assoc. Prof., Scherbakov O.O., Lova M.R.****ANALYSIS OF IMAGE DATA SETS IN THE AREA OF COMPUTER VISION**

Комп'ютерний зір відноситься до області штучного інтелекту, який швидко розвивається. Для того, щоб полегшити роботу при проектуванні таких систем, зокрема, моделей та їхніх архітектур, деякі фірми поділились сформованими навчальними вибірками.

Одним з таких наборів даних є модифікована база даних рукописних цифр Національного інституту стандартів і технологій (MNIST). Цей набір представляє собою елементарну колекцію даних для комп'ютерного зору, що містить 70 тисяч зображень рукописних цифр для кожного числа, тобто 0 – 9, які відформатовані у форматі сірого 28x28. Дані попередньо розділені на навчальну вибірку (60 тисяч) і тестовий набір (10 тисяч). Усі цифри розміщені у центрі зображення. Даний набір є корисним для фундаментального проекту комп'ютерного зору, коли необхідно автоматично оцифрувати рукописний текст

Набори даних CIFAR-10 і CIFAR-100 підготовлені Канадським інститутом перспективних досліджень. CIFAR-10 складається з 60 тисяч зображень розподілених за 10-ма класами. Дана колекція включає зображення літаків, автомобілів, птахів, котів, оленів, собак, жаб, коней, кораблів і вантажівок. До складу CIFAR-100 входять 60 тисяч зображень, але тепер вже передбачено 100 класів, тобто 600 зображень на кожен клас. Колекції CIFAR-10 і CIFAR-100 є зручними для використання, оскільки всі вони відформатовані у форматі 32x32 пікселі та попередньо розділені на навчальний набір із 50 тисяч зображень та тестовий набір із 10 тисяч зображень з рівною часткою даних з усіх класів.

Набір даних IMDB-Wiki містить 520 тисяч зображень обличчя, виділених з IMBD та Вікіпедії. Дані містять важливу метадані, наприклад розташування обличчя на зображенні, ім'я, дата народження та стать людини на фотографіях. Цей набір даних зазвичай використовується для завдань прогнозування статі та віку.

ImageNet набір даних створено спільно Стенфордським університетом та Принстонським університетом для типового змагання з комп'ютерного зору під назвою ImageNet Large Scale Visual Recognition Challenge (ILSVRC), де командам-учасникам було запропоновано 5 основних завдань, тобто класифікація об'єктів, локалізація об'єктів, виявлення об'єктів, виявлення та розпізнавання об'єктів з відео за допомогою набору даних ImageNet. Цей набір даних побудований на основі ієрархії WordNet (лексична база даних для англійської мови), де містяться лише іменники. В середньому на кожен вузол ієрархії припадає понад 500 зображень. Всього є понад 1,4 мільйона зображень понад 220 тисяч класів. Наразі це найбільший доступний набір зображень, який є відкритим і загальнодоступним.

Набір PASCAL VOC даних було відкрито дослідницьким інститутом PASCAL, що фінансується Європейським Союзом. Дана колекція містить зображення за 4-ма основними темами і включає транспортні засоби, домашнє господарство, тварин та людей.

При реалізації комп'ютеризованої системи виявлення справжності ID-картки доцільним є використання наведених наборів даних, що дозволить забезпечити виявлення об'єктів у різних предметних областях.

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ**

МАТЕРІАЛИ

ІХ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

**«ІНФОРМАЦІЙНІ МОДЕЛІ,
СИСТЕМИ ТА ТЕХНОЛОГІЇ»**



8–9 грудня 2021 року

**ТЕРНОПІЛЬ
2021**

- Ю.З. Лещини, З.В. Кузик**
 МЕТОДИ СТВОРЕННЯ МАКРОСІВ ДЛЯ АВТОМАТИЗОВАНОЇ РОЗРОБКИ
 ТЕХНІЧНОЇ ДОКУМЕНТАЦІЇ МЕРЕЖЕВИХ КАБЕЛЬНИХ СИСТЕМ
Yu. Leshchyshyn, Z. Kuzik
 METHODS OF MACROS DESIGN FOR AUTOMATED DEVELOPMENT OF
 NETWORK CABLE SYSTEM TECHNICAL DOCUMENTATIONS 113
- Р.О. Жаровський, М.Р. Лова, О.О. Щербак**
 ЗАСТОСУВАННЯ ІНДЕКСУ СТРУКТУРНОЇ ПОДІБНОСТІ ЗОБРАЖЕНЬ
 ПРИ ЇХ АНАЛІЗІ
R.O. Zharovskyy, M.R. Lova, O.O. Scherbakov
 APPLICATION OF THE STRUCTURAL SIMILARITY INDEX MEASURE IN
 THE IMAGES ANALYSIS 114
- Ю.З. Лещини, О.О. Марушчак**
 МЕТОДИ ТА ЗАСОБИ ПОБУДОВИ КОМП'ЮТЕРНОЇ СИСТЕМИ
 ОЦІНЮВАННЯ ХАРАКТЕРИСТИК ФОНОКАРДІОСИГНАЛІВ
Yu. Leshchyshyn, O. Marushchak
 METHODS AND MEANS OF THE DEVELOPMENT OF A
 PHONOCARDIOGRAPHIC SIGNALS CHARACTERISTICS EVALUATION
 COMPUTER SYSTEM 115
- Р.В. Ларіоник, Н.С. Луцик, А.М. Паламар**
 СИСТЕМА ДЛЯ МОНІТОРИНГУ ЯКОСТІ АТМОСФЕРНОГО ПОВІТРЯ НА
 БАЗІ IOT
R.V. Larionyk, N.S. Lutsyk, A.M. Palamar
 IOT-BASED AIR QUALITY MONITORING SYSTEM 116
- А.І. Маційовський**
 ДОСЛІДЖЕННЯ ВИСОКОНАВАНТАЖЕНИХ МЕРЕЖ ПЕРЕДАЧІ ДАНИХ
A.I. Matsiyovskiy
 RESEARCH OF HIGHLY LOADED DATA TRANSMISSION NETWORKS 117
- М.В. Оконський, С.А. Лупенко, А.М. Паламар**
 ІНФОРМАЦІЙНО-ВИМІРЮВАЛЬНА СИСТЕМА ДЛЯ КОНТРОЛЮ
 МЕТЕОРОЛОГІЧНИХ ПАРАМЕТРІВ НА ОСНОВІ ІНТЕРНЕТУ РЕЧЕЙ
M.V. Okonskyi, S.A. Lupenko, A.M. Palamar
 INFORMATION AND MEASURING SYSTEM FOR CONTROL OF
 METEOROLOGICAL PARAMETERS BASED ON THE INTERNET OF
 THINGS 118
- О.В. Осійчук, Є.В. Тиш, канд. техн. наук**
 АНАЛІЗ ПОПУЛЯРНОСТІ ВИКОРИСТАННЯ ОДНОПЛАТНИХ
 КОМП'ЮТЕРІВ
O.V. Oseechuk, Ye.V. Tysh, Ph.D
 ANALYSIS OF THE POPULARITY OF USING SINGLE-PAID
 COMPUTERS 119
- Х. Ольховецька**
 КОМП'ЮТЕРИЗОВАНА СИСТЕМА КОНТРОЛЮ ЯКОСТІ ПРОЦЕСУ
 ФЕРМЕНТАЦІЇ ВИННИХ ПРОДУКТІВ
Kh. Olkhovetska
 COMPUTERIZED QUALITY CONTROL SYSTEM OF WINE PRODUCTION
 FERMENTATION PROCESS 120

УДК 004.89

Р.О. Жаровський канд. техн. наук, М.Р. Лова, О.О. Щербаков

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ЗАСТОСУВАННЯ ІНДЕКСУ СТРУКТУРНОЇ ПОДІБНОСТІ ЗОБРАЖЕНЬ ПРИ ЇХ АНАЛІЗІ

UDC 004.89

R.O. Zharovskyy PhD, M.R. Lova, O.O. Scherbakov**APPLICATION OF THE STRUCTURAL SIMILARITY INDEX MEASURE IN THE IMAGES ANALYSIS**

Структурний індекс подібності зображень використовується як метрика для вимірювання міри схожості між двома заданими зображеннями. Оскільки ця технологія існує з 2004 року, то наявними у всесвітній мережі є доволі багато наукових і прикладних публікацій, що пояснюють дану концепцію і теорію SSIM («Structural Similarity Index Model») на загальному рівні.

Система візуального сприйняття людини здатна ідентифікувати структурну інформацію зі сцени і, отже, визначити відмінності між інформацією, добутою з сталонної та сцени зразку.

Отже, метрика, яка повторює таку поведінку, буде краще виконувати завдання, які передбачають розрізнення зразка та сталонного зображення. Показник структурної схожості виділяє з зображення 3 ключові характеристики: яскравість, контраст, структура.

Порівняння двох зображень здійснюється на основі цих 3 ознак. На рис. 1, наведеному нижче, показано розташування та алгоритм вимірювання структурної подібності. Сигнали X і Y надсилають сталонні зображення та зображення, які потрібно перевірити.



Рисуюнок 1. Система вимірювання показника структурної подібності зображень

Ця система обчислює індекс структурної схожості між двома заданими зображеннями, значення якого належать інтервалу від -1 до +1. Значення «+1» вказує, що дані зображення дуже схожі або однакові, а значення «-1» – вказує, що вони дуже різні. Часто ці значення представляються в діапазоні [0, 1], де крайні значення мають ту саму інтерпретацію.