

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

Охорона (назва факультету)

Кафедра комп'ютерних систем та мереж

(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

Магістр

(назва освітнього ступеня)

на тему:

*Методи та засоби розпізнавання статі та віку людини
у відеопотоці комп'ютеризованою системою*

Виконав: студент 6 курсу групи СІм-61
спеціальності

123 «Комп'ютерна інженерія»

(шифр і назва спеціальності (напряму підготовки))

Мишаківський Я. О.

(підпис)

(прізвище та ініціали)

Керівник

Баран І.О.

(підпис)

(прізвище та ініціали)

Нормоконтроль

Луцик Н.С.

(підпис)

(прізвище та ініціали)

Завідувач

кафедри

Осухівська Г.М.

(підпис)

(прізвище та ініціали)

Рецензент

(підпис)

(прізвище та ініціали)

м. Тернопіль – 2024

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

Факультет комп'ютерно-інформаційних систем і програмної інженерії

(повна назва факультету)

Кафедра комп'ютерних систем та мереж

(повна назва кафедри)

ЗАТВЕРДЖУЮ

Завідувач кафедри

доц. Осухівська Г.М.

(підпис) (прізвище та ініціали)

«11» листопада 2024 р

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

на здобуття освітнього ступеня

магістр

(назва освітнього ступеня)

за спеціальністю

123 Комп'ютерна інженерія

студенту

Мишаківському Ярославу Олександровичу

(прізвище, ім'я, по батькові)

1. Тема роботи

Методи та засоби розпізнавання статі та віку людини
у відеопотоці комп'ютеризованою системою

Керівник роботи

Баран Ігор Олегович., к.т.н., доц.

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом по університету від «29» листопада 2024 року № 4/7-1144

2. Термін подання студентом роботи 16.12.2024

3. Вихідні дані до роботи

наукові літературні джерела

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)

1 Аналіз предметної області дослідження. 2 Теоретична частина.

3. Практична частина

4 Охорона праці та безпека в надзвичайних ситуаціях

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

1. Тема, мета, задачі, об'єкт, предмет, новизна дослідження. 2. Актуальність дослідження.

3. Послідовність етапів визначення статі та віку

4. Блок-схема роботи розроблених алгоритмів розпізнавання статі та віку.

5. Блок-схема роботи розроблених алгоритмів створення індивідуальних відбитків за графічними ознаками для повторного розпізнавання

6. Діаграма класів функціонального блоку отримання ознак для розпізнавання статі та віку..

7. Приклад вихідного зображення та розмітки сегментів тіла людини

8. Результати експериментів в таблицях. 9. Залежності значення метрики СМС-1 від розміру треклета. 10. Висновки. Результати проведеного дослідження..

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
<i>Охорона праці</i>	<i>Осухівська Г.М., доцент</i>		
<i>Безпека в НС</i>			

7. Дата видачі завдання 11.11.2024

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Термін виконання етапів роботи	Примітка
1.	<i>Аналіз існуючих методів та засобів розв'язання науково-технічних проблем, що відповідають тематиці кваліфікаційної роботи.</i>	<i>20.11.2024р. – 30.11.2024р.</i>	<i>Виконано</i>
2.	<i>Обґрунтування актуальності дослідження</i>	<i>05.12-08.12.24</i>	<i>Виконано</i>
3.	<i>Аналіз предмету дослідження та предметної області</i>	<i>09.12-11.12.24</i>	<i>Виконано</i>
4.	<i>Проведення дослідження методів та засобів аналітичного опрацювання даних</i>	<i>11.12-12.12.24</i>	<i>Виконано</i>
5.	<i>Оформлення розділу «Аналіз предметної області дослідження»</i>	<i>12.12-13.12.24</i>	<i>Виконано</i>
6.	<i>Оформлення розділу «Теоретична частина»</i>	<i>13.12-14.12.24</i>	<i>Виконано</i>
7.	<i>Оформлення розділу «Практична частина»</i>	<i>14.12-15.12.24</i>	<i>Виконано</i>
8.	<i>Оформлення розділу «Охорона праці та безпека в надзвичайних ситуаціях»</i>	<i>07.12-12.12.24</i>	<i>Виконано</i>
9.	<i>Нормоконтроль</i>	<i>11.12-14.12.24</i>	
10.	<i>Попередній захист роботи</i>	<i>20.12.24</i>	<i>Виконано</i>
11.	<i>Захист кваліфікаційної роботи</i>	<i>23.12.24</i>	

Студент

(підпис)

Мишаківський Я.О.

(прізвище та ініціали)

Керівник роботи

(підпис)

Баран І.О.

(прізвище та ініціали)

АНОТАЦІЯ

Мишаківський Я.О. Методи та засоби розпізнавання статі та віку людини у відеопотоці комп'ютеризованою системою: робота на здобуття кваліфікаційного ступеня магістра: спец. 123 — комп'ютерна інженерія / наук.кер. І. О. Баран. Тернопіль: Тернопільський національний технічний університет імені Івана Пулюя, 2024. — 78 с.

Ключові слова: відеоаналітика, глибоке навчання, дескриптор, згорткова нейронна мережа, класифікація, розпізнавання, треклет

Кваліфікаційна робота присвячена дослідженню математичних і програмних основ розпізнавання людей за статтю та віком у відеопотоці даних для комп'ютерних систем. Як дескриптор ходи особи використовується енергетичне зображення ходи.

Розроблено та представлено власні варіанти алгоритмів розпізнавання статі та віку осіб на кадрах у відеопотоці із застосуванням ознак силуету та руху та їх повторної ідентифікації. Проведені експериментальні дослідження дозволили як виконати відбір найкращих методів і їх гіперпараметрів для функціонування запропонованих алгоритмів, так і покращити загальну якість процесу відеоаналітики.

Розроблені алгоритми здатні успішно працювати на даних відеозаписів у загальних міських умовах, а також бути впроваджені в ядро платформи, що забезпечує детектування та трекінг об'єктів на відеозаписах та відеопотоках, знятих за допомогою камери статичного відеоспостереження.

ANNOTATION

Myshakivskyi Ya. Methods and means of recognizing a person's gender and age in a video stream by a computerized system. Master's Graduation Thesis: speciality 123 — Computer engineering / supervisor I.O. Baran. Ternopil: Ternopil Ivan Puluj National Technical University, 2024. — 78 p.

Keywords: video analytics, deep learning, descriptors, convulsive neural networks, classification, recognition, tracklet

This thesis deals with the study of the mathematical and software foundations of recognizing people by gender and age in a video data stream for computer systems. The energy image of the gait is used as a descriptor of a person's gait.

Own variants of algorithms for recognizing the gender and age of people in frames in a video stream using silhouette and motion features and their re-identification have been developed and presented. The experimental studies conducted have allowed both to select the best methods and their hyperparameters for the functioning of the proposed algorithms and to improve the overall quality of the video analytics process.

The developed algorithms are able to successfully work on video data in general urban conditions, as well as to be implemented in the core of the platform that provides detection and tracking of objects in video recordings and video streams recorded using a static video surveillance camera.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ СКОРОЧЕНЬ І ТЕРМІНІВ	8
ВСТУП.....	9
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ДОСЛІДЖЕННЯ	11
1.1. Методи розпізнавання статі та віку за зображенням особи	11
1.2. Методи розпізнавання статі та віку за ознаками силуету та руху	14
1.2.1. Обрізання та масштабування обмежувальних вікон	16
1.2.2. Сегментація об'єктів та бінаризація зображень	16
1.2.3. Класифікація за кутами спостереження.....	18
1.2.4. Конструювання GEI	19
1.2.5. Визначення статі та віку	21
1.3. Висновки до розділу	24
РОЗДІЛ 2. ТЕОРЕТИЧНА ЧАСТИНА	25
2.1. Теоретичні дослідження методів повторної ідентифікації	25
2.1.1. Постановка задачі.....	25
2.1.2. Обмеження розв'язання задачі	26
2.1.3. Етапи роботи алгоритмів повторного розпізнавання	27
2.1.4. Отримання графічних ознак.....	27
2.1.5. Розрахунок дескриптора схожості.....	30
2.1.6. Ранжування	32
2.2. Розробка алгоритму розпізнавання статі та віку	34
2.2.1. Вибір методів реалізації алгоритму	34
2.2.2. Концепція алгоритму	35
2.3. Розробка алгоритму повторної ідентифікації	38
2.3.1. Вибір методів реалізації алгоритму	38
2.3.2. Концепція алгоритму	40
2.4. Висновки до розділу	45
РОЗДІЛ 3. ПРАКТИЧНА ЧАСТИНА.	46

3.1. Інструменти розробки	46
3.2. Реалізація програмного модуля розпізнавання статі та віку	47
3.3. Розробка та розмітка навчальних вибірок	50
3.3.1. Дані для завдання сегментації об'єктів	50
3.3.2. Дані для задачі визначення статі та віку	51
3.3.3. Дані для завдання створення індивідуальних відбитків об'єктів за графічними ознаками	53
3.4. Експериментальні випробування	54
3.4.1. Умови проведення випробувань	54
3.4.2. Випробування моделей сегментації для розпізнавання статі та віку	54
3.4.3. Випробування алгоритмів розпізнавання статі та віку	56
3.4.4. Випробування алгоритмів повторної ідентифікації	59
3.5. Висновки до розділу	62
РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	63
4.1. Охорона праці	63
4.2. Вплив електромагнітного імпульсу ядерного вибуху на елементи виробництва та заходи захисту	66
4.3. Висновки до розділу	69
ВИСНОВКИ.....	70
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	71
ДОДАТОК А. Тези конференції	

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ СКОРОЧЕНЬ І ТЕРМІНІВ

BGEI (Back-filled Gait Energy Image) – енергетичне зображення ходи зі зворотним зафарбовуванням.

CMC (Cumulative Match Characteristic) – графічне представлення, яке використовується для оцінки продуктивності системи пошуку зображень на основі вмісту, оціночна метрика для методів повторної ідентифікації особи.

CNN (Convolutional Neural Networks) – згорткова нейронна мережа.

GEI (Gait Energy Image) – енергетичне зображення ходи – монохромне зображення, що агрегує в собі безліч представлень об'єкта, отриманих у час його руху (ходи) – дескриптор ходи об'єкта.

GEV (Gait Energy Volume) – енергетичне об'ємне зображення ходи.

SLP (Silhouette Longitudinal Projection) – поздовжні проекції силуету.

STP (Silhouette Traverse Projection) – поперечні проекції силуету.

SOTA (State of the art) – відповідний сучасному рівню розвитку – умовна назва моделей та/або алгоритмів, що вирішують певне завдання з найкращими на даний момент характеристиками.

Аугментація (augmentation) - створення штучного прикладу даних шляхом внесення змін до реального прикладу.

Дескриптор – загалом це деякий набір ознак, що описують певний об'єкт або величину для її подальшого розпізнавання та/або ідентифікації.

Треклет (tracklet) – деяка обмежена частина траєкторії.

Трекінг (tracking) – процес відстеження об'єкта на відеопослідовності, визначення траєкторії об'єкта.

ВСТУП

Актуальність теми. Зараз дедалі більшою популярністю користуються інтелектуальні системи відеоспостереження, здатні автоматизувати виконання завдань, вирішуваних зазвичай людиною. Такі системи засновані на використанні алгоритмів комп'ютерного зору та методах глибокого навчання і використовуються в багатьох прикладних областях: робототехніка, підрахунок пішохідного та автомобільного трафіку, розумні будинки, охорона безпеки життєдіяльності та інших.

Найчастіше інтелектуальні системи вирішують завдання детектування та трекінгу об'єктів на кадрах відеопотоку. На даних, отриманих в результаті вирішення цих завдань, проводиться інтелектуальний опис об'єктів та подій на зображеннях відеопослідовності.

Найбільш корисними завданнями інтелектуальної відеоаналітики є визначення статі та віку людей, а також створення анонімних індивідуальних графічних відбитків об'єктів задля забезпечення можливості повторної ідентифікації. Однак більшість рішень, що демонструють хорошу якість відеоаналітики, використовують як ознаки зображення осіб людини. Такі рішення не здатні працювати з даними відеоспостереження із довільними зовнішніми умовами з часто виникаючих труднощів виділення особи на кадрах.

Мета дослідження: реалізація методів класифікації людей за статтю та віком та їх повторної ідентифікації у відеопотоці за допомогою технологій глибокого навчання.

Завдання кваліфікаційної роботи:

- провести аналіз існуючих алгоритмів інтелектуального опису об'єктів на кадрах відеопотоку;
- розробити власні алгоритми інтелектуального опису об'єктів, що працюють на даних, типових для міського середовища;
- провести дослідження та налаштування алгоритмів за допомогою розроблених програмних засобів;

– виконати впровадження та тестування розроблених алгоритмів до складу ядра платформи.

Об'єкт дослідження: розпізнавання статі та віку за зображенням особи у відеопотоці.

Предмет дослідження: методи та алгоритми розпізнавання статі та віку за зображенням особи у відеопотоці.

Методи дослідження: фундаментальні твердження комп'ютерної інженерії, методи - аналітичний, порівняльний, експериментальний.

Наукова новизна отриманих результатів:

– запропонована власна архітектура алгоритмів на базі графічних ознак на виході нейромережевого детектора.

Практичне значення одержаних результатів. Розширення функціональності рішення для автоматизації підрахунку пішохідного трафіку в частині інтелектуального опису об'єктів. Підрахований системою трафік використовується для задач оцінки пасажиропотоку та вибору локацій для відкриття нових торгових точок на основі порівняльного аналізу.

Публікації. Окремі результати дослідження апробовано на XII науково-технічній конференції «Інформаційні моделі, системи та технології» як опубліковані тези:

1. Мишаківський Я.О. Методи для реалізації алгоритму розпізнавання статі та віку людини у відеопотоці даних // Інформаційні моделі, системи та технології: Праці XII наук.-техн. конф. (Тернопіль, 18-19 грудня 2024 р.). с. 138.

2. Мишаківський Я.О. Методи розпізнавання статі та віку за зображенням особи // Інформаційні моделі, системи та технології: Праці XII наук.-техн. конф. (Тернопіль, 18-19 грудня 2024 р.). с. 140.

Структура роботи. Робота складається з пояснювальної записки та графічної частини [1]. Пояснювальна записка складається з вступу, 4 розділів, висновків, списку використаної літератури та додатків. Обсяг роботи: пояснювальна записка – 81 арк. формату А4, графічна частина – 10 аркушів формату А1.

РОЗДІЛ 1

АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1. Методи розпізнавання статі та віку за зображенням особи

Розпізнавання по обличчю є популярним і точним підходом до завдання визначення статі та віку по фото або відео. Причиною цього є присутність на обличчі всіх візуальних особливостей, які дозволяють визначати стать, вік, расу, настрій та багато інших рис людини, візуально видимі в основному (або виключно) на вигляд і подання особи. На відео розпізнавання по особі переважно для визначення статі та віку у випадках, коли на кадрах особа людини видно чітко.

У послідовності дій для вирішення описаної задачі можна виділити три етапи (рис. 1.1):

- визначення людей на кадрах відеопотоку та побудова детекцій;
- детектування осіб;
- розпізнавання статі та віку по особі.

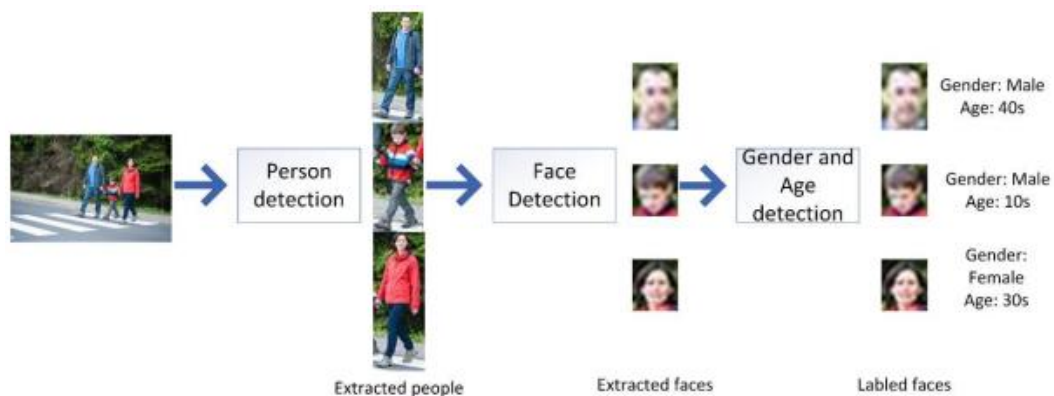


Рис. 1.1. Послідовність етапів визначення статі та віку

Виявлення особи (face detection) є досить відомим завданням, у якого є множина різних підходів [4]. Для роботи системі необхідно обмежувати область пошуку особи. Це необхідно для можливості однозначного зіставлення виявленої особи з людиною. Так само це рішення дозволить прискорити розробку та загальну швидкість роботи системи, оскільки відсутня необхідність у написанні та

тестуванні додаткового модуля в порівнянні обмежуючих вікон детектованих осіб та обмежуючих вікон пішоходів. Нижче розписані підходи до задачі детектування особи на кадрі.

Класичні методи визначення осіб. Підходи, які широко використовувалися на початку 2000х років. Використовувалися, тому числі, класичні статистичні методи, аналіз головних компонентів, лінійні та нелінійні класифікатори. Також використовувалися методи зображення, що відображають на площину і за допомогою знаходження відстані Махаланобіса, знаходили ступінь близькості поточного виявленого об'єкта з референс вектор обличчя. Незважаючи на хорошу теоретичну базу для відповідних методів, вони майже не застосовувалися в реальних промислових застосунках. Проблемою була вкрай низька здатність подібних підходів адаптуватись до інваріантності середовища.

Методи, засновані на застосуванні нейронних мереж. Deep CNN добре себе зарекомендували в завданні детектування об'єктів на зображеннях. Вони здатні визначати характерні риси об'єкта безпосередньо з пікселів зображення [5]. До найвідоміших архітектур CNN можна віднести AlexNet, GoogLeNet, VGGNet, ResNet. Для навчання CNN потрібен величезний обсяг даних, тому для застосування їх до вузькоспрямованих завдань, зазвичай, використовують передбачені на відомих датасетах моделі, а потім, застосовуючи метод переносу навчання (transfer learning), розраховують ваги на останніх шарах мережі. Підходи на основі CNN є найбільш точними і в той же час найвибагливішими до обчислювальних потужностей.

Методи, засновані на вейвлетах Хаара та Габора. За допомогою вейвлетів із зображення отримували найбільш сильні сигнали, фактично виділяючи найважливіші ознаки на зображенні, відкидаючи інше. Вейвлети Габора були представлені в дослідженнях алгоритмів розпізнавання осіб [6; 7]. Різні вейвлети застосовувалися з ідеєю виділення візуальних ознак незалежно від завдання. Надалі дані візуальні ознаки використовувалися в авторських алгоритмах, у тому числі і для знаходження особи, як, наприклад, в алгоритмі Віоли-Джонса [8]. Використання ознак, отриманих вейвлетом Габора, є складним та часо затратним

[9]. Існуючі спрощені вейвлети та алгоритми на їх основі працюють швидше, проте сильно залежать від освітлення та інших зовнішніх факторів [10].

Методи, що ґрунтуються на знаходженні дескрипторів особи. Замість використання всього зображення алгоритми розбивають його на підобласті, у кожній з яких розраховують гістограми і знаходять найбільш яскраві та важливі ознаки для вирішуваної на даний момент завдання.

Надалі отримані ознаки можуть бути використані для завдання визначення особи, так і в задачі ідентифікації, якщо є приклад, для якого також можна отримати ознаки подібним способом. Наприклад, автори застосовували підхід на основі локальних дескрипторів, включаючи різні локальні бінарні шаблони, для завдання ідентифікації по обличчю. При цьому знаходилися такі дескриптори, які були стійкі до вікових змін особи. Дані методи компактні та швидко працюють, добре справляються із змінами у висвітленні та різними емоціями на обличчі. Однак ці підходи дуже чутливі до шуму на зображенні.

Методи, що базуються на тривимірних зображеннях та картах глибини. Багато сучасних камер, і навіть смартфонів, оснащені лідарами і мають можливість швидко та точно отримувати карту глибини зображення [13]. Цей метод зйомки дозволяє значно знизити чутливість алгоритмів до освітлення та зовнішніх факторів зйомки. Моделі, що працюють із тривимірними зображеннями, набагато більш точні та інваріантні до змін ніж працюючі з двовимірними [14]. Методи використовують текстури зображення [15], компенсують варіативне освітлення [16], а також компенсують різні пози людини за допомогою прагнення до симетричного обличчя особи під час аналізу [17]. Такі системи, що вимагають особливого обладнання, ретельного калібрування та синхронізації даних, що робить їх складно застосовними в системах загального призначення.

Визначення статі та віку. Якщо визначення статі є завданням бінарної класифікації, то завдання визначення віку можна вирішувати як мультикласову класифікацію за віковими кластерам, так і як завдання регресійного передбачення кількості років. Наприклад у [18] для визначення віку автори донавчили модель VGG-face методом перенесення навчання на наборі даних із зображеннями людей,

для кожного з яких була мітка про вік. На вихідному шарі нейронної мережі було 8 виходів, відповідних до вікових груп. У [19] для завдання визначення віку також використовувалися CNN (Ranking -CNN), але з бінарним виходом, а проставлення мітки про вік людини відбувалося агрегацією виходів "базових" нейромереж. Кожна з базових згорткових мереж була навчена на мітках приналежності людини до групи "старші" або "молодші" певного віку.

Якщо для завдання визначення особи на кадрі методи не на основі нейромереж мають досить хорошу точність і можуть скласти конкуренцію CNN, то для завдання визначення віку по особі згорткові мережі є неперевершеними лідерами. Так, у 2015 році було представлено мережу AgeNet [20], що демонструє високу точність у завданнях визначення віку по особі та посіла призове місце у змаганні «ChaLearn 2015 apparent age competition» [21]. Підхід є глибокою згортковою нейронною мережею, в якій як функція втрат використовуються як Евклідова норма, так і крос-ентропія. Більшість сучасних методів для визначення статі та віку використовують відомі архітектури глибоких згорткових нейромереж, таких як VGG-16 [22] або ImageNet [19].

1.2. Методи розпізнавання статі та віку за ознаками силуету та руху

Потреба у визначенні статі та віку людини на основі ознак руху або силуету виникає у ситуаціях, коли обличчя людини на кадрах погано помітне у зв'язку із віддаленістю камери, або зовсім не видно. Такі ситуації характерні для систем моніторингу, що встановлюються в місцях значного скупчення людей (як то стадіони, аеропорти, вокзали), щоб охопити на відео якнайбільше простору.

Усі алгоритми розпізнавання статі та віку використовують результати етапів детекції та трекінгу об'єктів. Завдання розпізнавання статі зазвичай представляється як завдання бінарної класифікації, а завдання розпізнавання віку - або як завдання регресії [23], або як завдання багатокласової класифікації [24], коли виділяються вікові групи (наприклад, 0-10, 11-20, 21-30 і т.д.).

Ключовою ознакою, що використовується у підходах до визначення статі та віку на основі ознак форми тіла чи ходи є GEI – зображення, що надає динамічну та статичну інформацію про ходьбу людини. Це один із способів отримання біометричних даних на відстані. GEI, зазвичай, відображає просторово-часові характеристики ходи, такі як довжина кроку, швидкість кроку, ширина кроку, час кроку, час маху, час опори (середні характеристики ходи) та їх відповідну мінливість та асиметрію (динамічні характеристики ходи). Приклад GEI-зображення представлений на рис. 1.2.

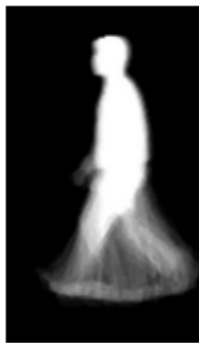


Рис. 1.2. Приклад GEI-зображення

Підходи на основі GEI показують високі метрики якості як у завданнях визначення статі [25, 26], і у завданнях визначення віку [27]. Запишемо основні етапи підходів на основі GEI. Вхідними даними є зображення обмежувачих вікон об'єктів протягом спостереження за ними, отримані на етапах детектування та трекінгу. Нижче наведено основні етапи підходів.

- обрізання та масштабування обмежувальних вікон;
- сегментація об'єктів та бінаризація зображень;
- класифікація за кутами спостереження (опціональний);
- конструювання GEI;
- визначення статі та віку

1.2.1. Обрізання та масштабування обмежувальних вікон. На даному етапі відбувається обрізання об'єктів за відповідними координатами обмежувачих вікон.

На рис. 1.3 представлений приклад зображень, що демонструють результати роботи цього етапу.



Рис. 1.3. Зображення з об'єктом, обрізане за обмежуючим вікном

Після отримання обрізаного зображення об'єкта відбувається його масштабування. У [26] всі зображення наводяться до формату 224x224 пікселів. У [25] - 240x240 пікселів. Щоб не спотворювати пропорції картинки, прийнято додавати чорні відступи (padding) з боків.

У разі низької якості зображення додатково застосовуються підходи збільшення роздільної здатності. У [28] був застосований двоступінчастий метод відновлення зображення з урахуванням CNN. Під час першого кроку із зображення видаляються шум і розмиття за допомогою моделі Image Restoration CNN (IRCNN). На другому кроці відбувається покращення роздільної здатності зображення (Super-Resolution Reconstruction, SR) за допомогою дуже глибоких мереж згортання SR (Very deep convolution networks SR, VDSR).

1.2.2. Сегментація об'єктів та бінаризація зображень. На цьому етапі відбувається віднімання фону, і таким чином зображення сегментується. Методи віднімання фону можна розділити на дві групи: методи, які засновані на нейронних мережах та методи, що не використовують нейромережі. Останні широко застосовувалися до 2010-х років, до появи сучасних згорткових архітектур нейронних мереж. Наприклад, у [29] запропоновано ефективний підхід для визначення фону та видалення тіней на чорно-білому відео. Для визначення об'єктів, що рухаються використовується локальна просторова когерентність.

Підхід ґрунтується на припущенні про те, що пікселі заднього фону на чорно-білому відео, знятому статичною камерою, мають нормальний розподіл. Тоді пікселі об'єкта, що рухається, будуть створювати викиди у розподілі. Тому для визначення параметрів розподілу використовується метрично усереднене середнє. Після первинного визначення фону застосовуються морфологічні оператори для видалення тіней та ізольованих помилок підходу.

Найбільш ефективні підходи до віднімання фону – підходи на основі згорткових нейронних мереж. Першими дослідниками, які застосували такий підхід до завдання виділення фону були Braham і Van Droogenbroeck. Побудована ними нейронна мережа описана в [30] і зветься ConvNet. Для навчання моделі використовувався набір даних з відеозаписами CDnet 2014, що використовується як бенчмарк для змагань [31, 32].

Автори статті [33] використовували різницю між еталонним тлом зображенням (до появи об'єкта) та поточним зображенням та створили свою реалізацію алгоритму віднімання фону. Для адаптації до динамічних змін сцени використовували зображення з інфрачервоної камери.

Після відокремлення об'єкта від фону зображення бінаризується: пікселі зображення, що належать об'єкту, фарбуються в білий колір, а решта в чорний. Приклад бінаризованих зображень наведено на рис. 1.4.

Результатом етапу є послідовність обрізаних та відмасштабованих двійкових зображень силуету (як правило, білий силует об'єкта на чорному тлі).



Рис. 1.4. Приклад бінаризованих зображень об'єкта

1.2.3. Класифікація за кутами спостереження. Потреба у цьому етапі залежить від специфіки місця, у якому ведеться спостереження. У таких місцях як торгові центри люди схильні змінювати напрямок руху упродовж невеликого проміжку часу, повертаючись до камери різними сторонами. Для опису такого типу пересування існує термін "вільна хода" (freestyle walk) – пересування, при якому об'єкт, що спостерігається, змінює напрямок руху і, відповідно, кут руху щодо спостерігача протягом усього тимчасового проміжку спостереження. Протилежність ході в одному напрямку (fixed-direction walk) [26].

Даний етап покликаний збільшити точність розпізнавання статі та віку при вільній ходьбі. Замість того, щоб об'єднувати всі зображення силуетів, витягнуті з усієї прогулянки в характеристику GEI, під час цього етапу відбувається класифікація силуетних зображень на групи у відповідність до напрямку руху об'єкта щодо камери (кут спостереження). Побудова кількох GEI (один GEI для однієї групи силуетних зображень) ефективно показує унікальні характеристики людини. Це пов'язано з тим, що це зменшує перешкоди від зображень силуетів, знятих під різними кутами спостереження. Крім того, деякі зображення силуетів можуть не відображати все тіло суб'єкта через обмежене поле зору камери. Ці часткові силуетні зображення тіла вважаються шумом. Виключення таких силуетних зображень із побудови GEI призводить до більш високої якості GEI.

У статті [26] використовується класифікаційна модель для категоризації силуету зображення на 10 груп кутів спостереження. Кожна група представляє діапазон 36 градусів ($360^\circ / 10$) кута спостереження. Також використовується спеціальна група для часткових зображень силуету тіла, які не повинні бути задіяні в отриманні GEI. Як класифікаційну модель використовується VGG16. Функція активації вихідного шару – softmax на 11 класів. У статті також використовується визначення кута перед побудовою GEI. Для цього автори використовують CNN, навчену на датасетах CASIA-(B) [35], OU-ISIR [36], та OU-MVLP [37].

1.2.4. Конструювання GEI. Після класифікації силуетних зображень прогулянки на групи, відповідні кутам спостереження, силуетні зображення у

кожній групі об'єднуються у усереднені зображення (GEI). Для кожного GEI його значення в позиції (x, y) визначається за формулою (1.1):

$$GEI(x, y) = \frac{1}{N} \sum_{t=1}^N I_t(x, y), \quad (1.1)$$

де N – число зображень силуетів групи, $I_t(x, y)$ – бінарне значення, записане в пікселі зображення на позиції (x, y) i -го силуету, де 1 – піксель належить силуету, 0 – піксель не належить силуету.

Та ж формула використовується і у випадку, коли крок класифікації по кутах спостереження було пропущено. У цьому випадку підсумкове зображення матиме один канал.

У [25] після складання GEI покращити якість моделі розпізнавання статі допомогло застосування локальних бінарних шаблонів з нечіткою логікою (Fuzzification of Local Binary Patterns). Метод допоміг зменшити вплив шумів на бінарні патерни та створити більш інформативні ознаки за рахунок включення нечіткої логіки до процедури розрахунку гістограми змінних кодів.

Існують підходи, які використовують модифікації GEI. Наприклад, в [38] описується підхід, у якому розраховується характеристика, що представляє ходу у тривимірному просторі (GEV) з використанням зображень глибини (depth image). Це модифікація GEI у тривимірному просторі, де аналогом бінарних пікселів силуету служать бінарні вокселі. Отримання таких зображень потребує наявності камер глибини. На рис. 1.5 представлені побудовані 3D- зображення з вокселями і розрахована на їх основі GEV.

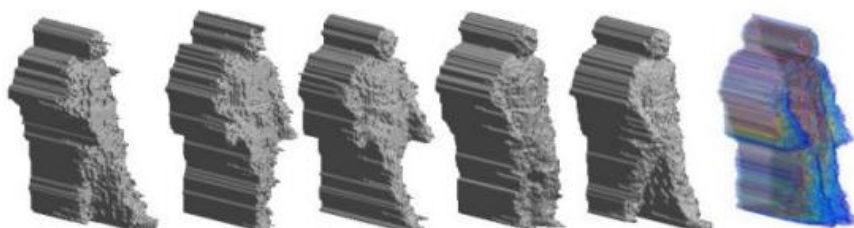


Рис. 1.5. Приклад GEV

У [38] описаний підхід, де розраховується BGEI - представлення ходи, що використовує припущення, що фронтальна поверхня моделі містить всю інформацію про ходу. При використанні BGEI є можливість все-таки втратити інформацію, оскільки задня нога більше представлена у розрахунку характеристики, однак така інформація в деяких випадках може бути непотрібною або вносити мінімальний внесок у точність розпізнавання статі та віку. BGEI можна будувати як з GEI, і з GEV. Для силуетів виду збоку BGEI будується за методом зворотного заливання бінарних силуетів. Для цього знаходиться піксель, що характеризує фронтальну поверхню об'єкта на профільній проекції в кожному рядку, і від нього рядок заповнюється до кінця зображення. Заповнені бінарні силуети вирівнюються по центроїду фронтальної поверхні. На рис. 1.6-1.7 продемонстровано силуети, побудовані за наведеним методом на основі GEI та GEV відповідно.

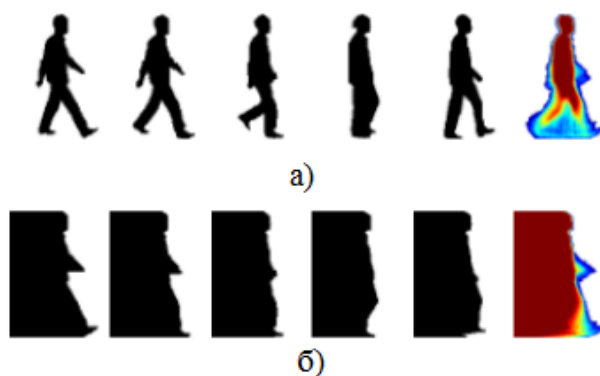


Рис. 1.6. Приклади силуетів з розрахованими: а) GEI, б) BGEI

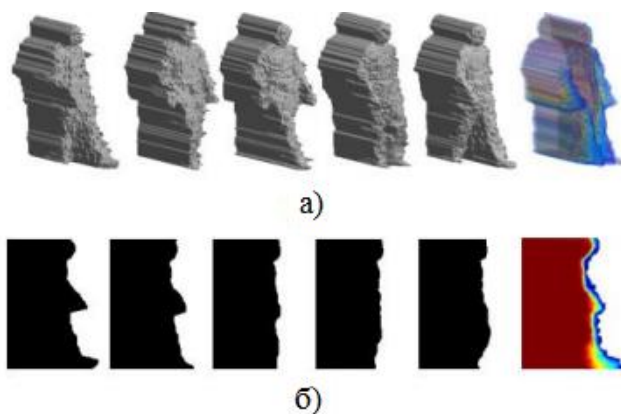


Рис. 1.7. Приклади об'ємних силуетів з розрахованими: а) GEV, б) BGEI

1.2.5. Визначення статі та віку. Перед застосуванням безпосередньо алгоритмів класифікації або регресії для визначення статі чи віку побудовані попередніх етапах N GEI зображень об'єднуються в одне зображення N кількістю каналів. Якщо класифікація за кутами спостережень не проводилася, то кількість каналів дорівнюватиме одному. На рис. 1.8 наведено приклад результуючого зображення розмірністю $224 \times 224 \times 10$ де будувалося 10 GEI [26].

Після отримання результуючого зображення застосовуються методи класифікації чи регресії.

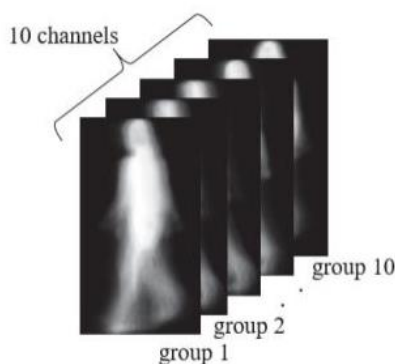


Рис. 1.8. Вхідне зображення для алгоритму класифікації

Це можуть бути як алгоритми машинного навчання, так і нейронні мережі. Для завдання визначення статі в [39] автори використовували згорткову нейронну мережу VGG16. У [26] для класифікації статі автори розробили свою архітектуру CNN, що має 7 шарів. В [25] був застосований метод опорних векторів з лінійним ядром.

Для визначення віку автори статті [24] тестували кілька підходів: як методи класифікації, і регресії. Для класифікації було виділено вікові групи (0-5, 6-10, 11-15, 16-20, 21-30, 31-40, 41-50, 51-60, і більше 60 років). Потім розраховували L2-відстань на основі GEI між сусідніми парами груп, та експериментально встановивши граничне значення для такої відстані, автори поєднали деякі вікові групи в одну. Після об'єднання вийшло п'ять груп. Для класифікації зображень на групи застосовувалася модифікація методу опорних векторів під назвою DAGSVM,

що дозволяє розширити застосування класичного SVM для завдання багатокласової класифікації.

Підхід на основі поздовжніх та поперечних проекцій силуету. GEI є характеристикою, що добре зарекомендувала себе щодо задачі визначення статі та віку, що відображає динамічні та образні особливості ходи. Крім підходів, що базуються на GEI, є також інші підходи, котрі демонструють високі метрики якості завдання визначення віку. Так, наприклад, розглядалося завдання бінарної класифікації людей на «молодих» та «старих» [27]. Як характеристики, що дозволяють визначити, до якої групи відноситься людина, використовувалися SLP і STP відповідно. Для їх побудови спочатку застосовувався підхід до локалізації вертикального положення точок тіла без використання спеціальних датчиків. Далі аналізувалась довжина кроку людини під час уривка відео його ходи – рис. 1.9.

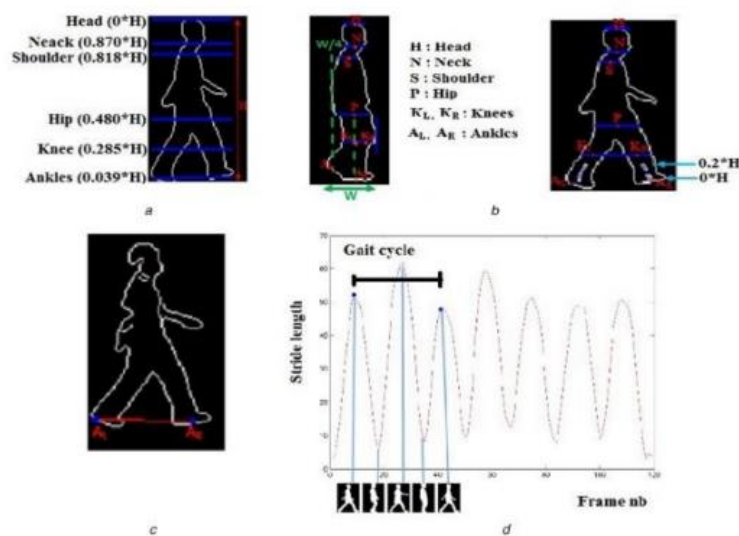


Рис. 1.9. Визначення довжини кроку

Потім на бінарних зображеннях силуетів будувалися SLP та STP з наступною агрегацією та отримання результуючого дескриптора – рис. 1.10.

Авторами підтверджено, що цей дескриптор відображає особливості ходи молодих і людей похилого віку: широкий крок, характерний для молодих людей, проявляється на поздовжній проекції, а ширший силует, як правило, характерний для людей старшого віку, знаходить свій відбиток на поперечній проекції.

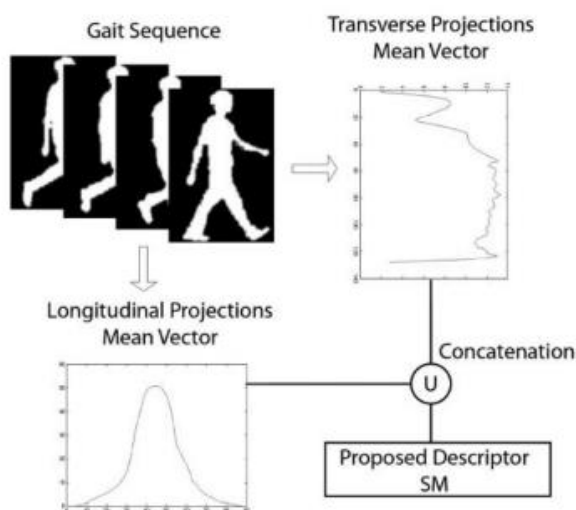


Рис. 1.10. Схема результуючого дескриптора на основі SLP та STP

Підхід із використанням камер різного типу. Типові зображення об'єктів обмежені у можливостях розпізнавання статі та віку через такі фактори, як освітлення, фон, тіні, різні типи одягу та аксесуарів. Тому у статті [28] для визначення статі використовували комбінацію зображень з камер видимого світла та інфрачервоних камер, заснованих на температурній різниці між людським тілом та фоном. У цій статті для зображень з кожної камери автори навчили дві згорткові нейронні мережі ResNet-101 із входньою роздільною здатністю зображень 197x447 пікселів, що враховує пропорції людини. Для прийняття рішення про поле людини ваги моделей на останніх шарах спочатку нормалізувалися, потім агрегувалися для отримання результуючого показника. Якщо значення показника перевищувало експериментальним чином підібране граничне значення, то людина на зображенні класифікувалась як чоловік, інакше – як жінка.

1.3. Висновки до розділу

У першому розділі виконано аналітичний огляд предметної області та теоретичних досліджень методів розпізнавання статі та віку за зображенням особи і за ознаками силуету та руху.

РОЗДІЛ 2

ТЕОРЕТИЧНА ЧАСТИНА

2.1. Теоретичні дослідження методів повторної ідентифікації

2.1.1. Постановка задачі. Створення індивідуальних відбитків об'єктів за графічними ознаками застосовується для забезпечення можливості запам'ятати об'єкт та розпізнати його повторно як на наступних кадрах відеопотоку (для покращення якості роботи алгоритму трекінгу), так і при появі цього об'єкта у відеопотоці через тривалий проміжок часу. Таким чином, за допомогою створення таких відбитків реалізується функціональність повторного анонімного розпізнавання об'єктів без потреби мати попередньо заповнену базу даних деяких ознак об'єктів. У більшості джерел таке завдання має назву задачі ре-ідентифікації об'єкта в відеопотоці [40, 41]. Далі для позначення даного завдання будуть використовуватися обидва терміни.

Завдання ре-ідентифікації об'єкта (Person/Target/Object re-identification, ReID) визначається як завдання пошуку візуальних (графічних) описів об'єкта у галереї, яка є набором таких же описів множини об'єктів, що спостерігалися в відеопотоці, за запитом піднабору візуальних (графічних) описів (query of images / query of features) того ж об'єкта, взятого з інших сцен та / або в інші моменти часу на тій самій сцені відеопотоку. Також завдання передбачає ранжування описів об'єктів галереї зі спадання ступеня схожості з вхідним/запитуваним об'єктом (object of interest) [42, 43]. На рис. 2.1 представлений приклад наборів зображень (query of images) для людей на сцені відеопотоку в різні моменти часу. Результатом роботи алгоритму, що вирішує задачу ре-ідентифікації, є зіставлення запитуваних наборів зображень (праворуч) з набором зображень з галереї, отриманим в інший момент часу (ліворуч).



Рис. 2.1. Приклад результатів роботи алгоритму реідентифікації об'єктів

2.1.2. Обмеження розв'язання задачі. Найбільш точні методи повторного розпізнавання людей ґрунтуються на біометричних ознаках, які можна отримати з графічної інформації відображення об'єктів у відеопотоці (зображення обличчя, особливості ходи та силуету). Однак ці ж методи сильно страждають від умов зовнішнього оточення: взаємні перекриття людей, різні погодні умови та умови освітлення, різна відстань від камери до об'єктів, різна якість відео. Такі умови можуть створити значні складнощі для отримання графічних відбитків об'єктів, заснованих на біометрії. З огляду на те, що для якісної роботи методів реідентифікації графічні ознаки повинні досить відрізнятися для того, щоб успішно розрізняти об'єкти, що однаково виглядають, і одночасно нести загальну інформацію, щоб бути тиражованими на різні умови, то графічні ознаки, що визначаються зовнішнім виглядом об'єктів, є кращими для використання у роботі алгоритмів повторного розпізнавання [45, 46]. Графічні ознаки, що ґрунтуються на біометричних показниках людини, можуть бути застосовані для уточнення результатів ре-ідентифікації лише у тих випадках, коли на камерах відеоспостереження вдається виділити відповідні дані. Рішення про розгляд даних методів для формування індивідуальних відбитків було перенесено на наступний етап робіт.

Далі будуть розглянуті методи повторного розпізнавання (ре-ідентифікації) об'єктів, які використовують як графічний відбиток зовнішні візуальні ознаки об'єктів без біометричної інформації.

2.1.3. Етапи роботи алгоритмів повторного розпізнавання. Усі алгоритми повторного розпізнавання за відбитком графічних ознак працюють із результатами методів детектування, тобто. з обмежувачими вікнами об'єктів. Вихідна платформа, що включає детектор множини об'єктів YOLOv4, повністю задовольняє початковим вимогам запровадження подібних алгоритмів. У загальному випадку конвеєр роботи алгоритмів ре-ідентифікації за графічними ознаками об'єкта ділиться на 3 основні етапи [45]:

- вилучення графічних ознак об'єкта з вікна, що обмежує, на кадрі відеопотоку – отримання графічного відбитка;
- розрахунок дескриптора схожості графічних відбитків - отримання метрики для пошуку у галереї;
- ранжування галереї в порядку зменшення подібності графічних відбитків об'єктів у ній та пошук найближчого поєднання для отримання результату – вилучення з галереї найбільше підходить під вхідний графічний опис об'єкта.

Далі будуть детально розглянуті методи, що застосовуються на різних етапах роботи алгоритмів повторного розпізнавання

2.1.4. Отримання графічних ознак. На даному етапі здійснюється розрахунок графічних відбитків об'єкта, пошук якого необхідно здійснити у галереї. Зазвичай такий відбиток є деяким вектором ознак, отриманим із зовнішнього графічного уявлення об'єкта на кадрі чи кадрах відеопотоку.

Одними з перших для отримання графічних відбитків об'єктів стали застосовуватись ознаки кольору у вигляді колірних гістограм [47 - 49], а також ознаки текстури (GLCM – Gray Level Co-occurrence Matrices, LBP – Local Binary Patterns, Haralik, Laws та інші) [47, 49, 50]. Ці методи дозволяли досить успішно

визначати схожість об'єктів при ре-ідентифікації на множині камер (за готовими розміченими наборами даних), а також розрізняти об'єкти в натовпі. Однак, дані методи сильно залежали від умов зовнішнього середовища, освітлення та якості відеоматеріалу – їх не можна було легко тиражувати на загальні умови, властиві камерам відеоспостереження.

Одним із розвитків згаданих вище методів є боротьба з обмеженнями довкілля. Наприклад, це відбувається шляхом колірної нормалізації, яка дозволяє згладити вплив непостійного освітлення та затінення об'єктів на зображення з різних камер [47]. Крім цього, для задач мультикамерної реідентифікації об'єктів застосовуються методи перенесення яскравості між декількома камерами, що дозволяє усереднити умови зовнішнього освітлення між порівнюваними відбитками графічних ознак [48, 50], покращити загальну якість алгоритму реідентифікації та розширити застосовність запропонованих методів вилучення графічних ознак завдання повторного розпізнавання.

Як додаткова інформація для відбитків графічних ознак, покликаної поліпшити якість ре-ідентифікації об'єктів певних класів [51], використовується підхід виділення областей важливості об'єктах класу людина. Для виділення таких областей використовується алгоритм на графах, що дозволяє виділити карту груп пікселів, що значно відрізняються від навколишніх інших пікселів. Виділені області використовуються як ваги графічних ознак, що описують обмежуваче вікно об'єкта. Іншим підходом, що широко використовується для графічного опису людей в алгоритмах повторного розпізнавання, є використання ознак форми: як додаткова інформація, що описує графічний відбиток, використовується інформація про розміри об'єктів або дані про позу [47, 49, 52], що приймається людиною. Таке збагачення ознак дозволяє легше зіставляти між собою об'єкти на близько розташованих кадрах однієї камери чи на кадри множині камер. Ще більш просунутим методом опису графічних ознак об'єкта класу «людина» на наборі кадрів відеопотоку є використання ознак GEI, що поєднують динамічну та статистичну інформацію про пересування об'єкта в межах треклету [53.]

Іншим напрямом привнесення допоміжної інформації у відбитки об'єктів на

основі графічних ознак є облік середовища, в якому знаходиться об'єкт. У [45, 52] згадуються методи, які крім ознак безпосередньо об'єкта використовують графічні ознаки контексту сцени та спираються на них для одержання точнішої інформації щодо розташування об'єкта в різні часові періоди або на багатьох камерах. Більше просунутий підхід застосовується у роботі [54], де замість ознак лише одного об'єкта класу людина розглядаються ознаки групи людей, поруч із якими знаходиться ця людина. Такий підхід виходить із припущення, що людина, що рухається у відеопотоці на сусідніх кадрах або різних камерах буде перебувати у складі групи людей протягом відносно тривалого проміжку часу. При цьому на наступних етапах роботи алгоритму, зіставлення об'єктів проводиться багатоступінчасто - спочатку ре-ідентифікується група людей, а потім кожна людина в цій групі.

Найбільш популярними та одночасно найбільш точними методами вилучення графічних відбитків для повторного розпізнавання об'єктів на сьогодні є нейромережеві алгоритми, що використовують глибоке навчання для розрахунку найбільш інформативних ознак із обмежуючих вікон. Найбільш успішні архітектури нейронних мереж для таких завдань базуються на згорткових шарах популярних архітектур, таких як VGG16 або ResNet [55, 42]; попередньо навчені на великих наборах даних ваги, після додаткового тонкого налаштування (Fine Tuning) на розмічених даних задачі ре-ідентифікації, показують одні з найбільш найкращих результатів за будь-яких зовнішніх умов. Таким чином, основні дослідження на даний момент спрямовані на покращення якості базових архітектур нейронних мереж шляхом їх більш тонкого налаштування [56, 57], генерації реалістичних зображень для розширення навчальної вибірки [57], аугментації навчальних даних [58], розробки спеціальних шарів для задачі реідентифікації [59] та стійких функцій помилки, наприклад, таких як triplet-loss1 [61, 62], які дозволяють одночасно навчати моделі на подібних і таких, що значно різняться, даних графічних відбитків об'єктів – рис. 2.2.

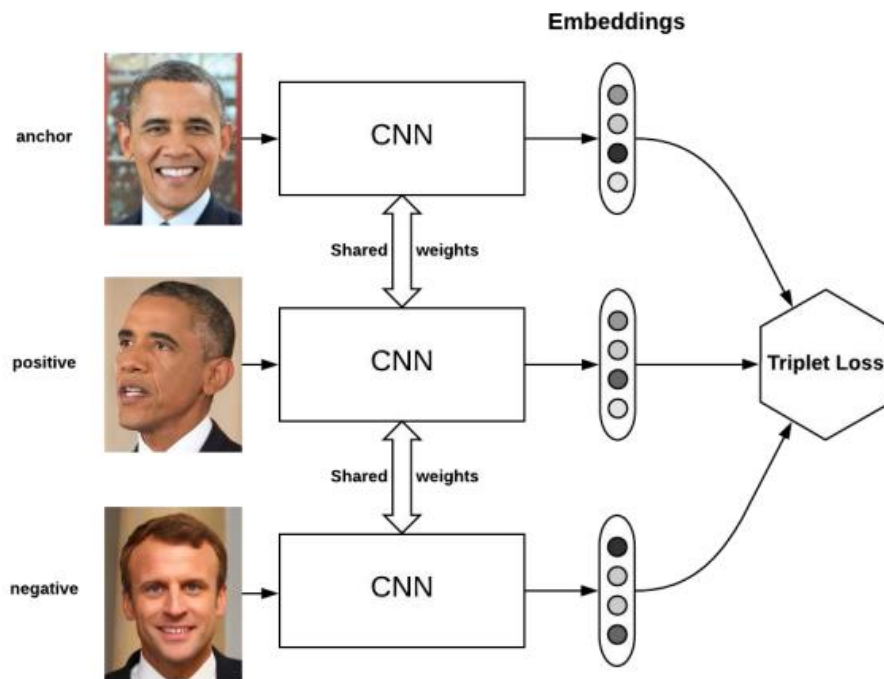


Рис. 2.2. Демонстрація роботи функції помилки triplet-loss

У сучасних моделях, що показують найкращу якість розв'язання задачі реідентифікації – SOTA-рішення – застосовуються підходи окремого вилучення графічних ознак з різних частин, що обмежує вікна, що дозволяє успішно боротися із проблемами різного масштабу об'єктів (об'єкти, що віддаляються і наближаються у відеопотоці), різними умовами зйомки та різною якістю відеоматеріалів [63, 64, 55]. Крім цього, застосовується узагальнення методів отримання графічних ознак для відеоданих: графічні відбитки об'єктів витягуються не з одного кадру, а агрегуються виходячи з інформації про набір кадрів переміщення об'єкта [64]; набір кадрів переміщення об'єкта з треклету використовується як вхідні дані для навчання нейронних мереж з рекурентними шарами [65].

2.1.5. Розрахунок дескриптора схожості. На даному етапі конвеєра алгоритмів ре-ідентифікації вилучені графічні ознаки об'єкта використовуються для розрахунку деякої метрики схожості з аналогічними графічними ознаками галереї. Розрахована метрика є основним інструментом для подальшого пошуку подібних об'єктів у галереї. Найбільш популярними підходами до створення дескриптора схожості є метрики оцінки відстані між двома векторами: L2-норма

(Евклідова відстань) або косинусна відстань [45]. Дані метрики є одними з базових математичних норм розрахунку подібності векторів, які настільки зарекомендували себе у вирішенні завдань комп'ютерного зору і, зокрема, ре-ідентифікації, що застосовуються досі в тому числі і в SOTA-рішеннях [55].

Іншим також досить поширеним на даний момент підходом є застосування моделей глибокого машинного навчання. Так для розрахунку метрик схожості застосовуються сіамські нейронні мережі [67], що дозволяють на основі двох пред'явлених на вхід векторів графічних ознак об'єктів розрахувати для них дескриптор схожості. Розвитком сіамських нейронних мереж для завдання ре-ідентифікації об'єктів у відеопотоці є застосування архітектур рекурентних нейронних мереж типу Long Short-Term Memory (LSTM) для отримання графічних ознак у часовій залежності з одночасним навчанням на основі цих ознак нейронної мережі для розрахунку дескриптора схожості [66, 68]. Приклад архітектури таких мереж представлений на рис. 2.3.

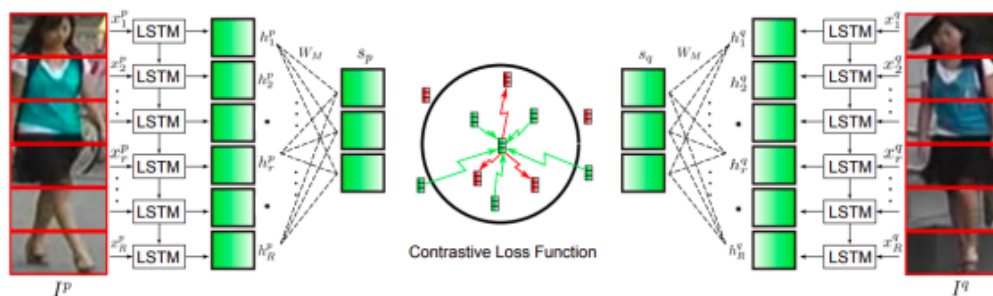


Рис. 2.3. Приклад архітектури сіамської нейронної мережі із застосуванням шарів LSTM

Одним з найбільш розвинених підходів до розрахунку метрики схожості має наступна архітектура нейронної мережі [64]. У ній для розрахунку метрики схожості графічних ознак об'єкта використовується механізм уваги (Attention Mechanism), що дозволяє розрахувати набір просторово-часових метрик між набором вхідних ознак об'єкта та набором ознак з галереї під час навчання мережі – рис. 2.4. Це дозволяє значно підвищити якість повторного розпізнавання об'єктів у різних часових відрізках або відеопотоках з множини камер [64]

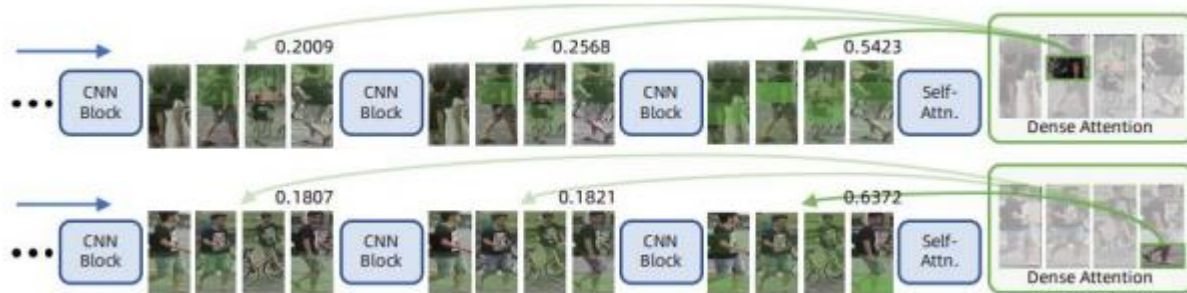


Рис. 2.4. Візуалізація роботи Механізму Уваги для навчання зіставлення окремих ознак зображення ознаками з галереї

2.1.6. Ранжування. На цьому заключному етапі роботи алгоритмів повторного розпізнавання об'єктів за графічними ознаками на основі розрахованої метрики схожості здійснюється ранжування представлених у галереї векторів ознак і визначення найбільш схожого об'єкта (якщо є) на основі перших K векторів галереї, де K – є гіперпараметром алгоритму.

Методи ранжирування поділяються на кілька типів за способом порівняння вхідного набору ознак (query of images) з галереєю та за способом навчання моделі порівнювати ці набори із галереєю. Алгоритми, які для порівняння та ранжування використовують лише один вектор ознак і здійснюють ранжування також за одним вектором ознак з галереї утворюють підмножину one-shot підходів. Алгоритми, які порівнюють і ранжують набори векторів ознак, є multi-shot підходи. За характером навчання алгоритми поділяються на ті, що під час навчання порівнюють і ранжують розмічені приклади навчальної вибірки (навчання з учителем), а також ті, які під час навчання вчать лише виділяти сильні ознаки добре відокремлюють різні об'єкти і дозволяють зіставляти одні й самі об'єкти (навчання без вчителя), а саме зіставлення та ранжування здійснюється за допомогою не нейромережових підходів [45].

Для етапу ранжирування алгоритми ре-ідентифікації пропонують як використання більш оптимальних підходів до впорядкування галереї на основі метрик схожості, так і спрощення самого алгоритму впорядкування шляхом групування у векторні ознаки в галереї. Типовими методами, за допомогою яких здійснюється ранжування та зіставлення вхідного вектора ознак з ознаками з

галереї є алгоритми перебору, пошуку найближчих сусідів (наприклад, k-nearest neighbours) і безпосередньо ранжування (наприклад, Ranking SVM [45], де класичний алгоритм SVM використовується для попарного порівняння вхідного вектора з набором векторів з галереї). До більш сучасних підходів можна віднести алгоритм [69], де застосовується метод інваріантного ранжування опорних векторів (Invariant Ranking SVM), що пропонує навчання алгоритму SVM на об'єднаних графічних ознаках об'єктів для забезпечення інваріантності ранжування щодо різних поз, форм, видів представлення об'єктів та ракурсів камери. Крім модифікацій безпосередньо ранжування, в [70] застосовується нейронна мережа для кластеризації об'єктів у семантичні групи на основі ознак об'єктів з окремих кадрів та зміни цих ознак згодом у відеопотоці. Поділ об'єктів на групи дозволяє покращити якість та оптимізувати процес ранжирування – спочатку відбираються групи, потім об'єкти цих групах.

Сучасні методи ранжирування, що реалізують підходи навчання з учителем, найчастіше є нейромережеві методи, як згадувалися раніше [64, 65], або підходи, що реалізують усі етапи роботи алгоритмів ре-ідентифікації в одній моделі [71, 72] - end-to-end підходи. В останньому методі реалізується автоматичне отримання треклетів об'єктів, вилучення графічних ознак та порівняння їх з ознаками треклета інших об'єктів в даних (multi-shot підхід). Метод узагальнює підходи використання розмічених даних для завдання ре-ідентифікації як в одному відеопотоку, так і на множині камер: здійснюється ранжування та пошук найбільш схожих треклетів об'єктів одночасно у кількох відеопотоках.

2.2. Розробка алгоритму розпізнавання статі та віку

2.2.1. Вибір методів реалізації алгоритму. На основі проведеного літературного огляду, а також враховуючи додатковий аналіз існуючих на сьогоднішній день методів розпізнавання статі та віку людини, які не використовують зображення особи, вилучення ознак з треклету об'єкта було обрано алгоритм побудови GEI зображення, який вимагає проведення етапу сегментації та

безпосередньо створення самого зображення (мапи ознак) [73]. Вибір на користь саме цього підходу було зроблено завдяки ефективності, показаній у кількох успішних роботах та відповідність підходу до особливостей даних, що завантажуються на платформу – відеодані переміщень людей у певному напрямку (людина протягом свого треклету рухається в одну сторону).

Для здійснення етапу сегментації обмежуючого вікна людини, необхідного для побудови GEI-зображення, було обрано рішення PaddleSeg. Дані та інструменти сегментації, що містять у своєму складі, включають зображення людей на повний зріст, на відміну від більшості інших, де зйомки здійснювалися під менш відповідним під дані платформи гострим кутом до горизонту. Завдяки цьому, моделі сегментації не потрібно навчати з нуля, для побудови моделі може використовуватись підхід перенесення навчання.

Завдання розпізнавання віку було вирішено не як завдання регресії та розрахунку точного числового значення, а як завдання класифікації та визначення однієї з трьох вікових категорій. Вибір на користь такого рішення було зроблено на підставі практично повної відсутності відкритих даних, умови зйомки яких наближені до даних, властивих платформі детектування та трекінгу об'єктів, і великий трудомісткості створення власного набору, що відповідає всім вимогам розробки алгоритму розпізнавання статі. Рішення підзадачі класифікації статі та віку було прийнято здійснювати за допомогою CNN, яка використовує для отримання ознак попередньо навчені ваги нейронної мережі esNet, що показує хороші результати у різних завданнях класифікації на зображення реального світу.

2.2.2. Концепція алгоритму. Запропонована версія розробленого алгоритму розпізнавання статі та віку ґрунтується на обраних на поточному етапі робіт методах, покликаних вирішити проблеми, виявлені на етапі розробки першої експериментальної версії алгоритму, і зробити підхід більш універсальним, що працює у найбільш загальних зовнішніх умовах, властивих даних, що завантажуються на платформу детектування та трекінгу об'єктів.

Логіка роботи алгоритму, показана для однієї ітерації обробки одного кадру

відеопотоку, з урахуванням даних та їх доступності на різних етапах роботи вже наявних у складі ядра компонентів для детектування та трекінгу представлена на рис. 2.5.

Основна частина алгоритмів працює з результатом трекінгу – базою відстежуваних на відеопотоці об'єктів, для яких вже було здійснено оновлення стану, відповідне поточному оброблюваному кадру. Таким чином, вхідними даними є набір оновлених відповідно до поточного оброблюваного кадру об'єктів (треків), що зберігають інформацію про останнє вікно, що обмежує. Розроблений алгоритм розпізнавання статі та віку працює з кожним об'єктом, що відстежується незалежно та для кожного i -го об'єкта з N відстежуваних на поточний момент виконує послідовність дій.

Визначає належність об'єкта до класу “Людина”. Тільки для об'єктів цього класу можливе і потрібне визначення статі та віку. Фільтрування класів виступає етапом валідації вхідних даних для алгоритму.

Для кожного об'єкта класу “Людина” вибирається останнє вікно, що обмежує, відповідне поточному оброблюваному кадру i , у термінах ядра платформи, здійснюється витяг графічних ознак з обмежуючого вікна об'єкта, яке в даному випадку має конкретну реалізацію - сегментація людини на зображенні обмежує вікна – виділення пікселів, які стосуються людини.

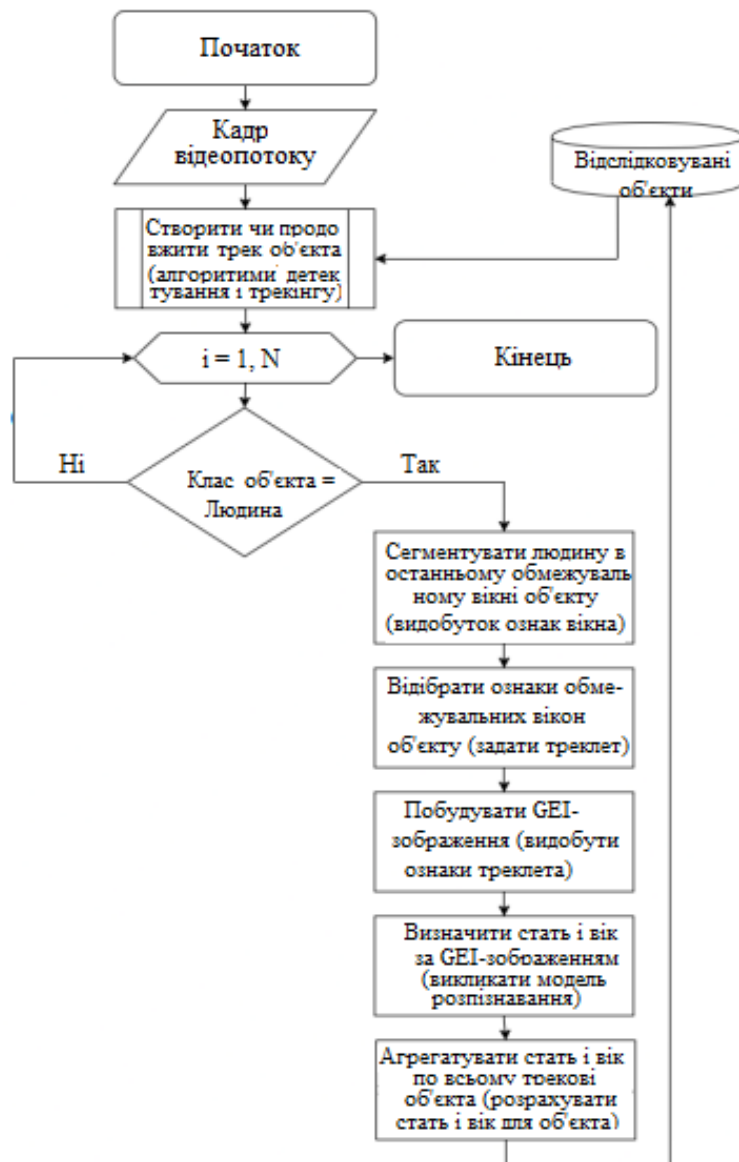


Рис. 2.5. Блок-схема роботи розроблених алгоритмів розпізнавання статі та віку

Ця операція здійснюється за допомогою моделі сегментації MobileSeg з бібліотеки PaddleSeg. Результатом роботи моделі сегментації є бінарне зображення, де значення інтенсивності кольору пікселя, що дорівнює 1, говорить про те, що піксель належить сегменту силуету, а значення 0 не належить. Вилучені ознаки прив'язуються до об'єкту, який зберігає історію таких ознак, що відповідає попереднім обробленим кадрам відеопотоку. Серед набору витягнутих ознак об'єктів, що обмежують вікнами, вибирається піднабір, що описує обмежений тривалістю треклет об'єкта (весь об'єкт описується повним треком). На поточному етапі робіт використовується правило вибору певної кількості останніх об'єктів, що

обмежують вікон.

На основі ознак обмежуючих вікон будуються ознаки треклету, якими у цьому випадку виступає GEI-зображення, яке будується згідно з формулою 1.1, що визначає кожне значення зображення позиції (x, y) .

Здійснюється виклик моделі CNN, попередньо навченої методом перенесення навчання на базі архітектури ResNet50 для визначення статі та віку людини за ознаками його треклету – за GEI-зображенням. Запропонована архітектура мережі, що поєднує вирішення завдань як визначення статі, так і визначення віку, представлена на рис. 2.6.

Результатом роботи моделі класифікації статі та віку є два набори чисел: код класу статі людини з відповідним значенням ймовірності віднесення приклад до цього класу, розрахований моделлю, а також код класу віку людини з аналогічним значенням ймовірності.

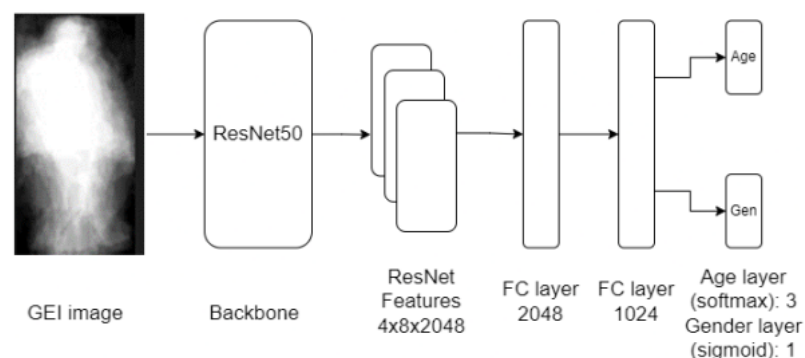


Рис. 2.6. Архітектура CNN, що використовується як класифікатор для розпізнавання статі та віку

На основі всієї історії результатів розпізнавання статі та віку поточного об'єкта здійснюється агрегація як вибір класу, котрий найчастіше зустрічається, і приймається остаточне рішення про значення статі та віку для об'єкта. Агрегація необхідна для того, щоб весь час намагатися уточнити стать та вік об'єкта на відеопотоці, оскільки у різні моменти часу це можна зробити з більшою чи меншою впевненістю. Усі проміжні та фінальні результати роботи алгоритму зберігаються у відстежуваних об'єктах (треках) для використання в наступних ітераціях обробки

кадрів відеопотоку

Загалом конвеєр перетворень даних для одного піднабору обмежуючих вікон об'єкта виглядає так, як це представлено на рис. 2.7.

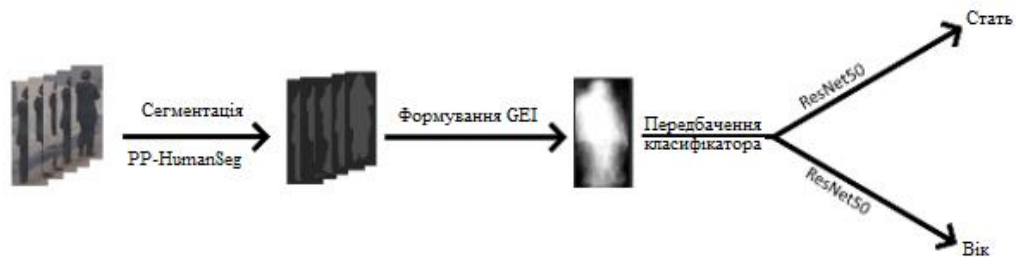


Рис. 2.7. Загальний вигляд конвеєра перетворень даних треклету об'єкта в алгоритмі розпізнавання статі та віку

Запропонована версія розробленого алгоритму розпізнавання статі та віку адаптована під архітектуру роботи ядра платформи, що використовує найбільш характерну для даних платформи інформацію про рух та силует людини, слабо залежить від зовнішніх умов зйомки та може застосовуватися у випадках, коли неможливо розпізнати особу людини, або навіть у разі відсутності інформації про кольори, що описують людину.

2.3. Розробка алгоритму повторної ідентифікації

2.3.1 Вибір методів реалізації алгоритму. Теоретичний огляд можливих методів виконання завдання створення індивідуальних відбитків об'єктів за графічними ознаками для наступного повторного розпізнавання найбільш повно розглядає алгоритми, що застосовуються на сьогоднішній день на різних етапах роботи конвеєра для повторного розпізнавання об'єктів. Вибір підходів, зроблений на першому етапі роботи, є повністю обґрунтованим, проте розробка алгоритмів, необхідна для усунення недоліків експериментальної версії, що вимагає розширення досліджень у вибраній галузі. Так, наприклад, нові дослідження повинні торкнутися вибору найбільш вдалої моделі для отримання графічних ознак з об'єктів обмежуючих вікон, способів представлення ознак всього треклету,

дескрипторів схожості графічних ознак, можливих способів агрегації кандидатів, вибраних із галереї.

Модель, яка вчиться розрізняти вектори ознак об'єктів, наведена на рис. 2.8.



Рис. 2.8. Принцип роботи моделі сіамської нейронної мережі

Методи агрегації кандидатів з галереї, отриманих на основі розрахунку та порівняння дескрипторів схожості, не є обов'язковим етапом алгоритму повторного розпізнавання. Тут можуть застосовуватись підходи на основі різних правил голосування, як це було зроблено на першому етапі робіт, а також різні правила агрегації: середня, максимальна, мода і т.д. Можуть також застосовуватись правила зважування кандидатів для розрахунку агрегованого значення.

Враховуючи обрані напрямки розробки алгоритмів створення індивідуальних відбитків об'єктів за графічними ознаками для повторного розпізнавання та можливості у виборі та реалізації конкретних методів для завдання отримання найбільш інформативних графічних ознак з обмежувальних вікон об'єкта було обрано ряд моделей (ResNet50, ShuffleNet, OSNet, MuDeep,

MLFN) для подальшого проведення експериментів.

Як метод вилучення графічних ознак треклету було обрано поелементний розрахунок середнього значення – від використання для цього архітектур рекурентних нейронних мереж було вирішено відмовитись на користь швидкості роботи першого методу; для експериментів з підбором найбільш відповідного алгоритму розрахунку дескриптора схожості між векторами ознак були відібрані наступні заходи (як часто вживаються у завданнях повторного розпізнавання): косинусна відстань, L1- та L2-норма, відстань Чебившева, відстань Махаланобіса – від використання сіамських нейронних мереж було вирішено відмовитися через наявність невеликого досвіду їх застосування (при розробці першої версії платформи детектування та трекінгу об'єктів проводилися експерименти для покращення алгоритму трекінгу з їх використанням) та економії ресурсів для збереження швидкодії; середнє значення, мода та виважене на кількість однойменних кандидатів середнє значення було обрано для експериментів з методами агрегації відносного всього треку об'єкта.

2.3.2. Концепція алгоритму. Запропонована версія розроблених алгоритмів створення індивідуальних відбитків об'єктів за графічними ознаками для повторного розпізнавання використовує конвеєр, що включає відібрані в результаті експериментів методи, що показують найкращу ефективність з точки зору якості та продуктивності, позбавляючи рішення від недоліків експериментальної версії

Розроблені алгоритми враховують як найбільш загальні зовнішні умови зйомки, властиві даним, що завантажуються на платформу детектування та трекінгу, так і особливості їх впровадження в модуль ядра платформи: доступність даних, структуру та формат, а також інтерфейси наявних програмних компонентів. Крім цього, алгоритми спроектовані таким чином, щоб були широкі можливості щодо подальшої оптимізації продуктивності для досягнення заявлених показників подальшої оптимізації продуктивності показників швидкості роботи платформи.

Логіка роботи конвеєра розроблених алгоритмів створення індивідуальних графічних відбитків для повторного розпізнавання об'єктів, показана на одній

ітерації обробки одного кадру відеопотоку, з урахуванням даних та їх доступності на різних етапах роботи вже наявних у складі ядра компонентів для детектування та трекінгу представлена на рис. 2.9.

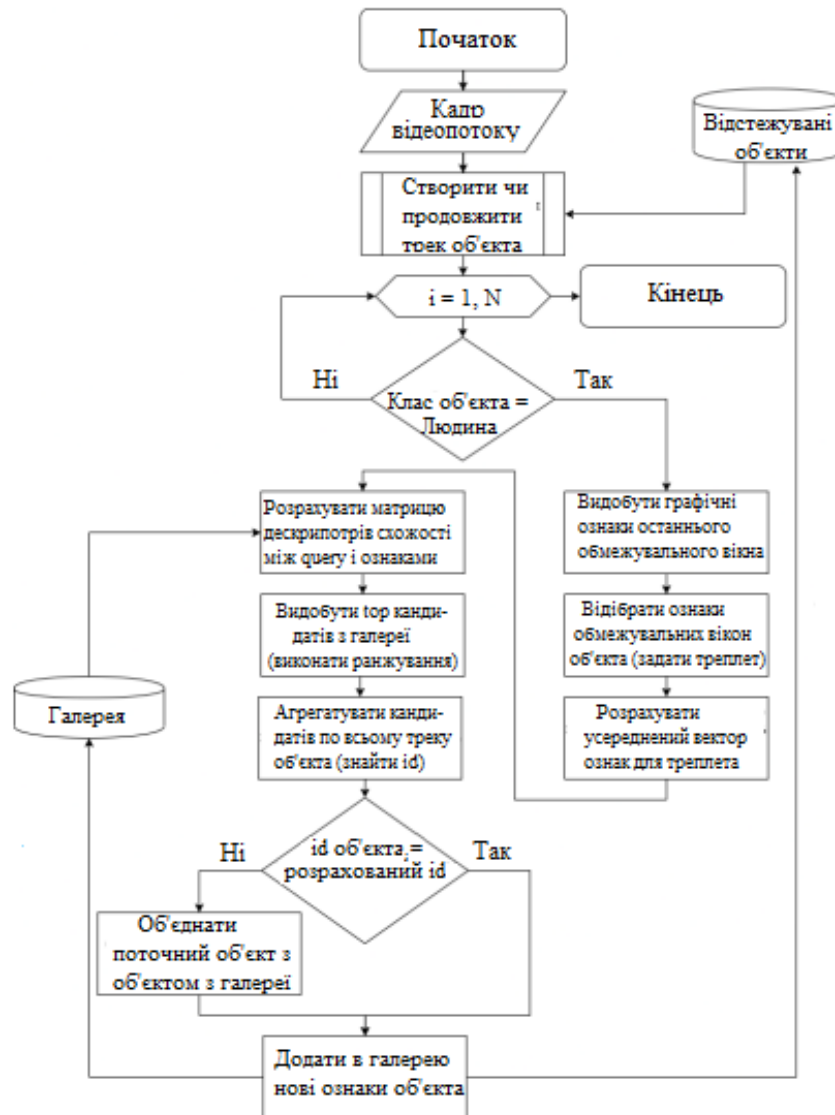


Рис. 2.9. Блок-схема роботи розроблених алгоритмів створення індивідуальних відбитків за графічними ознаками для повторного розпізнавання

Принцип роботи конвеєра розроблених алгоритмів повторного розпізнавання схожий на принцип роботи алгоритму розпізнавання статі та віку. Він також працює з результатами обробки поточного кадру алгоритмами детектування та трекінгу – оновленими даними відстежуваних об'єктів (треків), що зберігають інформацію про останнє обмежувальне вікно. Розроблені алгоритми працюють з

кожним об'єктом, що відстежується, незалежно і для кожного і-го об'єкта з N, що відстежуються на поточний момент виконує послідовність дій.

1. Визначає приналежність об'єкта до класу “Людина”. Тільки для таких об'єктів на даний момент потрібна реалізація повторного розпізнавання, а також тільки на об'єктах цього класу були навчені моделі для вилучення графічних ознак. Фільтрування класів виступає етапом валідації вхідних даних.

2. Для кожного об'єкту відібраного класу здійснюється витяг графічних ознак з останнього вікна, що обмежує. Цю операцію виконує відібрана за результатами експериментів та попередньо навчена за допомогою перенесення навчання на наборі спеціально підготовлених даних, модель MLFN. Результатом роботи даної моделі є вектор графічних ознак, відповідний останньому обмежуючому вікні об'єкта. Вилучені ознаки додаються до об'єкта для зберігання разом з історією таких же ознак, відповідних попереднім обробленим кадрам відеопотоку.

3. Для відбору треклету здійснюється вибір графічних ознак обмеженої кількості вікон об'єкта, що обмежують. На поточному етапі робіт використовується правило вибору останніх об'єктів, що обмежують.

4. На основі відібраних ознак треклетів обмежуючих вікон здійснюється розрахунок ознак всього треклету шляхом усереднення відповідних значень векторів ознак обмежуючих вікон згідно з формулою 2.2 (формула показує розрахунок значення вектора усереднених значень ознак TFV, що знаходиться у позиції x);

$$TFV(x) = \frac{1}{N} \sum_{i=1}^N I_i(x), \quad (2.2)$$

де N – кількість векторів ознак обмежуючих вікон треклету, $I_i(x)$ – значення ознаки позиції x і-го вектора. Отриманий середній вектор ознак є запитом (query), необхідним для здійснення розрахунку дескрипторів схожості з ознаками треклетів з галереї, їх подальшого ранжування та відбору, необхідного для повторного розпізнавання.

5. Маючи усереднений вектор графічних ознак, що описує треклет об'єкта, який є запитом (query) у задачі повторного розпізнавання, здійснюється розрахунок матриці дескрипторів схожості між query і ознаками галереї. Як дескриптор схожості використовується косинусна відстань між векторами ознак (формула 2.3), відібраний в результаті експериментів як метод, який показав найкращу якість:

$$D(u, v) = 1 - \frac{(u) \cdot (v)}{|u|_2 \cdot |v|_2}, \quad (2.3)$$

де u та v вектори. Матриця дескрипторів схожості є результатом виконання цього етапу алгоритму.

6. На основі отриманої матриці дескрипторів схожості здійснюється ранжування галереї та вилучення обмеженого набору (top) найбільш підходящих кандидатів. Ранжування здійснюється простим перебором та порівнянням значень дескрипторів схожості. Результатом етапу ранжування є top найбільш відповідних кандидатів графічних ознак треклетів об'єктів з галереї, кожен з яких прив'язаний до певного ідентифікатора (id) об'єкта.

7. На основі всієї історії вилучених кандидатів для одного об'єкта за допомогою методу агрегації, описаного формулою 2.4, відібраної за результатами експериментальних досліджень, ухвалюється рішення про ідентифікатор поточного об'єкта:

$$S^{id} = \frac{\frac{1}{N} \sum_{i=1}^N d_i^{id}}{\left(\frac{N^{id}}{N}\right)^c}, \quad (2.4)$$

де S^{id} – значення метрики для агрегації (чим вона менша для кандидата з певним id , тим паче придатним є цей кандидат), d_i – значення дескриптора схожості i -го кандидата з query, що має ідентифікатор id , N – кількість кандидатів, N^{id} – кількість кандидатів з ідентифікатором, що збігається з ідентифікатором id , що належить query, c – регулювальний гіперпараметр. Розрахований за допомогою агрегації

ідентифікатор може збігатися з поточним ID об'єкта, а може і не збігатися. Якщо id не збігається, це означає, що в галереї вже зберігається інформація про поточний об'єкт, що відстежується, що, у свою чергу, означає, що поточний об'єкт з'явився на відеопотоці повторно і йому потрібно присвоїти ідентифікатор, відповідний його першій появі на відеопотоці - об'єднати поточний об'єкт з об'єктом із галереї. Результатом даного етапу роботи конвеєра алгоритму є визначення коректного поточного ID об'єкта - ре-ідентифікація об'єкта.

8. Після завершення етапу ре-ідентифікації незалежно від того, чи сталося об'єднання поточного об'єкта з об'єктом з галереї, ознаки треклету поточного об'єкта додаються до галереї для забезпечення можливості повторного розпізнавання цього ж об'єкта на кадрах відеопотоку.

Процес отримання та перетворення даних, що відбувається в розробленому алгоритмі створення індивідуальних відбитків об'єктів за графічними ознаками для повторного розпізнавання, показано на рис. 2.10.

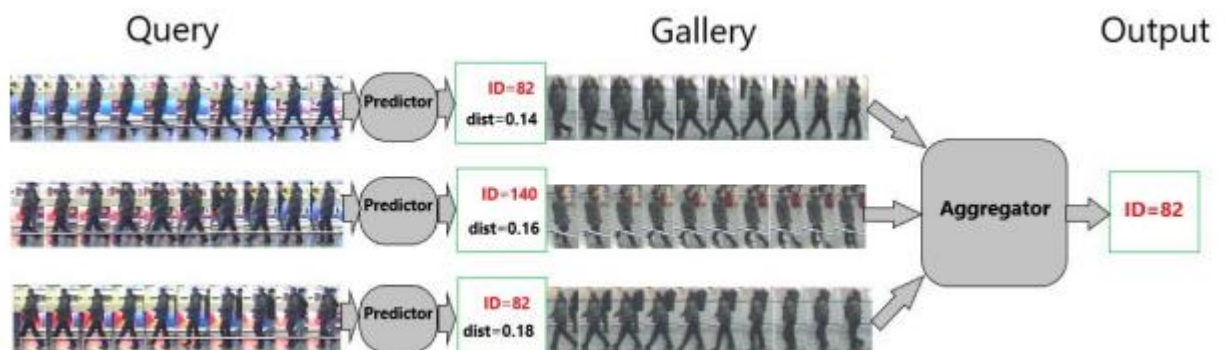


Рис. 2.10. Схема перетворення даних у конвеєрі розроблених алгоритмів створення індивідуальних графічних відбитків об'єктів для повторного розпізнавання

На рис. 2.10 у компоненті Predictor укладено процес отримання ознак треклету з наступним розрахунком дескриптора схожості та ранжуванням; компонент Aggregator визначає спосіб агрегації результатів ранжування та видає остаточне рішення щодо значення ідентифікатора об'єкта.

Запропонована концепція роботи розроблених алгоритмів створення

індивідуальних відбитків об'єктів за графічними ознаками для їх подальшого повторного розпізнавання адаптована до загальних умов зйомки, властивих даним, завантажуваним на платформу детектування та трекінгу, і дозволяє досягти кращих показників якості (результати експериментів з оцінкою якості роботи алгоритмів наведені у розділі 3), усуваючи недоліки експериментальної версії. Додатковим плюсом є широкі можливості щодо оптимізації продуктивності алгоритмів: відбір вікон треклету, що обмежують, може здійснюватися на основі їх якості (вибирати тільки ті з них, де людина виразно помітна і не перекривається іншими об'єктами); можлива подальша оптимізація алгоритмів ранжирування (наприклад, з допомогою алгоритмів k-наближених сусідів чи дерев пошуку); додавання ознак в галерею також може враховувати якість треклету, що обмежують вікон, на основі яких він був обраний; обробка даних кожного об'єкта може виконуватися паралельно; завдяки роботі з треклетом об'єкта виклик всього конвеєра повторного розпізнавання може здійснюватися не на кожній ітерації обробки чергового кадру відеопотоку, а набагато рідше.

2.4. Висновки до розділу

В цьому розділі описано особливості методу повторної ідентифікації, докладно розглянуто всі етапи роботи алгоритмів повторного розпізнавання.

Проведено вибір методів для реалізації алгоритмів розпізнаванні статі та віку і етапи роботи алгоритмів повторного розпізнавання, досліджено концепції цих алгоритмів.

РОЗДІЛ 3

ПРАКТИЧНА ЧАСТИНА

3.1. Інструменти розробки

Для програмної реалізації алгоритмів розпізнавання статі та віку, а також алгоритмів створення індивідуальних графічних відбитків об'єктів для повторного розпізнавання, які використовують сучасні нейромережові підходи для вирішення завдання було обрано Python як основну мову програмування.

Для втілення алгоритмів розпізнавання статі та віку було обрано підхід, який використовує модель глибокої нейронної мережі для сегментації зображень обмежуючих вікон об'єктів, що створює GEI-зображення за сегментованими зображень треклету і використовуючи згорткову нейронну мережу на базі ResNet50 для одночасної класифікації статі та віку.

Як інструменти для розробки моделі сегментації була обрана відкрита бібліотека PaddleSeg, що надає великий вибір попередньо навчених моделей сегментації для різних завдань. Вибір на користь PaddleSeg був зроблений завдяки тому, що серед знайдених рішень, це дозволяло вирішити одну з найбільш критичних проблем експериментальної версії алгоритмів - нестача даних, релевантним завданням, властивим вихідній платформі. Для вирішення цієї проблеми PaddleSeg, на відміну від інших рішень, пропонує попереднє навчені моделі на даних, близьких до найбільш загальних умов зйомки, де люди потрапляють у кадр повний зріст. Використання моделей з PaddleSeg дозволяє здійснити підхід перенесення навчання та покращити показники якості розв'язання задачі сегментації людини на зображення. Як конкретну реалізацію було відібрано модель MobileSeg, MobileNetV3, що використовує як backbone-мережу, – це модель, яка забезпечує найкращий компроміс між якістю та продуктивністю і може бути легко впроваджено в модуль ядра платформи.

Для розробки моделі класифікації статі та віку на основі попередньо навчених ваг ResNet50 було обрано фреймворк TensorFlow з надбудовою Keras, так

як він вже використовується в роботі ядра платформи, і Keras відразу надає всі необхідні інструменти для підвантаження ваг попередньо навченої моделі та реалізації підходу перенесення навчання.

Для розробки алгоритмів створення індивідуальних відбитків об'єктів за графічними ознаками для повторного розпізнавання використовуються моделі та методи відкритої бібліотеки Torchreid, написаної на популярному фреймворку для реалізації моделей глибокого навчання PyTorch і включає в себе множину методів для вилучення та обробки графічних ознак для завдання повторного розпізнавання. Бібліотека Torchreid була обрана через те, що вона підтримує множину сучасних SOTA-моделей та наборів даних для завдання ре-ідентифікації; має відкриту реалізацію, що активно підтримується співтовариством; підтримує методи роботи як із зображеннями, так і з відеопотоками; має модульний дизайн з можливістю його простого розширення своєю функціональністю. Всі моделі та методи, що використовуються для розробки алгоритмів повторного розпізнавання графічними ознаками були взяті з бібліотеки Torchreid.

3.2. Реалізація програмного модуля розпізнавання статі та віку

Програмні компоненти алгоритмів розпізнавання статі та віку реалізовувалися таким чином, щоб потім бути легко вбудовуваними в конвеєр алгоритмів обробки відеопотоків модуля ядра: компоненти реалізовувалися згідно модульної парадигми проектування програмних систем, використовувані методи бути легко конфігурованими, вся архітектура підсистеми підтримується і є відкритою для подальших оптимізацій продуктивності та покращення якості.

Нижче наведені деякі діаграми класів основних функціональних блоків алгоритму розпізнавання статі та віку. Для даних діаграм було зроблено такі припущення:

- діаграми описують окремі логічні блоки, отримані шляхом умовного поділу для зручності подання загальної структури програмної системи;
- для класів, які вже були присутні на наведених раніше діаграмах або

детально описуються в інших діаграмах, прийнято спрощене відображення, що включає лише назву класу;

- на діаграмах показані лише найважливіші з погляду логіки роботи алгоритму розпізнавання статі та віку на основі GEI-ознак класи, їх поля та методи;
- поля та методи класів, що не належать до прямої функціональності алгоритму опущені;
- з метою спрощення на діаграмах опущено деякі бібліотеки функцій;
- опущено деякі аргументи методів класів, які не впливають на основну логіку роботи;
- значення, визначені у класах-перерахуваннях, наведені у вигляді рядкових змінних;
- на діаграмах прийнято наступне спрощення відображення конструкторів класів: опущено аргументи конструкторів, які повністю збігаються за назвою з полями класу.

На рис. 3.1 представлена діаграма класів функціонального блоку вилучення ознак для розпізнавання статі та віку, до яких належать сегментація зображень обмежувальних вікон об'єктів, отримання треклету, створення GEI-зображення як операція вилучення ознак з треклету.

На діаграмі класів рис. 3.1 всі сутності можна поділити на блоки вилучення графічних ознак із зображення обмежуючого вікна (класи `PaddleSegmentationBBoxVFE` і `BBoxVFEBASE`, а також `PaddleSegmentation`), вибору треклету, що описується ознаками об'єктів, що обмежують вікон (класи `LastBboxVisualFeaturesTOFHP` і `TrackedObjectFeaturesExtractorBase`), вилучення ознак з треклетів у вигляді GEI-зображення (`GEITrackedObjectFeaturesExtractor` і `TrackedObjectFeaturesExtractorBase`), а також блок класів, що пов'язують роботу алгоритмів розпізнавання статі та віку з конвеєром відеоаналітики модуля ядра платформи (класи `StreamManager`, `Stream` та `Tracker`). Окремо показані класи `Resnet50AgeGenderRecognizer`, `AgeVotingAggregator` і `GenderVotingAggregator`, які докладніше описуються на наступній діаграмі.

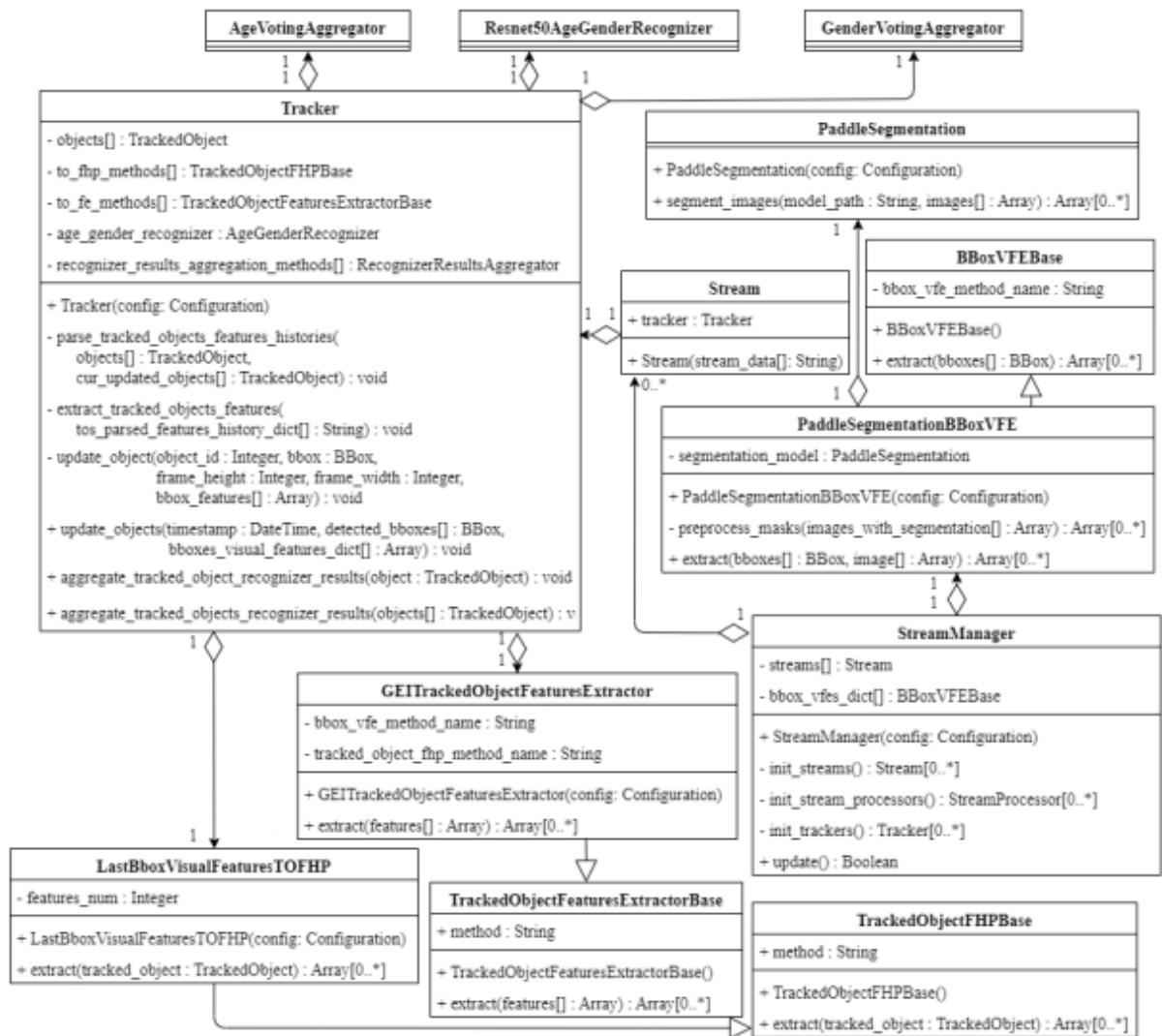


Рис. 3.1. Діаграма класів функціонального блоку отримання ознак для розпізнавання статі та віку

З точки зору конвеєра обробки відеопотоків модуля ядра, функціональність вилучення ознак з обмежувальних вікон є окремим компонентом, робота якого викликається за допомогою класу StreamManager у методі update, де зосереджено всю логіку обробки чергового кадру відеопотоку. У методі update відбувається виклик методу extract класу PaddleSegmentationBBoxVFE, спадкоємця класу BBoxVFEBase, де здійснюється звернення до методу segment_images класу PaddleSegmentation, який, у свою чергу, інкапсулює роботу з навченою моделлю MobileSeg бібліотеки PaddleSeg. Сегментовані зображення після обробки метод preprocess_masks повертаються в StreamManager, де прив'язуються до поточних відстежуваних об'єктів. Так здійснюється витяг графічних ознак з обмежуючих

вікон об'єктів, необхідне для подальшої побудови GEI зображення по треклету. Вся наступна логіка роботи алгоритму розпізнавання статі та віку здійснюється шляхом виклику методів відповідних класів із класу Tracker, оскільки необхідною умовою роботи алгоритму є завершення оновлення положення об'єктів на поточному кадрі, що обробляється – після отримання результатів від алгоритму трекінгу.

3.3. Розробка та розмітка навчальних вибірок

3.3.1. Дані для завдання сегментації об'єктів. В даний час у відкритому доступі знаходиться величезна кількість розмічених наборів даних, призначених для навчання та тестування моделей, які вирішують різні підзавдання сегментації об'єктів на зображеннях. Для підготовки моделі сегментації, з урахуванням результатів якої здійснювалося побудова GEI-зображення об'єкта класу людина, необхідне розробки алгоритмів визначення статі та віку людей, використовувався набір даних UP-S31. Це специфічний набір даних, призначений безпосередньо для вирішення задачі сегментації тіла людини на зображенні та відокремленні його від решти фону. Особливість даного набору полягає в розмітці сегментів не просто тіла людини, але окремих її частин – приклад показаний на рис. 3.2.

Вихідний набір даних містить 8515 розмічених зображень. Для роботи з набором не використовувалася інформація про окремі частини тіла – всі частини бралися в сукупності визначення сегмента всього тіла людини. Через малу кількість вихідних даних, вони піддавалися аугментації за допомогою наступних перетворень: зміна пропорцій зображення, кадрування зображення, ефект блюра, колірна нормалізація, корекція кольору. Аугментації застосовувалися випадково на етапі навчання моделі сегментації. Саме навчання проводилося на 80% даних від їх початкової кількості (6812 прикладів у навчальній вибірці), а якість валідувалася на 20%, що залишилися (1703 прикладів у валідаційній вибірці).

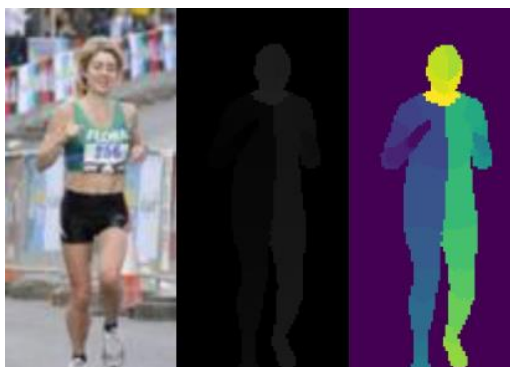


Рис. 3.2. Приклад вихідного зображення та розмітки сегментів тіла людини з набору даних UP-S31

3.3.2. Дані для задачі визначення статі та віку. Через особливості методів визначення статі та віку, що застосовуються в розроблених версіях алгоритмів, які використовують GEI-ознаки для вирішення завдання, кожен приклад набору даних повинен складатися з набору зображень руху однієї й тієї ж людини (треклет) в одному напрямку, а розмітка повинна містити інформацію про сатаь та вік цієї людини. Приклади таких наборів зображень представлені на рис. 3.3. Незважаючи на те, що у відкритому доступі існує безліч наборів даних, що містять таку інформацію, всі вони вимагають попередньої обробки і, можливо, додатковий розмітки для того, щоб бути готовими до використання у рамках розробки алгоритмів розпізнавання статі та віку.



Рис. 3.3. Приклади треклетів зображень однієї людини

Для розмітки та організації наборів даних із відкритих наборів запропонована файлова структура організації треклетів зображень, що включає розмітку і представлена у такому вигляді (рис. 3.4):

```

[dataset-1]
  [subset-1]
    [person-id_gender-id_age-id]
      img1.jpg
      img2.jpg
      ...
    ...
  [subset-2]
  ...
[dataset-2]
...

```

Рис. 3.4. Приклад файлової структури треклетів

Структура розмічених даних є набором папок з наборами даних (dataset-1, dataset-2 і т.д.), всередині розташовані підпапки піднаборів даних (subset-1, subset-2 – для відділення один від одного, наприклад, навчальної та валідаційної вибірок), підпапками піднаборів даних є папки з треклетами, в назві яких присутні код-ідентифікатор людини (person_id), код-ідентифікатор статі людини (gender_id) та код-ідентифікатор віку (age_id).

Як можливі значення ідентифікаторів статі використовувалися значення 0 – жіноча стать та 1 – чоловіча стать. Як можливі значення ідентифікаторів віку використовувалися значення 1 – вік у діапазоні до 15 років, 2 – вік у діапазоні від 16 до 45 років, 3 – вік понад 45 років.

Використані відомі набори даних, підготовлені для навчання та валідації розроблених алгоритмів розпізнавання статі та віку: PETA – Town Centre та MARS. Характеристики набору MARS наступні: кількість зображень = 1191003; кількість унікальних об'єктів = 1028; співвідношення класів статі (0 / 1) складає 605 / 423; співвідношення класів віку (1/2/3) складає 56 / 900 / 72. Крім оригінальних зображень треклетів у час навчання моделей визначення статі та віку також використовувалися та аугментовані зображення, де як методи аугментації застосовувалися дзеркалювання та афінні перетворення. Для навчання та тестування моделей набори даних були поділені на навчальну та валідаційну підвибірki у співвідношенні 75 до 25% відповідно.

3.3.3. Дані для завдання створення індивідуальних відбитків об'єктів за графічними ознаками. Типові дані для вирішення подібних завдань є треклетами, з якими співвідноситься інформація про ідентифікатор об'єкта; при цьому в наборі міститься набір різних треклетів однієї і тієї ж людини, як частин одного треку або як результатів запису з використанням кількох камер. До даних для повторного розпізнавання, також як і даних для визначення статі та віку, застосовувався алгоритм попередньої обробки, що усуває проблеми неточної розмітки меж.

У табл. 3.1 представлені дані, що використовуються для розробки алгоритмів створення індивідуальних графічних відбитків об'єктів для забезпечення можливості повторного розпізнавання, а також їх основні характеристики.

Таблиця 3.1

Відкриті набори даних, підготовлені для побудови моделей створення індивідуальних відбитків об'єктів за графічними ознаками

Назва набору даних	Кількість зображень	Кількість унікальних об'єктів
CAVIAR	177759	121
I-LIDS-VID	43059	300
Market-1501	32668	1259
MARS	1191003	1261
PRID2011	96121	934
DukeMTMC-VideoReID	811776	1812

Для навчання представлені дані в рівному співвідношенні розбивалися на піднабори для навчання, для опису галереї та для опису вхідного набору ознак (query). При цьому, набір даних CAVIAR використовувався для тестування і проведення експериментів, оскільки дані у цьому наборі максимально наближені до зовнішніх умов зйомки, властивих для вхідних даних вихідної платформи.

3.4. Експериментальні випробування

3.4.1. Умови проведення випробувань. Усі випробування та дослідження здійснювалися на фізичних обчислювальних машинах з наступними характеристиками:

- процесор (CPU) Intel Core i5-7600, 4 x 3.50 ГГц;
- графічний прискорювач (GPU) Nvidia GeForce GTX 1080 Ti;
- оперативна пам'ять (RAM) 16 (DDR4) Гб;
- операційна система Ubuntu 18.04 (64 bit).

3.4.2. Випробування моделей сегментації для розпізнавання статі та віку. Як найбільш відповідна нейронна мережа для вирішення поставленого завдання було обрано модель MobileSeg із відкритої бібліотеки PaddleSeg, яка використовує як backbone-мережу легковагу MobileNetV3, яка дозволяє зменшити вимоги до продуктивності модуля ядра при її впровадженні роботу платформи.

Для доналаштування ваг моделі сегментації використовувалися зібрані дані.

Оцінка якості роботи моделі на валідаційній підвибірці здійснювалася за допомогою кількох метрик: Mean Intersection Over Union (MIOU) – усереднене значення метрики IOU для всіх прикладів валідаційної вибірки, Ассурасу – частка правильно виділених пікселів зображення обмежує вікна людини, Precision – величина, зворотна частці виділених пікселів, що не відповідають сегменту людини (частка відсутності хибних спрацьовувань для пікселів, що належать сегменту людини), Карра - коефіцієнт узгодженості, що враховує випадкові збіги.

Всі наведені вище метрики оцінки якості мають область значень від 0 до 1. Чим вище значення метрики, тим вища підсумкова якість. Мета експерименту полягала в оцінці кількості епох донавчання нейронної мережі для сегментації пікселів зображення обмежуючого вікна, що належать людині за допомогою наведених вище метрик. Результати експериментів представлені у табл. 3.2.

Результати експериментів з оцінки кількості епох донавчання нейронів мережі для сегментації людини на зображеннях обмежуючих вікон

Кількість епох	MIOU	Accuracy	Precision	Kappa
500	0,804	0,936	0,794	0,773
1000	0,814	0,940	0,797	0,786
1250	0,825	0,944	0,822	0,800
1500	0,823	0,943	0,815	0,797
2000	0,827	0,946	0,832	0,803
3000	0,831	0,947	0,831	0,808
5000	0,834	0,948	0,830	0,812
10000	0,838	0,949	0,846	0,816
20000	0,839	0,949	0,844	0,813

За результатами, наведеними в табл. 3.2, видно, що кількість епох для донавчання моделі сегментації має бути не менше 10000. При цьому значенні вимірювані метрики якості мають найкращі показання. За значення 20000 метрики стають трохи гіршими – мережа перенавчається. Візуалізація результатів роботи моделі також показує помітні поліпшення після її навчання – рис. 3.5.

Таким чином, результатом проведених експериментів є визначення найкращої кількості епох (10000) для донавчання моделі сегментації людини на зображення обмежуючих вікон, наближених до найбільш загальних умов зйомки, властивою платформі детектування та трекінгу. Від якості роботи цієї моделі залежить якість роботи всього розробленого алгоритму розпізнавання статі та віку.



Рис. 3.5. Приклад сегментації за допомогою вибраної моделі: а) вихідне зображення з набору даних, що не брало участь у навчанні та валідації, б) результат сегментації моделі без донавчання, в) результату сегментації моделі з донавчанням на 10000 епохах

3.4.3. Випробування алгоритмів розпізнавання статі та віку. Завданням проведення експериментів є оцінка якості роботи розробленого алгоритму при використанні різних методів, що входять до його складу, та вибір найкращих з них для використання у складі модуля ядра вихідної платформи.

Оцінка якості роботи алгоритмів розпізнавання статі та віку здійснюється за трьома основними метриками, що широко використовуються в задачах класифікації: точність (precision), повнота (recall) та міра F1. Дані метрики розраховуються для кожного класу окремо, який можна назвати цільовим.

Для проведення випробувань розроблених алгоритмів розпізнавання статі та віку використовувалися наступні методи з відповідними гіперпараметрами.

Архітектура моделі згорткової нейронної мережі на базі ResNet50:

- ResNet50 з додаванням повнозв'язкових шарів та одного виходу для класифікації статі чи віку;
- ResNet50 з додаванням повнозв'язних шарів і двох виходів для одночасної класифікації статі та віку.

Розмір треклета відстежуваного об'єкта для розрахунку GEI-зображення, використовуюваною як ознаки для моделей класифікації статі та віку (максимальний

розмір треклета визначався обсягом доступної пам'яті GPU для навчання та розумним часом навчання моделі):

- не обмежений – для формування GEI використовується довільної кількості обмежувальних вікон об'єкта;
- обмежений і дорівнює числу в діапазоні від 1 до 6.

Метод розрахунку метрик для агрегації кількох результатів роботи алгоритму розпізнавання статі та віку за треклетом та прийняття остаточного рішення для всього відстежуваного об'єкта:

- середнє значення ймовірностей приналежності до класу кожному з класів. Вибирається клас із найбільшою середньою ймовірністю (mean);
- мода серед класів. Як остаточне рішення вибирається клас, що найчастіше зустрічається (mode);
- середнє зважене значення ймовірностей приналежності до класу для кожного класу. Зважування середнього значення складає величину, рівну частоті найпопулярнішого серед агрегованих результатів класу (weighted mean).

Експерименти умовно поділено на 2 частини: відбір моделей для класифікації статі та віку, відбір використовуваної з обраною моделлю функції агрегації.

Результати експериментів з відбору моделі класифікації статі та віку із врахуванням різних підходів до формування треклету представлені в табл. 3.3. За результатами, наведеними в табл. 3.3, видно, що більш висока якість задачі класифікації статі, так і в задачі класифікації віку показує модель з двома виходами. Це може пояснюватися тим, що під час навчання відразу на двох виходах, згорткові шари архітектури ResNet отримують такі ознаки, які одночасно інформативні як задачі розпізнавання статі, так задачі розпізнавання віку. Також це може говорити про те, що визначити, наприклад, вік людини простіше за ознаками силуету та руху, знаючи її стать. В результаті, найкращою моделлю, що показує метрики якості класифікації статі 0.84 за мірою F1 та віку в середньому 0.49 за мірою F1, є нейронна мережа з двома виходами для одночасної класифікації статі та віку, що використовує GEI-зображення, побудовані за треклетами об'єктів фіксованого розміру 6.

Таблиця 3.3

Результати експериментів щодо відбору моделі класифікації статі та віку

№	Кільк. виходів	Розмір треклета	Precision				Recall				F1			
			Вік	<15	15-45	>45	Вік	<15	15-45	>45	Вік	<15	15-45	>45
1	1	не обмеж.	0.75	0.40	0.50	0.63	0.65	0.48	0.22	0.35	0.69	0.43	0.40	0.50
2	1	6	0.76	0.46	0.50	0.65	0.70	0.53	0.25	0.35	0.72	0.51	0.46	0.51
3	2	не обмеж.	0.81	0.35	0.51	0.44	0.76	0.27	0.22	0.54	0.78	0.36	0.37	0.51
4	2	6	0.82	0.52	0.55	0.48	0.85	0.38	0.28	0.44	0.84	0.52	0.46	0.50

Результати експериментів щодо відбору методів агрегації для множини рішень моделі класифікації статі та віку в рамках історії одного об'єкта, що відстежується, у відеопотоці представлені в табл. 3.4.

Таблиця 3.4

Результати експериментів щодо відбору методів агрегації множини рішень моделі класифікації статі та віку

№	Функція	Precision		Recall		F1	
		стать	серед. вік	стать	серед. вік	стать	серед. вік
1	mean	0,79	0,77	0,89	0,63	0,84	0,69
2	mode	0,85	0,79	0,94	0,71	0,89	0,74
3	w. mean c=0.1	0,82	0,76	0,94	0,66	0,87	0,71
4	w. mean c=0.3	0,83	0,76	0,94	0,65	0,88	0,70
5	w. mean c=0.5	0,82	0,77	0,94	0,63	0,87	0,70
6	w. mean c=0.7	0,82	0,78	0,94	0,68	0,87	0,72
7	w. mean c=0.9	0,82	0,78	0,94	0,69	0,87	0,73
8	w. mean c=1.1	0,82	0,78	0,94	0,69	0,87	0,73

За результатами табл. 3.4 видно, що найкращим методом агрегації множини рішень моделі класифікації для одного об'єкта є вибір клас, який найбільш часто зустрічається. При цьому дозволяє поліпшити підсумкові значення метрик якості:

для розпізнавання статі – до 0,89 у міру F1, для розпізнавання віку – до 0,74 у міру F1, значно покращуючи показники метрик як Recall, так і Precision.

3.4.4. Випробування алгоритмів повторної ідентифікації. Завданням проведення експериментів є оцінка якості роботи розробленого алгоритму при виборі різних значень параметрів методів, що входять до нього, і вибір найкращих з них для використання у складі модуля ядра вихідної платформи.

Усі експерименти, проведені з розробленим алгоритмом створення індивідуальних графічних відбитків об'єктів для їхнього повторного розпізнавання розділені на 2 частини: експерименти з методами створення відбитків графічних ознак) та розрахунку дескрипторів схожості – експерименти над логічним компонентом Predictor та експерименти з методами агрегації кандидатів, відібраних з галереї за запитом (query) для ухвалення остаточного рішення про ідентифікатор об'єкта під час його повторного розпізнавання – експерименти над логічним компонентом Aggregator.

Результати експериментів із вибором значення регульовального параметра з метрики weighted mean представлені на рис. 3.6 , 3.7.

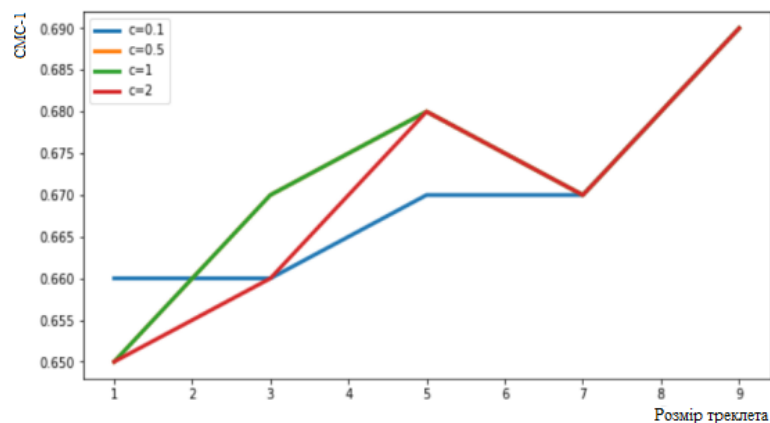


Рис. 3.6. Залежність значення метрики СМС-1 від розміру треклета при різних значеннях регульовального параметра c метрики агрегації weighted mean та кількості кандидатів, що дорівнює 1

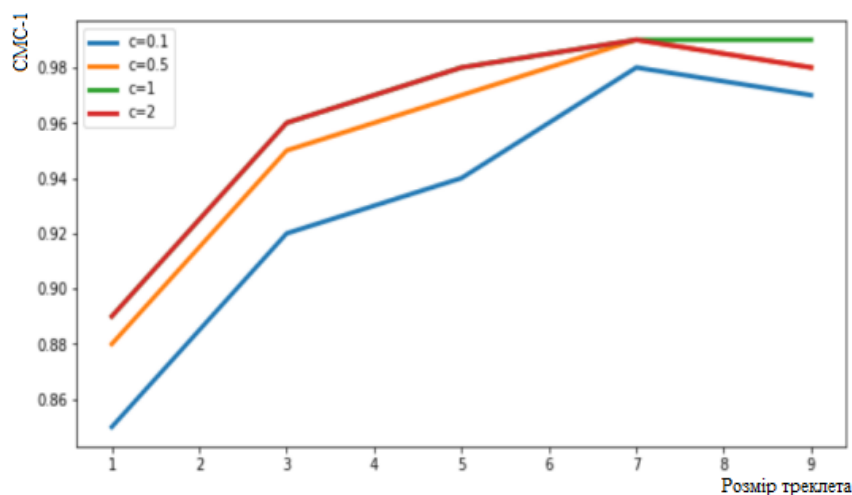


Рис. 3.7. Залежність значення метрики СМС-1 від розміру треклета за різних значень регулювального параметра c метрики агрегації weighted mean та кількості кандидатів, що дорівнює 4

З рис. 3.6, 3.7 видно, що зі збільшенням розміру треклета та кількості кандидатів, за якими проводиться агрегація, якість за метрикою СМС-1 зростає, при цьому, найкращої якості у всіх випадках дозволяє досягти метрика агрегації weighted mean при значенні регулювального параметра c , що дорівнює 1. Таким чином, за результатами експериментів як значення параметра з метрики агрегації weighted mean вибрано Значення 1.

Експерименти з відбором найкращих значень параметрів розміру треклету кількості кандидатів для агрегації та методів розрахунку метрик для агрегації були проведені з урахуванням налаштування параметра c методу weighted sum, та їх результати наведені на рис. 3.8, 3.9.

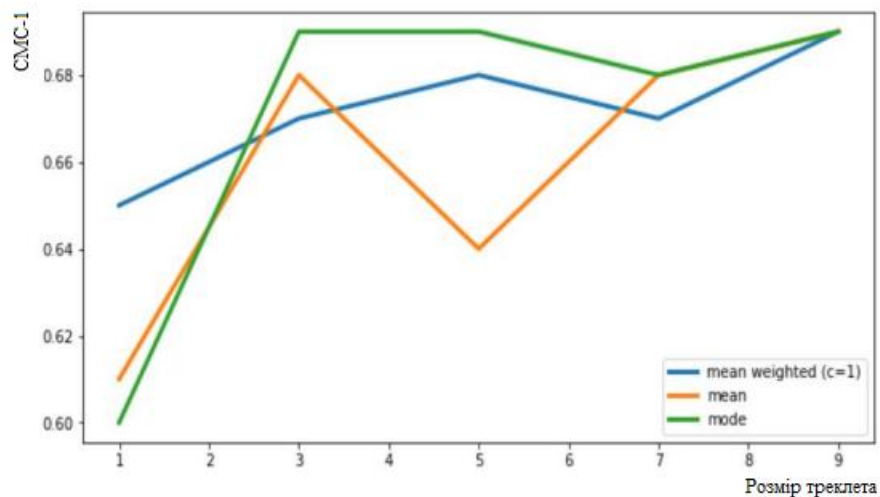


Рис. 3.8. Залежність значення метрики СМС-1 від розміру треклета під час використання різних метрик для агрегації та кількості кандидатів, що дорівнює 1

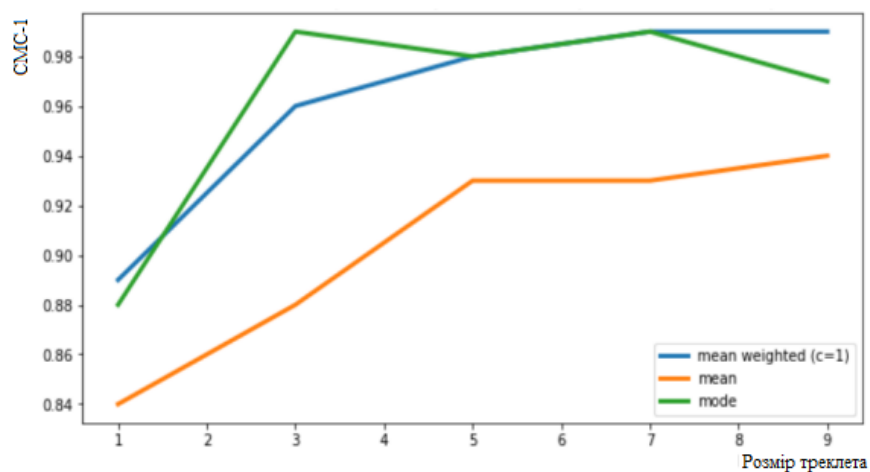


Рис. 3.9. Залежність значення метрики СМС-1 від розміру треклета під час використання різних метрик для агрегації та кількості кандидатів, що дорівнює 4

Виходячи з даних, котрі наведені на рис. 3.8, 3.9, можна зробити висновок, що метод weighted mean показує найкращі значення метрики якості СМС-1, що особливо помітно зі збільшенням як розміру треклета, і кількості кандидатів для агрегації. Оптимальний розмір треклета дорівнює 7, однак, зменшення значень метрики якості на великих розмірах можна пояснити розміром треклета, яким будуються графічні ознаки, відібраному під час проведення експериментів над компонентом Predictor.

За результатами всіх експериментів, проведених з компонентом Aggregator розробленого алгоритму створення графічних відбитків об'єктів для їх повторного розпізнавання, були відібрані найкращі щодо метрики якості СМС-1 значення гіперпараметрів: метрика методу агрегації weighted mean зі значенням регулювального параметра c , рівним 1; максимальна довжина треклету, що дорівнює 7; мінімальна кількість кандидатів для агрегації – 4. Конвеєр розроблених алгоритмів повторного розпізнавання з відібраними значеннями гіперпараметрів досягає значення метрики якості СМС-1, що дорівнює 0,98.

3.5. Висновок до розділу

У третьому розділі здійснено вибір програмного інструментарію, описана розробка програмних засобів, що реалізують алгоритми.

Також описані розробка та розмітка навчальних вибірок, і проведення випробувань алгоритмів.

Результати проведених експериментів дали змогу не тільки відібрати найкращі методи та їх гіперпараметри для роботи запропонованих алгоритмів, але й покращити загальну якість процесу відеоаналітики.

РОЗДІЛ 4

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Охорона праці

Метою кваліфікаційної роботи магістра є реалізація методів класифікації людей за статтю та віком та їх повторної ідентифікації у відеопотоці за допомогою технологій глибокого навчання. Оскільки, проведення робіт з розробки та використання методів передбачає застосування комп'ютерної техніки, зокрема ПК та периферійних пристроїв, то обов'язковим є дотримання вимог з охорони праці і техніки безпеки.

Для ефективної і безпечної роботи колективу працівників з розробки ПЗ комп'ютерних систем, в тому числі і фахівців з підвищення ефективності контролю доступу в приміщення, необхідно організувати безпечні умови праці. При цьому керівник організації несе безпосередню відповідальність за порушення нормативно-правових актів з охорони праці [74]. Окрім цього, на робочих місцях працівників необхідно забезпечити дотримання вимог, затверджених Наказом Мінсоцполітики від 14.02.2018 за № 207 «Про затвердження Вимог щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Згідно Вимог приміщення, де розміщені робочі місця операторів, крім приміщень, у яких розміщені робочі місця операторів великих ЕОМ загального призначення (сервер), мають бути оснащені системою автоматичної пожежної сигналізації відповідно до цих вимог;

– переліку однотипних за призначенням об'єктів, які підлягають обладнанню автоматичними установками пожежогасіння та пожежної сигналізації, затвердженого наказом Міністерства України з питань надзвичайних ситуацій та у справах захисту населення від наслідків Чорнобильської катастрофи від 22.08.2005 N 161, зареєстрованого в Міністерстві юстиції України 05.09.2005 за N 990/11270 (НАПБ Б.06.004-2005);

– Державних будівельних норм "Інженерне обладнання будинків і споруд. Пожежна автоматика будинків і споруд", затверджених наказом Держбуду України від 28.10.98 N 247 (далі - ДБН В.2.5-56:2014, з димовими пожежними сповіщувачами та переносними вуглекислотними вогнегасниками.

В інших приміщеннях допускається встановлювати теплові пожежні сповіщувачі. Приміщення, де розміщені робочі місця операторів, мають бути оснащені вогнегасниками, кількість яких визначається згідно з вимогами ДСТУ 4297:2004 «Пожежна техніка. Технічне обслуговування вогнегасників». Загальні технічні вимоги і з урахуванням граничнодопустимих концентрацій вогнегасної рідини відповідно до вимог НАПБ А.01.001-2014. Приміщення, в яких розміщуються робочі місця операторів сервера загального призначення, обладнуються системою автоматичної пожежної сигналізації та засобами пожежогасіння відповідно до вимог ДБН В.2.5-56:2014, НАПБ А.01.001-2014 і вимог нормативно-технічної та експлуатаційної документації виробника. Проходи до засобів пожежогасіння мають бути вільними.

Лінія електромережі для живлення комп'ютера та периферійних пристроїв повинні бути виконаними як окрема групова трипровідна мережа шляхом прокладання фазового, нульового робочого та нульового захисного провідників. Нульовий захисний провідник використовується для заземлення (занулення) електроприймачів. Не допускається використовувати нульовий робочий провідник як нульовий захисний провідник. Нульовий захисний провідник прокладається від стійки групового розподільного щита, розподільного пункту до розеток електроживлення. Не допускається підключати на щиті до одного контактного затискача нульовий робочий та нульовий захисний провідники.

Площа перерізу нульового робочого та нульового захисного провідника в груповій трипровідній мережі має бути не менше площі перерізу фазового провідника. Усі провідники мають відповідати номінальним параметрам мережі та навантаження, умовам навколишнього середовища, умовам розподілу провідників, температурному режиму та типам апаратури захисту, вимогам НПАОП 40.1-1.01-97.

У приміщенні, де одночасно експлуатуються понад п'ять комп'ютерів, на помітному, доступному місці встановлюється аварійний резервний вимикач, який може повністю вимкнути електричне живлення приміщення, крім освітлення. Комп'ютери повинні підключатися до електромережі тільки за допомогою справних штепсельних з'єднань і електророзеток заводського виготовлення.

У штепсельних з'єднаннях та електророзетках, крім контактів фазового та нульового робочого провідників, мають бути спеціальні контакти для підключення нульового захисного провідника. Їхня конструкція має бути такою, щоб приєднання нульового захисного провідника відбувалося раніше, ніж приєднання фазового та нульового робочого провідників. Порядок роз'єднання при відключенні має бути зворотним. Не допускається підключати комп'ютери до звичайної двопровідної електромережі, в тому числі – з використанням перехідних пристроїв. Електромережі штепсельних з'єднань та електророзеток для живлення комп'ютерної техніки повинні бути виконаними за магістральною схемою, по 3-6 з'єднань або електророзеток в одному колі. Штепсельні з'єднання та електророзетки для напруги 12 В та 42 В за своєю конструкцією мають відрізнятися від штепсельних з'єднань для напруги 127 В та 220 В. Штепсельні з'єднання та електророзетки, розраховані на напругу 12 В та 42 В, мають візуально (за кольором) відрізнятися від кольору штепсельних з'єднань, розрахованих на напругу 127 В та 220 В.

При підвищенні ефективності контролю доступу в приміщення, де для забезпечення безпеки мешканців, співробітників і збереження майна використовуються ДС, важливим, з точки зору охорони праці, є забезпечення достатньої величини природного та штучного освітлення, які визначені у НПАОП 0.00-7.15-18. Організація робочого місця фахівця із дослідження методів та програмно-апаратних засобів оптимізаційних процесів на основі ГА повинна забезпечувати відповідність усіх елементів робочого місця та їх розташування ергономічним вимогам ДСТУ 8604:2015 «Дизайн і ергономіка. Робоче місце для виконання робіт у положенні сидячи. Загальні ергономічні вимоги». Відстань від

екрана до ока фахівців, які працюють за комп'ютером визначається згідно з вимогами ДСанПіН 3.3.2.007-98.

Розміщення принтера або іншого пристрою введення-виведення інформації на робочому місці має забезпечувати добру видимість екрана комп'ютера, зручність ручного керування пристроєм введення-виведення інформації в зоні досяжності моторного поля згідно з вимогами ДСанПіН 3.3.2.007-98.

Таким чином, у результаті аналізу вимог щодо охорони праці користувачів комп'ютерів, визначено особливості організації робочих місць, вимог з електробезпеки, природного та штучного освітлення для ефективної і безпечної роботи фахівців розробки та реалізації методів класифікації людей за статтю та віком та їх повторної ідентифікації у відеопотоці за допомогою технологій глибокого навчання.

4.2. Вплив електромагнітного імпульсу ядерного вибуху на елементи виробництва та заходи захисту

У воєнний час при застосуванні ядерної зброї проти України на електронно-обчислювальне обладнання в першу чергу буде впливати електромагнітний імпульс (ЕМІ) ядерного вибуху у вигляді короткого імпульсу, який вражає головним чином електричну та електронну апаратуру. ЕМІ виникають в основному в результаті взаємодії гамма-випромінювання з атомами навколишнього середовища. На утворення ЕМІ йде невелика кількість ядерної енергії, але він здатен викликати високі імпульси струмів та напруг в кабелях повітряних і підземних ліній зв'язку, сигналізації, управління, електропередачі, в антенах радіостанцій. Вплив ЕМІ може привести до згорання чутливих електронних та електричних елементів, зв'язаних з великими антенами чи відкритими дротами, а також до порушень в обчислювальних пристроях. Вплив ЕМІ необхідно враховувати для всіх електричних та електронних систем [75].

Особливістю ЕМІ, як вражаючого фактору є його здатність розповсюджуватись на десятки і сотні кілометрів в оточуючому середовищі. Тому

ЕМІ може вплинути своєю дією на об'єкти, там де вибухова хвиля, світлове випромінювання, проникаюча радіація втрачають своє значення, як вражаючі фактори. При наземних та низьких повітряних вибухах в лініях зв'язку та електрозабезпечення виникають напруги, які можуть викликати пробій ізоляції провідників та кабелів відносно землі, пробій ізоляції елементів приладів підключених до повітряних і підземних ліній..

Найбільш піддані впливу ЕМІ системи зв'язку, сигналізації, управління. Використані в цих системах кабелі та апаратура мають обмежену електричну міцність не більше 10кВ імпульсної напруги, тоді як наведені імпульси напруги від ЕМІ можуть перевищувати ці значення. Найбільш піддана впливу ЕМІ апаратура виконана на напівпровідниках та інтегральних схемах, працюючих на малих струмах і напругах, і значить відчутних до впливу зовнішніх електричних і магнітних кіл, в тому числі і елементи програмного засобу для управління процесом міграції віртуальних машин в обчислювальній хмарі. ЕМІ пробиває ізоляцію, спалює елементи електричних схем радіоапаратури, викликає коротке замикання в радіопристроях, іонізацію діелектриків, змінює або повністю стирає магнітний запис. ЕМІ пошкоджує також резистори, викликає іскріння в їх міжконтактних з'єднаннях і деяких областях провідної поверхні. Найбільшу небезпеку ЕМІ представляє для апаратури, яка встановлена в особливо міцних спорудах, які витримують великі тиски ударної хвилі. В цих спорудах апаратура не виходить з ладу від механічних пошкоджень, але ЕМІ може вивести з ладу всю незахищену апаратуру системи зв'язку, сигналізації і керування. Найбільших значень досягають напруги, які наводяться між кабелем і землею.

Розглянемо можливі шляхи рішення задачі захисту від ЕМІ. Ідеальним захистом від ЕМІ виявилось б повне укриття приміщення, в якому розміщена радіоелектронна апаратура, металевим екраном [75]. Водночас зрозуміло, що практично забезпечити такий захист у ряді випадків неможливо, тому що для роботи апаратури часто потрібно забезпечити її електричний зв'язок із зовнішніми пристроями. Тому використовуються менш надійні засоби захисту, такі, як струмопровідні сітки, або плівкові покриття для вікон, щільникові металеві

конструкції для повітрезабірників і вентиляційних отворів і контактні пружинні прокладки, розміщені по периметру дверей і люків.

Більш складною технічною проблемою рахується захист від проникнення ЕМІ в апаратуру через різноманітні кабельні входи. Радикальним рішенням даної проблеми міг би стати перехід від електричних мереж зв'язку до практично не схильних до впливу ЕМІ волоконно-оптичних. Проте заміна напівпровідникових приладів у всьому спектрі виконуваних ними функцій електронно-оптичними пристроями можлива тільки у віддаленому майбутньому. Тому в даний час в якості засобів захисту кабельних входів найбільш широко використовуються фільтри, у тому числі волоконні, а також іскрові розрядники, металоокисні варистори і високошвидкісні зенеровські діоди [75].

Всі ці засоби мають як переваги, так і недоліки. Так, ємнісно-індуктивні фільтри достатньо ефективні для захисту від ЕМІ малої інтенсивності, волоконні фільтри захищають у відносно вузькому діапазоні надвисоких частот. Іскрові розрядники мають значну інерційність й в основному придатні для захисту від перевантажень, що виникають під впливом напруг і струмів, що наводяться в обшивці літака, кожусі апаратури й оплетенні кабеля.

Металоокисні варистори є напівпровідниковими приладами, що різко підвищують свою провідність при високій напрузі. Проте, при застосуванні цих приладів у якості засобів захисту від ЕМІ варто враховувати їх недостатньо високу швидкодію і погіршення характеристик при кількаразовому впливі навантажень. Ці недоліки відсутні у високошвидкісних зенеровських діодах, дія яких заснована на різкій лавиноподібній зміні опору від високого значення практично до нуля, при перевищенні прикладеної до них напруги граничного розміру. Крім того на відміну від варисторів характеристики зенеровських діодів після багатократних впливів високих напруг і переключень режимів не погіршуються.

Найбільш раціональним підходом до проектування засобів захисту від ЕМІ кабельних входів є створення таких роз'ємів у конструкції яких передбачені спеціальні заходи, що забезпечують формування елементів фільтрів і установку вмонтованих зенеровських діодів. Подібне рішення сприяє одержанню дуже малих

значень ємності й індуктивності, що необхідно для забезпечення захисту від імпульсів, що мають незначну тривалість і, отже, потужну високочастотну складову. Використання роз'ємів подібної конструкції дозволить вирішити проблему обмеження малогабаритних характеристик пристрою захисту. Складність рішення задачі захисту від ЕМП і висока вартість розроблених для цього засобів і методів змушують піти по шляху їхнього вибіркового застосування в особливо важливих системах зброї і військової техніки. Такий же шлях обраний і для захисту систем, що мають велику протяжність, керування і зв'язку. Проте основним методом вирішення проблеми спеціалісти вважають створення так званих розподілених мереж зв'язку [75].

4.3. Висновки до четвертого розділу

В цьому розділі проаналізовано важливі питання охорони праці та безпеки в надзвичайних ситуаціях, висвітлено питання впливу електромагнітного імпульсу ядерного вибуху на елементи виробництва та заходи захисту.

ВИСНОВКИ

Під час виконання кваліфікаційної роботи отримано такі результати:

- розроблено власну версію алгоритму розпізнавання статі та віку людей на кадрах у відеопотоці, що використовує ознаки силуету та руху. Запропонований конвеєр методів забезпечує якість розпізнавання статі, що дорівнює 0,89 для міри F1 і якість розпізнавання категорії віку, що дорівнює 0,74 для міри F1;
- розроблено власну версію алгоритмів створення індивідуальних відбитків об'єктів за графічними ознаками для повторного розпізнавання, працює з даними, наближеними до найбільш загальних умов зйомки. Реалізований підхід забезпечує якість вирішення завдання повторного розпізнавання об'єктів, що дорівнює 0,98 за метрикою СМС з рангом 1;
- алгоритми розпізнавання статі та віку, а також створення індивідуальних відбитків об'єктів за графічними ознаками для повторного розпізнавання впроваджено в роботу платформи, що забезпечує детектування та трекінг об'єктів;
- виконано розробку розширеної навчальної вибірки, що дозволила створити алгоритми на основі глибокого навчання, що враховують особливості даних відеозаписів у загальних умовах зйомки;
- розроблено програму та методику випробувань модуля ядра платформи з впровадженими у його роботу новими функціями. Проведення випробувань дозволило як відібрати найкращі методи та їх параметри для функціонування розроблених алгоритмів, так і покращити загальну якість всього конвеєра відеоаналітики, реалізованого у платформі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Луцик Н. С., Луцків А. М., Осухівська Г. М., Тиш Є. В. Методичні рекомендації до виконання кваліфікаційної роботи магістра для студентів спеціальності 123 «Комп'ютерна інженерія» другого (магістерського) рівня вищої освіти усіх форм навчання. Тернопіль. ТНТУ. 2024. 44 с.
2. Луцик Н.С., Луцків А.М., Осухівська Г.М., Тиш Є.В. Програма та методичні рекомендації з проходження практики за тематикою кваліфікаційної роботи для студентів спеціальності 123 «Комп'ютерна інженерія» другого (магістерського) рівня вищої освіти усіх форм навчання. Тернопіль. ТНТУ. 2024. 45 с.
3. Варавін А.В., Лещишин Ю.З., Чайковський А.В. Методичні вказівки до виконання курсового проєкту з дисципліни «Дослідження і проєктування комп'ютерних систем та мереж» для здобувачів другого (магістерського) рівня вищої освіти спеціальності 123 «Комп'ютерна інженерія» усіх форм навчання. Тернопіль. ТНТУ, 2024. 32 с.
4. Мишаківський Я.О. Методи розпізнавання статі та віку за зображенням особи // Інформаційні моделі, системи та технології: Праці XII наук.-техн. конф. (Тернопіль, 18-19 грудня 2024 р.). с. 138.
5. Liu W. A survey of deep neural network architectures and their applications / W. Liu, Z. Wang, X. Liu [et al.] // Neurocomputing. 2017. Vol. 234. Pp. 11-26.
6. Gu W. Facial expression recognition using radial encoding of local Gabor features and classifier synthesis / W. Gu, C. Xiang, Y. Venkatesh // Pattern Recognition. 2012. Vol. 45(1). Pp. 80-91.
7. Serrano S. Recent advances in face biometrics with Gabor wavelets: A review / S. Serrano, I. Diego, C. Conde // Pattern Recognition Letters. 2010. Vol. 31(5). Pp. 372-381.
8. Chandrappa D.N. Face Detection Using a Boosted Cascade of Features Using OpenCV / D.N. Chandrappa, G. Akshay, M. Ravishankar // Wireless Networks and Computational Intelligence (ICIP 2012). 2012. Vol. 292. Pp. 399-404.

9. Serrano S. Analysis of variance of Gabor filter banks parameters for optimal face recognition / S. Serrano, I. Diego, C. Conde [et al.]. // Pattern Recognition Letters. 2011. Vol. 31(15). Pp. 1998-2008.
10. Chen J. WLD: a robust local image descriptor / J. Chen, S. Shan, C. He, [et al.]. // IEEE Trans. Pattern Anal. Mach. Intell. 2010. Vol. 32(9) Pp. 1705–1720.
11. Микитишин А. Г., Митник М. М., Стухляк П. Д., Пасічник В. В. Комп'ютерні мережі. Книга 1 [навчальний посібник]. Львів : «Магнолія 2006», 2013. 256 с.
12. Микитишин А. Г., Митник М. М., Стухляк П. Д., Пасічник В. В. Комп'ютерні мережі. Книга 2. [навчальний посібник]. Львів : "Магнолія 2006", 2014. 312 с.
13. What is a LiDAR scanner, the iPhone 12 Pro's camera upgrade, anyway. // URL: <https://www.techradar.com/news/what-is-a-lidar-scanner-the-iphone-12-prosumored-camera-upgrade-anyway> (Дата звертання 21.11.2024).
14. Kevin W. A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition / W. Kevin, C. Kyong, F. Patrick // Computer Vision and Image Understanding. 2006. Vol. 101(1) Pp. 1-15.
15. Blanz V. Face Recognition Based on Fitting a 3D Morphable Model / V. Blanz, T. Vetter // IEEE Trans. Pattern Anal. Mach. Intell. 2003. Vol. 25(9) Pp. 1063–1074.
16. Zhang L. Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics / L. Zhang, D. Samaras // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006. Vol. 28(3) Pp. 351-363.
17. Passalis G. Using Facial Symmetry to Handle Pose Variations in Real-World 3D Face Recognition / G. Passalis, P. Panagiotis, T. Theoharis [et al.]. // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2011. Vol. 33(10) Pp. 1938–1951.
18. Zakariya Q. Deep Convolutional Neural Network for Age Estimation based on VGGFace Model / Q. Zakariya, A. Arafat, D. Buket [et al.] // ArXiv. 2017. 1709.01664.

19. Chen S. Using Ranking-CNN for Age Estimation / S. Chen, C. Zhang, M. Dong [at al.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. Pp. 742-751.
20. Liu X. AgeNet: Deeply Learned Regressor and Classifier for Robust Apparent Age Estimation / X. Liu, S. Li, M. Kan [at al.] // 015 IEEE International Conference on Computer Vision Workshop (ICCVW). 2015. Pp. 258-266.
21. Agbo-Ajala O. Deep learning approach for facial age classification: a survey of the state-of-the-art / O. Agbo-Ajala, S. Viriri // Artificial Intelligence Review. 2021. Pp. 179–213.
22. Antipov G. Apparent Age Estimation from Face Images Combining General and Children-Specialized Deep Learning Models / G. Antipov, M. Baccouche, S. Berrani [at al.] // 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2016. Pp. 801-809.
23. Makihara Y. Gait-based age estimation using a whole-generation gait database / Y. Makihara, M. Okumura, H. Iwama [at all] // Multimedia Tools and Applications. 2018. Vol. 77. Pp. 28333–28354
24. Li X. Gait-based human age estimation using age group-dependent manifold learning and regression / X. Li, Y. Makihara, C. Xu [at al.] // Multimedia Tools and Applications. 2018. Vol. 77. Pp. 28333–28354.
25. El-Alfy E.M. Silhouette-Based Gender Recognition in Smart Environments Using Fuzzy Local Binary Patterns and Support Vector Machines / E. M. El-Alfy, A. G. Binsaadoon // Procedia Computer Science. 2017. Vol.109. Pp. 164-171.
26. Kitchat K. Gender classification from gait silhouette using observation angle-based GEIs / K. Kitcha, N. Khamsemanan, C. Nattee // 2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM). 2018. Pp. 485-490.
27. Mansouri N. Gait-based human age classification using silhouette model / N. Mansouri, A.I. Mohammed, B.J. Yousra // IET Biom. 2018. Vol. 7(2). Pp. 116-124.

28. Baek N. Multimodal Camera-Based Gender Recognition Using Human-Body Image With Two-Step Reconstruction Network / N. Baek, S. W. Cho, J. H. Koo [at al.] // IEEE Access. 2017. Vol. 7. Pp. 104025-104044.
29. Jung C.R. Efficient Background Subtraction and Shadow Removal for Monochromatic Video Sequences / C. R. Jung // IEEE Transactions on Multimedia. 2009. Vol. 11(3). Pp. 571-577.
30. Braham M. Deep background subtraction with scene-specific convolutional neural networks / M. Braham, M.V. Droogenbroeck // 2016 International Conference on Systems, Signals and Image Processing (IWSSIP). 2016. Pp. 1-4.
31. Wang Y. CDnet 2014: an expanded change detection benchmark dataset / Y. Wang, P. Jodoin, F. Porikli [et al.] // 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2014. Pp. 393-400.
32. Bouwmans T. Deep Neural Network Concepts for Background Subtraction: A Systematic Review and Comparative Evaluation / T. Bouwmans, S. Javed, M. Sultana [at al.] // Neural Networks. 2019. Vol. 117. Pp. 8-66.
33. Dedeoğlu Y. Silhouette-Based Method for Object Classification and Human Action Recognition in Video / Y. Dedeoğlu, B.U. Töreyn, U. Güdükbay [at al.] // Lecture Notes in Computer Science. 2006. Vol. 3979. Pp. 64-77.
34. Elharrouss O., Gait recognition for person re-identification / O. Elharrouss, N. Almaadeed, S. Al-Maadeed [et al.] // The Journal of Supercomputing. 2021. Vol. 7. Pp. 3653- 3672.
35. Zheng S. Robust view transformation model for gait recognition / S. Zheng, J. Zhang, K. Huang [at al.] // 2011 18th IEEE International Conference on Image Processing. 2011. Pp. 2073-2076.
36. Iwama H. The OU-ISIR Gait Database Comprising the Large Population Dataset and Performance Evaluation of Gait Recognition / H. Iwama, M. Okumura, Y. Makihara [at al.] // IEEE Transactions on Information Forensics and Security. 2012. Vol. 7(5). Pp. 1511-1521.
37. Takemura N. Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition / N. Takemura, Y. Makihara, D. Muramatsu

[at al.] // IPSJ Transactions on Computer Vision and Applications. 2018. Vol. 10(4). Pp. 1-14.

38. Sivapalan S. The Backfilled GEI A Cross-Capture Modality Gait Feature for Frontal and Side-View Gait Recognition / S. Sivapalan, D. Chen, S. Denman [at al.] // 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA). 2012. Pp. 1- 8.

39. Liu T. Combining Convolutional Neural Network and Support Vector Machine for Gait-based Gender Recognition / T. Liu, X. Ye, B. Sun // 2018 Chinese Automation Congress (CAC). 2018. Pp. 3477-3481.

40. Li X. Object Re-Identification Based on Deep Learning / X. Li, Z. Zhou // Visual Object Tracking with Deep Neural Networks. 2019.

41. What is reidentification? // KLAP Augmenting fan experience for the sport industry. 2021. URL: <https://www.getklap.com/blog/what-is-reidentification> (Дата звертання 23.11.2024).

42. Ristani E. Features for Multi-Target Multi-Camera Tracking and Re-Identification / E. Ristani, C. Tomasi. // ArXiv 2018. 1803.10859.

43. Person Re-Identification // Papers With Code. 2022. URL: <https://paperswithcode.com/task/person-re-identification> (Дата звертання 24.11.2024).

44. Velychko D., Osukhivska H., Palaniza Y., Lutsyk N., Sobaszek L. Artificial Intelligence Based Emergency Identification Computer System. Advances in Science and Technology Research Journal, 18 no. 2, 2024, P. 296-304.

45. Person Re-Identification Shaogang Gong [Electronic resource] / Shaogang Gong [et al.]. Springer. 2014. 444 p.

46. Zhou K. Learning Generalisable Omni-Scale Representations for Person ReIdentification / K. Zhou, Y. Yang, A. Cavallaro [at al.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021. Pp. 1-1.

47. Cai Y. Exploring context information for inter-camera multiple target tracking / Y. Cai, G. Medioni. // IEEE Winter Conference on Applications of Computer Vision. 2014. Pp. 761-768.

48. Das A. Consistent Re-identification in a Camera Network / A. Das, A. Chakraborty, A.K. Roy-Chowdhury // Lecture Notes in Computer Science. 2014. Vol. 8690 Pp. 330-345).
49. Chen, X. Multitarget Tracking in Nonoverlapping Cameras Using a Reference Set / X. Chen, L. An, B. Bhanu // IEEE Sensors Journal. 2015. Pp. 2692-2704.
50. Skorenkyy Y., Kozak R., Zagorodna N., Kramar O., Baran, I. Use of augmented reality-enabled prototyping of cyber-physical systems for improving cybersecurity education. Journal of Physics: Conference Series. 2021. Vol. 1840, No. 1. DOI: <http://dx.doi.org/10.1088/1742-6596/1840/1/012026>
51. Martinel N. Saliency Weighted Features for Person Re-Identification / N. Martinel, C. Micheloni, G. L. Foresti. // European Conference on Computer Vision. 2015. Pp. 191–208.
52. Palamar A., Palamar M., Osukhivska H. Real-time Health Monitoring Computer System Based on Internet of Medical Things. CEUR Workshop Proceedings, 3rd International Workshop on Information Technologies: Theoretical and Applied Problems (ITTAP 2023), Ternopil, Ukraine, Opole, Poland, November 22–24, 2023. Vol. 3628. P. 106-115.
53. Ma X. Person Re-Identification by Unsupervised Video Matching / X. Ma, X. Zhu, S. Gong [et al.]. // ArXiv. 2016. 1611.08512v3.
54. Chen Y., Person Re-identification using group context / Y. Chen, S. Duffner, A. Stoia [et al.] // HAL. 2018. hal-01895373.
55. Zhou K. Omni-Scale Feature Learning for Person Re-Identification // K. Zhou, Y. Yang, A. Cavallaro [et al.] // ArXiv. 2019. 1905.00953.
56. Quan R. Auto-ReID: Searching for a Part-Aware ConvNet for Person ReIdentification: materials from ICCV 2019 / R. Quan, X. Dong, Y. Wu [et al.] // ArXiv 2019. 1903.09776.
57. Zheng L. Person Re-identification: Past, Present and Future / L. Zheng, Y. Yang, A. G. Hauptmann. // ArXiv. 2016. 1610.02984v1.
58. Zhong Z. Unlabeled Samples Generated by GAN Improve the Person Re-identification Baseline in vitro / Z. Zheng, L. Zheng, Y. Yang. // ArXiv 2017. 1701.07717.

59. Zhong Z. Random Erasing Data Augmentation / Z. Zhong, L. Zheng, G. Kang [et al.]. // ArXiv 2017. 1708.04896v2.
60. Yatsyshyn V., Pastukh O., Palamar A., Zharovskyy R. Technology of relational database management systems performance evaluation during computer systems design. Scientific Journal of TNTU, Ternopil, Ukraine, 2023. Vol. 109, No 1. P. 54–65.
61. Zheng Z. Pedestrian Alignment Network for Large-scale Person Re-identification / Z. Zheng, L. Zheng, Y. Yang. // ArXiv. 2017. 1707.00408.
62. Hermans A. In Defense of the Triplet Loss for Person Re-Identification: materials from CVPR / A. Hermans, L. Beyer, B. Leibe. // ArXiv. 2017. 1703.07737.
63. Triplet Loss and Online Triplet Mining in TensorFlow // Olivier Moindrot blog. GitHub. 2018. URL: <https://omoindrot.github.io/triplet-loss> (Дата звертання 29.11.2024).
64. He T. Dense Interaction Learning for Video-based Person Re-identification / T. He, X. Jin, X. Shen [et al.]. // ArXiv. 2021. 2103.09013v3.
65. Wiecek M. On the Unreasonable Effectiveness of Centroids in Image Retrieval / M. Wiecek, B. Rychalska, J. Dabrowski. // ArXiv. 2021. 2104.13643v1.
66. Varior R.R. A Siamese Long Short-Term Memory Architecture for Human ReIdentification: materials from CVPR / R.R. Varior, B. Shuai, J. Lu [et al.] // ArXiv. 2016. 1607.08381v1.
67. Zheng L. Person Re-identification: Past, Present and Future / L. Zheng, Y. Yang, A. G. Hauptmann. // ArXiv. 2016. 1610.02984v1.
68. Deep Learning 018 The re identification problem in computer vision // URL: https://www.youtube.com/watch?v=_DqF4233m5Y&list=WL&index=13&t=2s (Дата звертання 30.11.2024).
69. Tan S. Dense invariant feature based support vector ranking for person reidentification / S. Tan, F. Zheng, L. Shao. // IEEE Transactions on Circuits and Systems for Video Technology. 2015. Pp. 687-691.
70. Yin J. Unsupervised Person Re-identification via Simultaneous Clustering and Consistency Learning / J. Yin, J. Qiu, S. Zhang [et al.]. // ArXiv. 2021. 2104.00202v1.

71. Li M. Unsupervised Person Re-identification by Deep Learning Tracklet Association / M. Li, X. Zhu, S. Gong. // ECCV 2018. Pp. 772–788.
72. Munjal B. Query-guided End-to-End Person Search / B. Munjal, S. Amin, F. Tombari [et al.]. // ArXiv. 2019. 1905.01203.
73. Мишаківський Я.О. Методи для реалізації алгоритму розпізнавання статі та віку людини у відеопотоці даних // Інформаційні моделі, системи та технології: Праці XII наук.-техн. конф. (Тернопіль, ТНТУ ім. І. Пулюя, 18-19 грудня 2024 р.) с. 140.
74. Зеркалов Д.В. Охорона праці в галузі: Загальні вимоги. Навчальний посібник. К.: Основа. 2011. 551 с.
75. Безпека в надзвичайних ситуаціях. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної та заочної (дистанційної) форм навчання / укл.: Стручок В. С. Тернопіль: ФОП Паляниця В. А., 2022. 156 с.

ДОДАТОК А

Тези конференції

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
імені ІВАНА ПУЛЮЯ
НАУКОВЕ ТОВАРИСТВО ім. ШЕВЧЕНКА

МАТЕРІАЛИ

ХІІ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

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



18-19 грудня 2024 року

ТЕРНОПІЛЬ
2024

І. О. Кухар, Г. М. Осухівська	я
ІНТЕГРАЦІЯ ВІДНОВЛЮВАЛЬНИХ ДЖЕРЕЛ ЕНЕРГІЇ В РОЗУМНІ ЕЛЕКТРИЧНІ МЕРЕЖІ З ВИКОРИСТАННЯМ ПЕРЕДОВИХ ТЕХНОЛОГІЙ	
I. O. Kulchar, H. Osukhivska	
INTEGRATION OF RENEWABLE ENERGY SOURCES INTO SMART GRIDS USING ADVANCED TECHNOLOGIES	135:
Ю. Лещишин, А. Герасименко, О. Герасименко	я
КОМП'ЮТЕРНА СИСТЕМА МОНІТОРИНГУ РАДІОЗВ'ЯЗКУ ВІД БПЛА	
Yu. Leshchynshyn, A. Herasymenko, O. Herasymenko	
COMPUTER SYSTEM FOR MONITORING RADIO COMMUNICATIONS FROM UAVS	136:
А. Луцків, А. Люлька	я
ЗАСТОСУВАННЯ АЛГОРИТМІВ БАЛАНСУВАННЯ НАВАНТАЖЕННЯ В ПРОЦЕСАХ СУЧАСНИХ ОПЕРАЦІЙНИХ СИСТЕМ	
A. Lutskiv, A. Liulka	
APPLICATION OF LOAD BALANCING ALGORITHMS IN THE PROCESSES OF MODERN OPERATING SYSTEMS	137:
Я. О. Мышаківський	я
МЕТОДИ РОЗПІЗНАВАННЯ СТАТІ ТА ВІКУ ЗА ЗОБРАЖЕННЯМ ОСОБИ	
Ia. O. Myshakivskiy	
METHODS OF GENDER AND AGE RECOGNITION FROM THE IMAGE OF A PERSON	138:
Я. О. Мышаківський	я
МЕТОДИ ДЛЯ РЕАЛІЗАЦІЇ АЛГОРИТМУ РОЗПІЗНАВАННЯ СТАТІ ТА ВІКУ ЛЮДИНИ У ВІДЕОПОТОЦІ ДАНИХ	
Ia. O. Myshakivskiy	
METHODS FOR IMPLEMENTING AN ALGORITHM FOR RECOGNITION OF A PERSON'S GENDER AND AGE IN A VIDEO DATA STREAM	140:
М. Є. Олійник, І. В. Мудрий, Н. С. Луцьк	я
МЕТОДИ ТА ЗАСОБИ ОПТИМАЛЬНОГО РОЗПОДІЛУ ЗАВДАНЬ В КОМП'ЮТЕРИЗОВАНИЙ СИСТЕМІ БЕЗПЛОТНОЇ ДОСТАВКИ	
M. Y. Oliynyk, I. V. Mudryi, N. S. Lutsyk	
METHODS AND TOOLS FOR OPTIMAL TASK ALLOCATION IN A COMPUTERIZED UNMANNED DELIVERY SYSTEM	141:
Н. Сороківська, В. Яцишин	я
ВИКОРИСТАННЯ АДАПТОВАНИХ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ІДЕНТИФІКАЦІЇ ПТАХІВ У ПРИРОДНИХ УМОВАХ	
N. Sorokivska, V. Yatsyshyn	
APPLICATION OF ADAPTED ARTIFICIAL INTELLIGENCE MODELS FOR BIRD IDENTIFICATION IN NATURAL ENVIRONMENTS	142:
А. А. Станько, А. В. Гончаренко, І. В. Журник	я
РОЗУМНІ МІСТА: ІНТЕГРАЦІЯ ТЕХНОЛОГІЙ, ДАНИХ І СТРАТЕГІЇ СТАЛОГО РОЗВИТКУ МІСЬКОЇ ЕКОСИСТЕМИ	
A. A. Stanko, A. V. Honcharenko, I. V. Zhuryk	
SMART CITIES: INTEGRATING TECHNOLOGIES, DATA AND STRATEGIES FOR SUSTAINABLE URBAN ECOSYSTEM DEVELOPMENT	143:
А. А. Станько, С. З. Кульчицький, Ю. В. Срібний	я
ФОРМУВАННЯ КЕРОВАНИХ ОЗЕР ДАНИХ: МЕТОДОЛОГІЇ, ІНСТРУМЕНТИ ТА ПРАКТИЧНІ АСПЕКТИ	
A. A. Stanko, S. Z. Kulchytskyi, Y. V. Sribnyi	
FORMATION OF MANAGED DATA LAKES: METHODOLOGIES, TOOLS AND PRACTICAL ASPECTS	144:
А. Чуба, В. Тимошук, А. Микитишин, В. Шиманська	я
СУЧАСНІ ПІДХОДИ ДО МОНІТОРИНГУ ТА УПРАВЛІННЯ В DOCSIS-МЕРЕЖАХ	
A. Chuba, V. Tymoshchuk, A. Mykytyshyn, V. Shymanska	
MODERN APPROACHES TO MONITORING AND MANAGEMENT IN DOCSIS NETWORKS	145:

МЕТОДИ ДЛЯ РЕАЛІЗАЦІЇ АЛГОРИТМУ РОЗПІЗНАВАННЯ СТАТІ ТА ВІКУ ЛЮДИНИ У ВІДЕОПОТОЦІ ДАНИХ

Ia.O.Myshakivskiy

METHODS FOR IMPLEMENTING AN ALGORITHM FOR RECOGNITION OF A PERSON'S GENDER AND AGE IN A VIDEO DATA STREAM

Провівши аналіз існуючих на сьогоднішній день методів розпізнавання статі та віку людини, які не використовують зображення особи, випущення ознак з треклету об'єкта було обрано алгоритм побудови GEI (Gait Energy Image – енергетичне зображення ходи), який вимагає проведення етапу сегментації та безпосередньо створення самого зображення (мапи ознак). Вибір на користь саме цього підходу було зроблено завдяки ефективності, показаний у кількох успішних роботах та відповідність підходу до особливостей даних, що завантажуються на платформу – відеодані переміщення людей у певному напрямку (людина протягом свого треклету рухається в одну сторону). Додатковою перевагою використання GEI-ознак є можливість побудови зображення навіть на основі кадрів, де видно лише силует людини (немає інформації про кольори, важко виділити ключові точки або текстурні ознаки) – такі дані у тому числі відносяться до загальних умов зйомки.

Для здійснення етапу сегментації об'єкта людини, необхідного для побудови GEI-зображення, було обрано рішення PaddleSeg. Дані та інструменти сегментації, що містять у своєму складі, включають зображення людей на повний зріст, на відміну від більшості інших, де зйомки здійснювалися під менш відповідним під дані платформи гострим кутом до горизонту. Завдяки цьому, моделі сегментації не потрібно навчати з нуля, для побудови моделі може використовуватись підхід перенесення навчання [1].

Завдання розпізнавання віку було вирішено не як завдання регресії та розрахунку точного числового значення, а як завдання класифікації та визначення однієї з трьох вікових категорій. Вибір на користь такого рішення було зроблено на підставі практично повної відсутності відкритих даних, умови зйомки яких наближені до даних, властивих платформі детектування та трекінгу об'єктів, і великий трудомісткості створення власного набору, що відповідає всім вимогам розробки алгоритму розпізнавання статі. Рішення підзадачі класифікації статі та віку було прийнято здійснювати за допомогою згорткової нейронної мережі, яка використовує для отримання ознак попередньо навчені ваги нейронної мережі esNet, що показує хороші результати у різних завданнях класифікації на зображення реального світу.

Розроблена версія алгоритму розпізнавання статі та віку буде враховувати особливості його впровадження в роботу ядра платформи детектування та трекінгу об'єктів (враховує доступність даних, їх структуру та формат, а також інтерфейси наявних програмних компонентів), а також можливості щодо подальшої оптимізації продуктивності для досягнення необхідних показників часу обробки відеопотоків.

Література

1. Liu, Yi & Chu, Lutao & Chen, Guowei & Wu, Zewu & Chen, Zeyu & Lai, Baohua & Hao, Yuying. (2021). PaddleSeg: A High-Efficient Development Toolkit for Image Segmentation. 10.48550/arXiv.2101.06175.]

МЕТОДИ РОЗПІЗНАВАННЯ СТАТІ ТА ВІКУ ЗА ЗОБРАЖЕННЯМ ОСОБИ

Ia.O.Myshakivskyi

METHODS OF GENDER AND AGE RECOGNITION FROM THE IMAGE OF A PERSON

Розпізнавання по обличчю є популярним і точним підходом до завдання визначення статі та віку по фото або відео. Причиною цього є присутність на обличчі всіх візуальних особливостей, які дозволяють визначати стать, вік, расу, настрій та багато інших рис людини, візуально видимі в основному (або виключно) на вигляд і подання особи. На відео розпізнавання по особі переважно для визначення статі та віку у випадках, коли на кадрах особа людини видно чітко. У послідовності дій для вирішення описаної задачі можна виділити три етапи: визначення людей на кадрах відеопотоку та побудова детекцій; детектування осіб; розпізнавання статі та віку по особі.

Виявлення особи (face detection) є досить відомим завданням, у якого є множина різних підходів. Для роботи системи необхідно обмежувати область пошуку особи. Це необхідно для можливості однозначного зіставлення виявленої особи з людиною. Так само це рішення дозволить прискорити розробку та загальну швидкість роботи системи, оскільки відсутня необхідність у написанні та тестуванні додаткового модуля в порівнянні обмежувачів вікон детектованих осіб та обмежувачів вікон піпходів. Нижче розписані підходи до задачі детектування особи на кадрі.

Класичні методи визначення осіб. Підходи, які широко використовувалися на початку 2000х років. Використовувалися, тому числі, класичні статистичні методи, аналіз головних компонентів, лінійні та нелінійні класифікатори. Також використовувалися методи зображення, що відображають на площину і за допомогою знаходження відстані Махаланобіса [1], знаходили ступінь близькості поточного виявленого об'єкта з референс вектор обличчя. Незважаючи на хорошу теоретичну базу для відповідних методів, вони майже не застосовувалися в реальних промислових застосунках. Проблемою була вкрай низька здатність подібних підходів адаптуватись до інваріантності середовища.

Методи, засновані на застосуванні нейронних мереж. Deep CNN добре себе зарекомендували в завданні детектування об'єктів на зображеннях. Вони здатні визначати характерні риси об'єкта безпосередньо з пікселів зображення [2]. До найвідоміших архітектур CNN можна віднести AlexNet, GoogLeNet, VGGNet, ResNet. Для навчання CNN потрібен величезний обсяг даних, тому для застосування їх до вузькоспрямованих завдань, зазвичай, використовують передбачені на відомих датасетах моделі, а потім, застосовуючи метод переносу навчання (transfer learning), розраховують ваги на останніх шарах мережі. Підходи на основі CNN є найбільш точними і в той же час найефективнішими до обчислювальних потужностей.

Методи, засновані на вейвлетах Хаара та Габора. За допомогою вейвлетів із зображення отримували найбільш сильні сигнали, фактично виділяючи найважливіші ознаки на зображенні, відкидаючи інше. Вейвлети Габора були представлені в досліджених алгоритмах розпізнавання осіб. Різні вейвлети застосовувалися з ідеєю виділення візуальних ознак незалежно від завдання. Надалі дані візуальні ознаки використовувалися в авторських алгоритмах, у тому числі і для знаходження особи, як, наприклад, в алгоритмі Біюлі-Джонса. Використання ознак, отриманих вейвлетом Габора, є складним та часо затратним. Існуючі спрощені вейвлети та алгоритми на їх

основі працюють швидше, проте сильно залежать від освітлення та інших зовнішніх факторів.

Методи, що ґрунтуються на знаходженні дескрипторів особи. Замість використання всього зображення алгоритми розбивають його на підобласті, у кожній з яких розраховують гістограми і знаходять найбільш яскраві та важливі ознаки для вирішуваної на даний момент задачі.

Надалі отримані ознаки можуть бути використані для завдання визначення особи, так і в задачі ідентифікації, якщо є приклад, для якого також можна отримати ознаки подібним способом. Наприклад, автори застосовували підхід на основі локальних дескрипторів, включаючи різні локальні бінарні шаблони, для завдання ідентифікації по обличчю. При цьому знаходилися такі дескриптори, які були стійкі до вікових змін особи. Дані методи компактні та швидко працюють, добре справляються із змінами у висвітленні та різними емоціями на обличчі. Однак ці підходи дуже чутливі до шуму на зображенні.

Методи, що базуються на тривимірних зображеннях та картах глибини. Багато сучасних камер, і навіть смартфонів, оснащені лідарами і мають можливість швидко та точно отримувати карту глибини зображення. Цей метод зйомки дозволяє значно знизити чутливість алгоритмів до освітлення та зовнішніх факторів зйомки. Моделі, що працюють із тривимірними зображеннями, набагато більш точні та інваріантні до змін ніж працюючі з двовимірними. Методи використовують текстури зображення, компенсують варіативне освітлення, а також компенсують різні пози людини за допомогою прагнення до симетричного обличчя особи під час аналізу. Такі системи, що вимагають особливого обладнання, ретельного калібрування та синхронізації даних, що робить їх складно застосовувати в системах загального призначення.

Визначення статі та віку. Якщо визначення статі є завданням бінарної класифікації, то завдання визначення віку можна вирішувати як мультикласову класифікацію за віковими кластерам, так і як завдання регресійного передбачення кількості років. Наприклад удля визначення віку автори донесли модель VGG-face методом перенесення навчання на наборі даних із зображеннями людей, для кожного з яких була мітка про вік. На вихідному шарі нейронної мережі було 8 виходів, відповідних до вікових груп. У для завдання визначення віку також використовувалися CNN (Ranking-CNN), але з бінарним виходом, а проставлення мітки про вік людини відбувалося агрегацією виходів "базових" нейромереж. Кожна з базових згорткових мереж була навчена на мітках приналежності людини до групи "старші" або "молодші" певного віку.

Якщо для завдання визначення особи на кадрі методи не на основі нейромереж мають досить хорошу точність і можуть скласти конкуренцію CNN, то для завдання визначення віку по особі згорткові мережі є неперевершеними лідерами. Так, у 2015 році було представлено мережу AgeNet, що демонструє високу точність у завданнях визначення віку по особі та посіла призове місце у змаганні «ChaLearn 2015 apparent age competition». Підхід є глибокою згортковою нейронною мережею, в якій як функція втрат використовуються як Евклідова норма, так і крос-ентропія. Більшість сучасних методів для визначення статі та віку використовують відомі архітектури глибоких згорткових нейромереж, таких як VGG-16 або ImageNet.

Література

1. Perlibakas V. Distance measures for PCA-based face recognition / V. Perlibakas // Pattern Recognition Letters. 2004. Vol. 25(6). Pp. 711–724.
2. Liu W. A survey of deep neural network architectures and their applications / W. Liu, Z. Wang, X. Liu [et al.] // Neurocomputing. 2017. Vol. 234. Pp. 11–26.