

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)

Кафедра Інформатики
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

**РОЗРОБКА ПРОГРАМНОГО ЗАСОБУ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ НА
ЗОБРАЖЕННЯХ**
(тема)

Виконав:
студент 4 курсу, групи ІТІНФ-18-2

Кушнір І.І.
(прізвище, ініціали)

Спеціальності 122 Комп'ютерні науки
(код і повна назва спеціальності)

Тип програми освітньо-професійна

Освітня програма Інформатика
(повна назва освітньої програми)

Керівник проф. Машталір В.П.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

Кобилін О.А.
(прізвище, ініціали)

2022 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)

Кафедра Інформатики
(повна назва)

Рівень вищої освіти перший (бакалаврський)

Спеціальність 122 Комп'ютерні науки
(код і повна назва)

Тип програми освітньо-професійна

Освітня програма Інформатика
(повна назва освітньої програми)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« ____ » _____ 2022 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Кушніру Іллі Ігоровичу
(прізвище, ім'я, по батькові)

1. Тема роботи Розробка програмного засобу для класифікації об'єктів на зображеннях

затверджена наказом університету від 16 травня 2022 року № 541Ст

2. Термін подання студентом роботи до екзаменаційної комісії 29 травня 2022 р.

3. Вихідні дані до роботи: Методи виділення облич на зображеннях, методи класифікації об'єктів, бібліотека OpenCV, мова програмування Python, фреймворк Tensorflow, нейронна мережа ResNet

4. Перелік питань, що потрібно опрацювати в роботі

1. Огляд методів виділення облич на зображеннях

2. Огляд сучасних програмних засобів класифікації об'єктів

3. Аналіз інформаційного забезпечення

4. Створення архітектури інформаційної системи

5. Програмна реалізація застосунку по розпізнаванню об'єктів на зображеннях з використанням ResNet та мови програмування Python

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) Ілюстрація етапів обробки зображень, приклади застосування фільтрації, схема тренування каскадів, ілюстрація процесу розпізнавання обличчя людини, приклади роботи бібліотеки OpenCV

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата
Консультант з дотримання діючих стандартів та норм	Доцент Белова Н.В.		

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	18.04.2022	
2	Аналіз завдання, підбір літератури	18.04.22-21.04.22	
3	Аналіз літератури з досліджуваної проблеми	22.04.22-25.04.22	
4	Аналіз технічних засобів	26.04.22-30.04.22	
5	Розробка інформаційної системи вибраної предметної області	01.05.22-14.05.22	
6	Програмна реалізація	15.05.22-23.05.22	
7	Оформлення пояснювальної записки	24.05.22-28.05.22	
8	Перевірка на плагіат	30.05.22	
9	Рецензування	29.05.22	
10	Підготовка презентації та доповіді	29.05.22-01.06.22	
11	Занесення роботи в електронний архів	02.06.22	
12	Попередній захист кваліфікаційної роботи	02.06.22	

Дата видачі завдання 18 квітня 2022 р.

Студент _____
(підпис)

Керівник роботи _____ проф. Мащталір В.П.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ/ABSTRACT

Пояснювальна записка до кваліфікаційної роботи: 68 с., 1 табл., 15 рис., 1 дод., 30 джерел.

КЛАСИФІКАЦІЯ ОБ'ЄКТІВ, ПЕРЕТВОРЕННЯ ХАФА, ПЕРЕТВОРЕННЯ РАДОНА, АЛГОРИТМИ МАШИНОГО НАВЧАННЯ, ВЕРОГІДНІСТНІ АЛГОРИТМИ, МОДЕЛІ ШТУЧНОГО ІНТЕЛЕКТУ, ОБРОБКА ДАНИХ, ЦІЛЬОВІ ФУНКЦІЇ, МАТЕМАТИЧНІ МОДЕЛІ.

Об'єктом роботи є процес обробки зображень для контекстної інтерпретації в системах технічного зору.

Метою роботи є отримання й обробка фото та відеоматеріалу, для виділення облич та зіставлення їх з наявною базою даних.

Робота присвячена застосуванню нейромережі на архітектурі ResNet з відповідними налаштуваннями для детектування та розпізнавання обличчя людини у відеопотоці з доповненою реальністю.

OBJECT CLASSIFICATION, HAF'S TRANSFORMATION, RADON'S TRANSFORMATION, MACHINE-LEARNING ALGORITHMS, PROBABILITY ALGORITHMS, ARTIFICIAL INTELLEGENCY MODELS, DATA SCIENCE, OBJECTIVE FUNCTION, MATHEMATICAL MODELS.

The object of the work is the image processing for contextual interpretation in technical vision systems.

The goal of the work is to obtain and process photos and videos, to highlight faces and compare them with the existing databases.

The work is devoted to the application of the neural network on the ResNet architecture with the appropriate settings for detecting and recognizing a person's face in an augmented reality video stream.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	6
Вступ	7
1 Аналіз проблемної області та постановка задачі	8
1.1 Аналіз проблеми, об'єкт та предмет дослідження	8
1.2 Огляд існуючих рішень, проблематика	12
1.3 Сучасні програмні засоби для класифікації об'єктів	16
1.4 Постановка задачі	22
2 Аналіз інформаційного забезпечення	23
2.1 Задачі та алгоритми	23
2.2 Перетворення Радона, перетворення Хафа, робастні оцінювачі..	26
2.3 Вивчення об'єктів за допомогою приблизного медіального фільтру та методу оптичного потоку	33
2.4 Розпізнавання та відслідковування об'єктів	36
2.5 Аналітика поведінки	39
2.6 Обґрунтування вибору програмних засобів	40
3 Розробка інформаційної системи, алгоритмічного та програмного забезпечення	44
3.1 Сучасні вимоги та потреби у класифікації об'єктів на зображеннях	44
3.2 Вимоги до програмного забезпечення	47
3.3 Програмні ресурси для створення інформаційної системи	53
3.4 Архітектура інформаційної системи	56
Висновки	62
Перелік джерел посилання	63
Додаток А Приклади розпізнавання різних об'єктів за допомогою OpenCV .	67

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

MIT – Massachusetts Institute of Technology (Массачусетський технологічний інститут)

CMU – Carnegie Mellon University (Університет Карнегі-Меллона)

FERET – Facial Recognition Technology (технологія розпізнавання облич)

MATLAB – пакет прикладних програм для вирішення задач технічних обчислень

МФ – медіанний фільтр

DFT – Discrete Fourier transform (дискретне перетворення Фур'є)

DCT – Discrete cosine transform (дискретне косинусне перетворення)

XML – Extensible Markup Language (розширювана мова розмітки)

UI – User interface (інтерфейс користувача)

SVM – Support-vector machine (опорно-векторна машина)

FLANN – Fast Approximate Nearest Neighbor Search Library (бібліотека швидкого пошуку найближчих сусідів)

EOM – машина для інженерних розрахунків

ML – Meta Language (метамова)

AI – Artificial intelligence (штучний інтелект)

GPU – Graphics processing unit (графічний процесор)

ВСТУП

Щодня люди розвиваються, а разом з розвитком приходять і нові технології, які спрощують наше життя. Якщо подивитися навколо, то стає зрозумілим, що всі створені нами предмети спрямовані на зручність та збереження такої цінності як час. Більшою мірою вже вдалося автоматизувати механічну працю (збирання деталей на фабриці, перевірка пропису текстів та інше), а зараз популярність набуває розробка аналітичних систем, що самі приймають рішення та розвиваються. Розробка програмного засобу для класифікації об'єктів на зображеннях – є одним з великих розділів цієї області.

Дана робота присвячена застосуванню нейромережі для розпізнавання об'єктів на фото та відеоматеріалах.

Для розробки програмного забезпечення для розпізнавання необхідно вивчити теоретичні відомості щодо алгоритмів, що використовуються технологіям і зробити огляд існуючих аналогів. А саме, навчання машинного зору, методи обробки та розпізнавання зображення, мова програмування, що використовується, і необхідні бібліотеки даних.

Я вважаю, що актуальність даної теми важко переоцінити. Адже останнім часом такі галузі як відеоспостереження, автомобільний контроль та медицина все більше і більше потребують вдосконалення даної технології та розвитку цього напрямку в цілому.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ

1.1 Аналіз проблеми, об'єкт та предмет роботи

В наш час фото і відеозображення використовуються у багатьох галузях і для виконання необхідних цілей: військових, наукових, розважальних, промислових, цивільних та медичних за допомогою камери телефону, відеокамери, радарів, сканерів та електронних мікроскопів. Для того, щоб машина змогла сприйняти інформацію з зображення воно має пройти цифрову обробку тому це дуже важливий етап без якого її подальша обробка буде неможливою.

З боку обчислюваної машини аналіз зображення має зовсім інший вигляд від звичайної фотографії, пристрій бачить тільки набір чисел у різній послідовності в залежності від стану і різновидності об'єкта. Цей процес називається оцифровкою зображення тобто перетворення образотворчої форми у числові дані. Знімок розділений на дрібні сегменти, які називаються пікселями або елементами зображення [1-3].

Щоб полегшити цей процес, оцифрування відбувається в координатах (x, y) і амплітуді яскравості. Зображення стає масивом цілих чисел. Рівні яскравості та розміри зображення іноді досить великі.

Кожне число в пікселі вказує на інтенсивність кольору в цій точці.

Якість зображення повністю залежить від розміру рівня та світлового малюнка (тобто сірого). Цей параметр також вказується на обсяг пам'яті, який займає велике зображення через розмір і кількість бітів. І навпаки, якщо розмір невеликий, це може викликати ефект зернистості.

Якщо зображення змінюється з 1920×1024 на 512×1024 , різниця між ними буде менш помітною, але якщо змінити його на 128×256 і знову збільшити до 512×1024 , ви можете помітити значну кореляцію.

Обробка зображень поділяється на чотири типи операцій:

– операції з пікселями: цей процес відбувається не зважаючи на всі інші пікселі у представленому екземплярі тобто все залежить лише від даного пікселю, який обрано як вхідний. Для цього використовується функція порогування, яка розподіляє їх в залежності від порогових рівнів білого та чорного. У деяких інших випадках вмикають інвертування зображення, підвищення та пониження яскравості, контрастне розтягнення;

– локальні (сусідні) операції: цей процес базується на найближчих вхідних даних від обраного пікселя. В більшості випадків операція адаптивна, тому що все залежить від точних значень для пікселів які перебувають на всій області матриці. Найкращий приклад це фільтри які використовують для покращення текстури або кольорів, відтінків і тому подібне;

– геометричні операції: процес має залежність від результатів геометричних перетворень вхідних даних для обраних пікселів. Це допомагає зменшити кількість необхідних сегментів на відміну від глобальних перетворень;

– глобальні операції: перехід до пікселів залежить від усіх сегментів зображення, при цьому здійснюється перетворення зображення зі зміщенням пікселів на мінімальну відстань до фонових сегментів.

Розпізнавання образів – це наукова дисципліна, яка передбачає виявлення характеристик і подальшу класифікацію зображень чи об'єктів у певній галузі. Методи розпізнавання образів є одними з найважливіших компонентів інтелектуальних систем і широко використовуються в різних сферах.

Зображення – це група об'єктів, які об'єднані загальними ознаками та ознаками для формування класифікації.

Розпізнавання зображень – це процес ідентифікації об'єктів на зображенні за допомогою алгоритму машинного навчання. Розпізнавання образів можна інтерпретувати як класифікацію даних, отриманих на основі вивчених знань або статистичної інформації. Для кожного класу всі знання даються для порівняння [4].

Причини використання технології розпізнавання образів:

- розпізнавання місцевості;
- класифікація змісту тексту;
- розпізнавання відео-матеріалів;
- система безпеки зі сканером ока.

У типовій системі всі отримані дані оновлюються у формі для правильного використання машини. Щоб система працювала належним чином, необхідний об'єкт обмежено однією областю, виділеною на зображенні. Якщо знайдено відповідність між об'єктом і знанням, об'єкту присвоюється позначення класу, що означає, що класифікація відбулася.

Поділ на частини допомагає знайти рішення і виходить цілий абзац. Кластеризація використовується в необробленому навчанні. Можна представити компоненти дискретними, неперервними, дискретними двійковими змінними. Обчислювальний алгоритм знаходження важливих параметрів об'єкта називається компонентом функції. На основі отриманих даних створюється вектор властивостей.

Необхідні вимоги до системи:

- чітко та швидко знаходити вивчений образ;
- знаходити закономірності у незнайомих об'єктах і їх класифікація;
- якісно аналізувати форму об'єкта з різних кутів;
- ідентифікувати образи якщо одна з частин відсутня;
- автоматично виводити результат роботи.

Навчання – це процес перетворення в адаптивну систему шляхом аналізу результатів тестування. Це дуже важливий процес, оскільки від алгоритму залежить якість і точність результатів системи розпізнавання зображень [5].

Необхідні дані включають навчальні та тестові набори, перший для підготовки системи та побудови основної функції, а другий для найбільш наближених до фактичних умов перевірки, що визначають якість у системі.

Використання навчального набору є частиною побудови моделі. Він має форму дисплея, який аналізується системою (тобто поїздами). Результати навчання визначаються алгоритмами та правилами навчання, які пояснюють зв'язок між входом і виходом. Система реалізує алгоритм, завдяки якому вихідні дані покращуються на основі навчального набору даних.

Потім систему було протестовано за допомогою тестового набору, спрямованого на перевірку результатів. Часто використовується п'ята частина файлу даних, що достатньо для вимірювання точності.

Головні переваги системи:

- розподіляє вихідні результати по класам;
- знаходить більш точну інформацію біометричних показників;
- допомагає ідентифікувати походження джерела звуку;
- дає можливість знайти та розпізнати обраний об'єкт з різних ракурсів.

Недоліки:

- сповільнення швидкості роботи через підвищення складності синтаксичного образу;
- у більшості випадків необхідний великий набір даних;
- неконкретизоване пояснення розпізнавання обраного об'єкта.

Застосування:

- обробка зображень, сегментація та аналіз. Розпізнавання образу людини через передачу інформації аналізу до машини;
- комп'ютерний зір: розпізнавання вже відомих об'єктів з бази даних та порівняння особливостей з новим об'єктом на фото або відео має дуже широке використання;
- сейсмічний аналіз: виділення необхідних частин знімку з образами, інтерпретація тимчасових об'єктів у записах сейсмічного аналізу;

- класифікація/аналіз радіолокаційного сигналу: розпізнавання збігу з вивченими даними та методи обробки вхідних сигналів має застосування у багатьох додатках класифікації радіолокаційного сигналу;
- розпізнавання мови: найкращий результат було досягнуто за допомогою шаблонів розпізнавання образів. Він має широке коло використання у різних алгоритмах та трактує більші одиниці (тобто слова);
- ідентифікація відбитків пальців: зараз представляє собою одну з самих актуальних технологій на біометричному ринку. Для цього використовуються найкращі методи та підходи до розпізнавання образів.

Об'єкт роботи – процес обробки зображень для контекстної інтерпретації в системах технічного зору.

Предмет роботи – кластерні алгоритми та метод сегментації зображень.

1.2 Огляд існуючих рішень, проблематика

Для вирішення проблеми розташування об'єктів у цифрових зображеннях і КСТ використовуються різні методи цифрової обробки зображень. Локалізація – це процес визначення площ об'єктів на зображеннях за заданими параметрами [6].

Найважливішими вимогами до КСТ є максимально можлива точність і швидкість обчислень під час його роботи. Звідси випливає, що вимога до мінімального часу роботи задається насамперед алгоритмом витягування деталей. Точність, у свою чергу, визначається складністю поділу класів на набуті компоненти точність.

У першому критеріями для порівняльного аналізу методів вилучення ознак можуть бути:

- час роботи алгоритму;
- складність у поділі ознак на класи.

Задача локалізації розуміється як задача віднесення певного прикладу до одного з декількох множин об'єктів, що попарно не перетинаються. В реальних ситуаціях при класифікації об'єктів виділяють три рівні складності:

- класи можна розділити прямими лініями (або гіперплощинами);
- лінійна подільність;
- класи можна розділити кількома гіперплощинами – нелінійна подільність;
- класи перетинаються, тому їх поділ можна трактувати тільки через ймовірнісне співвідношення.

Формалізувати і кількісно визначити ступінь складності в розділенні класів ознак практично неможливо. Тим не менш, зробити таку оцінку можна на підставі наступного міркування: чим простіше класи поділяються, тим нижче помилка класифікації. Чим нижче помилка класифікації, тим вищий коефіцієнт розпізнавання [8-10].

Оцінка коефіцієнта розпізнавання зазвичай проводиться на еталонних базах. Так, наприклад, для задачі пошуку облич на зображенні найбільш поширеними є бази MIT/CMU і BioID. Практика показує, що для бази MIT/CMU коефіцієнт розпізнавання вважається високим при його значенні 75% і вище, а для BioID – 90% і вище. Найбільш популярні бази для ідентифікації осіб – FERET, ORL, FACES, AR FACE.

Для них значення якості розпізнавання, що починається зі значення 90% з використанням додаткової інформації про координати очей, вважається високим. Таким чином, для проведення порівняльного аналізу методів вилучення ознак виділені наступні критерії:

- час роботи алгоритму (вимірюється в мілісекундах і обмежено значенням 0,025 мс);
- коефіцієнт розпізнавання (приймає значення від 0% до 100%).

Методи обробки цифрових зображень, що застосовуються для локалізації об'єктів на зображеннях.

Аналіз літератури, дозволив виділити декілька підходів до класифікації методів обробки зображень для локалізації об'єктів.

Методи обробки різняться в залежності від типу зображення локалізованого об'єкта (об'єкт може бути представлений точковим зображенням, групою точок, у вигляді майданного або протяжного зображення).

У залежності від етапу процесу локалізації (методи первинного і вторинного пошуку), залежно від застосовуваної процедури пошуку (прогресивна сканування всього зображення, порівняння виділених областей з еталоном, обробка окремих фрагментів скануючим «вікном», знаходження опорних точок, геометричний пошук).

Існують дві основні групи ознак кольорових зображень: це ознаки пікселя зображення, зазвичай це характеристики кольору пікселя, і ознаки шуканого об'єкта, тобто безліч пікселів, з яких і складається зображення об'єкта, зазвичай це геометричні характеристики.

При виборі значень параметрів локалізації, виходячи зі змінності умов отримання зображень, з метою виключення можливості втрати області об'єкту, значення параметрів задаються максимальним діапазоном можливих значень. Тому виключається можливість визначення місця помилок об'єктів [11].

На кожному етапі локалізації можна використовувати різні параметри об'єкта, тому не існує чистого методу локалізації для одного набору функцій.

Методи кореляційного аналізу зводяться до обчислення кореляції зображення заданого об'єкта зі спостережуваним зображенням. Суть локалізації за кореляційним критерієм полягає в тому, щоб виявляти пікселі (фрагменти і цілі зображення) об'єктів, що цікавлять тільки в тому випадку, якщо між їх значеннями яскравості в заданих каналах спостерігається стійкий кореляційний зв'язок [12].

Під час пошуку та локалізації за еталоном виконується кореляційна обробка ознак, отриманих від еталону і вхідного зображення як з

використанням порогів за величиною подібності, так і без встановлення порогу (коли шукається максимум подібності).

Кореляційні методи характеризуються великою обчислювальною складністю. Пов'язано це з масштабуванням і поворотами шуканого об'єкта. Головний недолік методу зіставлення з еталоном полягає в необхідності використання величезної кількості еталонів для обліку змін об'єктів, що виникають при їх повороті і збільшенні (зменшенні) розмірів, тому кореляція рідко використовується на практиці, якщо можливий довільний поворот шуканого об'єкта. З цієї причини при зіставленні з еталоном бажано обмежитися ознаками, які менше залежать від змін розміру і форми об'єкта, наприклад, структурні ознаки об'єкта.

Аналіз гістограми зображення полягає в побудові розподілу кольорових або статистичних властивостей пікселів на цифровому зображенні. Гістограма розподілу кольорових значень зображення показує процентний вміст точок певної яскравості.

Пошук і локалізація областей інтересу проводиться методом установки порогових значень яскравості, і всі точки зображення, що лежать в межах даних значень, виділяються на зображенні, як шукані області.

Якщо контраст об'єкт – фон є достатнім, тоді локалізувати такий об'єкт не складно. Основою таких методів є уявлення про те, що будь-яке зображення можна розбити на кінцеве число сегментів – ділянок, які мають будь-які характеристики. Методи сегментації застосовуються для пошуку об'єктів на цифровому зображенні (пошук країв і зв'язок з ними ліній і кіл), пошуку людей (пошук сегментів тіла за ознаками та об'єднання їх в єдине ціле), пошуку обличчя [13].

Суть структурних методів полягає у створенні основи непохідного елемента в зображенні структурного елемента об'єкта. Окремі структури є елементами зображення об'єкта, які складаються з частин прямих: проекцій прямокутних паралелепіпедів на площину (у найпростішому випадку ми

також шукаємо паралельні прями, прямокутники та паралелограми), які потрібні, наприклад, для визначення будівель і доріг.

1.3 Сучасні програмні засоби для класифікації об'єктів

Образ являє собою певну структуру в системі класифікації, яка акцентує конкретну групу об'єктів за вибраними ознаками. Образи мають характерні властивості, які проявляються в тому, що знайомство з кінцевою кількістю явищ з одного числа множини дозволяє дізнаватися про будь-яку кількість його представників.

Образи володіють об'єктивними властивостями в тому аспекті, що люди, які вчаться на різних матеріалах спостережень, переважним чином однаково та незалежно від інших поділяють одні й ті ж об'єкти за певними критеріями оцінювання. На практиці в задачі розпізнавального процесу універсальна множина ділиться на певні частини – образи [14].

При цьому відображення будь-якого об'єкта органами сприйняття системи розпізнавання незалежно від знаходження цього об'єкту щодо цих органів, називають зображенням об'єкта, а множини цих відображень, що об'єднані схожими якостями, називають образами.

Розпізнавання образів з англійської мови – *pattern recognition* – це теоретичний розділ науки штучного інтелекту, який вивчає способи класифікації об'єктів. Зазвичай об'єкт, який класифікується, являє собою образ (*pattern*). Образом може легко стати цифрова фотографія, наприклад, розпізнавання конкретних зображень; літера чи цифра, наприклад, у розпізнаванні символу; запис мови під час розпізнавання мови.

Розпізнавання образів – задача визначення об'єкта або ідентифікація певних якостей цього об'єкта за його зображенням, так зване оптичне розпізнавання, або за допомогою аудіо-запису в акустичному розпізнаванні, також за іншими властивостями [15].

Зображення також є сукупністю систем класифікації, яка підкреслює (зв'язує) конкретну групу об'єктів на стандарті.

Крім того, слід зазначити, що розпізнавання образів (дій, символів) є найбільш фундаментальною проблемою таких інтелектуальних систем, як наука. З іншого боку, роль розпізнавання образів відіграє велику практичну роль.

Для вирішення завдань розробки як автоматизованих, так і автоматизованих систем управління теорія розпізнавання образів завжди використовувала різні технології в інтелектуальній діяльності компанії як зразок. Наука у вигляді певних інструментів або алгоритмів широко використовує можливості людського інтелекту. На думку дослідників, такий стан речей найбільш розумний при вирішенні складних і складних математичних задач, оскільки в еволюції людина навчилася розвивати власні здібності до вирішення проблем. в автоматичному керуванні [16].

Очний процесор – це інтелектуальна інформаційна структура, яка може образно копіювати сучасну реальність, будучи здатною розуміти візуальну інформацію, що протікає в будь-якій формі. Крім того, очний процесор здатний розпізнавати і розпізнавати властивості події або властивості середовища об'єкта. Він тестує його обробку та автоматично видає рішення. Іноді рішення приймаються за участю оператора.

Система виконує такі дії:

- попередня фільтрація. Передбачається усунення шумів та слабозв'язаних точок зображення, виділення вихідного зображення;
- переміщення зображення. Під час даного процесу відбувається одночасне переміщення цілої картини на задану відстань по периметру цілої картини;
- зміна розміру зображення. Відбувається розширення картини на відповідну величину, а також його ущільнення;
- повороти картини;

- з'ясування центру картинки. Під центром мають на увазі центр мас. Де під одиницею маси мається на увазі значення ij . Відповідні точки в площині властивостей, що задаються;

- виконання логічних операцій над зображеннями;
- виокремлення частини зображення товщини, що задається.

Важливою рисою око-процесора є здатність приймати адаптивні рішення. Ця особливість зумовлюється наступними концептами:

- величезна кількість простору властивостей;
- присутність алгоритмів прийняття рішень, які несуть в собі вагові показники;
- наявність малопередбачуваних ситуацій в умовах реального функціонування;
- присутність обставин з пороговим значенням прийняття рішень;
- експлуатація за обставин апріорного невизначення.

Тому метод розпізнавання образів також популярний у сфері штучного інтелекту. Крім того, слід зазначити, що подія розпізнавання образів значною мірою заснована на збігу факторів. Але, як відомо, більшість речей не схожі одна на одну. Однак, як показала практика, між деякими суб'єктами завжди існує консенсус [17].

При цьому методи розпізнавання образів поділяють на два види:

- структурні;
- засновані на теорії рішень.

Як правило, при обчисленні кількісних величин використовується теорія рішень. Структурно орієнтовані зображення, зображення яких дає кращі властивості.

Важливу роль у процесі розпізнавання зображень відіграє шаблон навчання, тобто так званий упорядкований набір певних властивостей. Клас зображення – це набір об'єктів зі схожими властивостями.

Частіше використовуються такі форми атрибутів, як властивості об'єкта, векторні функції та рядкові функції. Відповідні методи є компонентами

векторних властивостей кожного класу об'єктів. Створене зображення має належати до найближчого до літака класу у заздалегідь визначеному масштабі.

Має сенс, що простим методом було б знайти мінімальну відстань, розраховану за правилами Евкліда, між певним набором векторів, що належить невідомому об'єкту, і набором векторів майбутньої моделі. Визначення того, чи належить об'єкт до класу, – це найшвидша відстань. Мінімальний розподіл відстані є кращим, якщо відстань між математичними очікуваннями класів більша за відстань розсіювання об'єктів для кожного окремого класу [18].

З іншого боку, велике значення мають методи розпізнавання образів на основі імовірнісних індикаторів з використанням збігів, що впливають на моделювання нових класів зображень.

Тому необхідно розробити оптимістичний алгоритм, здатний значно зменшити прогнозовану кількість можливих помилок. У той же час слід зазначити, що зрозуміти, як виглядати найкраще з точки зору комп'ютерного опису, непросто завдання.

Так, відомо, що всі існуючі способи поділяють на три ступені:

- процес первинної обробки;
- процес відсіювання та логічне з'ясування результатів фільтрації;
- процес прийняття рішення.

Зазвичай для науки розпізнавання образів, щоб ідентифікувати об'єкт або подію на фотографії, потрібно використовувати всі процеси в групі вище. Але іноді достатньо двох. Фаза скринінгу включає прийоми, які допомагають визначити об'єкти, які зараз цікавлять, без попереднього їх вивчення. Більшість із цих методів використовує алгоритм для всіх пікселів одночасно. На цьому рівні аналіз часто не використовується.

Найпростіший варіант – бінаризація порогового зображення. Наприклад, для зображень у відтінках сірого цим обмеженням є значення світла. Визначення та вибір двійкового порогу дозволяє самостійно визначити тип алгоритму. Бінаризація зазвичай відбувається під час процесу адитивного

встановлення. Так, така операція може бути незліченною можливістю. Слід додати, що класичні методи фільтрації використовуються в багатьох завданнях. Найпоширенішим із класичних методів перетворення є перетворення Фур'є.

Однак для таких зображень цей метод не використовується. Однак у дослідженні зображень іноді не вистачає простого одновимірного перетворення та потреби в ресурсномістких двовимірних перетвореннях [19].

Розрахунки за формулами дещо трудомісткі, тому доцільно застосувати згортку до потрібної області за допомогою низькочастотної або високочастотної фільтрації, в залежності від конкретного завдання. Крім того, слід зазначити, що цей процес спрощення не дозволяє виконувати широкий спектр операцій, таких як аналіз, але часто одного результату достатньо без подальшого коригування.

Перетворення вейвлетів є набагато кращим способом редагування зображень, ніж Фур'є. Цей метод спрощує процес складання, вивчення та надсилання кількох слайдів. Перетворення вейвлетів засноване на поділі зображень на хвилі з нижчою частотою, при цьому частота хвиль постійно рухається в часі порівняно з методом Фур'є. Це пояснюється тим, що перетворення Фур'є керується гармонічними алгоритмами [20].

З математичної точки зору фотографічні зображення є двовимірною матрицею відповідних заданих кількостей світла. Однак при переході від однієї його частини до іншої гістограми постійно змінюються. Прикладом використання вейвлет-перетворення є задача на знаходження відображення очей, де вейвлет насправді є відображенням.

В основі цього методу лежить кореляція, яка використовується в основному разом з іншими методами, але може використовуватися і окремо. Кореляція відіграє важливу роль у процесі розпізнавання зображень.

Ще один популярний тип фільтрації – фільтрація елементів. Фільтрація дозволяє знаходити частини простих функцій відображення. Найвідомішою є змінна Хаффа, яка допомагає знаходити різні ефективні обчислювальні

функції. Подібним методом є метод перетворення Радона, який оптимізує обчислення за допомогою швидкого перетворення Фур'є.

Інша частина фільтра фільтрує різні контури. Така фільтрація корисна в ситуації, коли справа дуже складна, але є чіткі межі. Тоді контурна фільтрація є одним з найважливіших механізмів роботи з зображеннями. Після фільтрації завантажується набір даних, який необхідно обробити. Але іноді вони потребують більш логічних змін. Тому потрібні методи, які можна використовувати для переміщення всього зображення до окремих частин його об'єктів.

Методи математичного перетворення є інструментом для переходу від фільтрації до логіки. Вони можуть видаляти шум з бінарної точки зору, змінюючи вже використовувані об'єкти. Існують також прийоми, які допоможуть вам визначити щось за контурами. Такий принцип у підході має назву контурного аналізу. Обрані точки являють собою визначальні індивідуалістичні властивості, що допомагають зіставити різні класи об'єктів.

Способи машинного навчання та прийняття рішень – це кінцева стадія в процесі розпізнавання. Ці способи знаходяться на межі статичної математики, способів оптимізації, а також класичних наук математики, але мають також і власну особливість, яка пов'язана з проблематикою обчислювальної продуктивності та процесом перенавчання.

Здебільшого суть навчання полягає в наступному: на базі певної навчальної вибірки з властивостями кожного класу необхідно спроектувати таку систему, завдяки якій комп'ютер зможе дати аналіз новому зображенню та визначити, який з об'єктів представлений на зображенні [21].

Слід зазначити, що людський зір легко адаптується до зору навколишнього середовища. Наприклад, ви можете відокремити основний об'єкт від його оточення за тонкими візерунками тіней.

Подивившись на фотографію групи людей, ви можете легко визначити точну кількість людей, а також передбачити, як вони зараз по виразу обличчя. Психологи, що вивчають процеси сприйняття, витрачають більше десяти років

на розуміння процесу зорової системи. Навіть використання спеціальних оптичних ілюзій у дослідженнях для пошуку деяких основних принципів та об'єктивних рішень у цій галузі сьогодні не існує.

1.4 Постановка задачі

Об'єктом роботи є процес обробки зображень для контекстної інтерпретації в системах технічного зору.

Метою роботи є отримання й обробка фото та відеоматеріалу, для виділення облич та зіставлення їх з наявною базою даних

Конкретизуючи дану постановку, можна сказати, що для вирішення поставленої задачі необхідно:

- формалізувати представлення розбиття зображень матрицями просторових відносин, які відповідають прототипам заданих класів;
- ввести і вивчити індексацію класів еквівалентності при сегментації зображень;
- обґрунтувати, синтезувати і досліджувати метрику для порівняння результатів сегментації зображень;
- провести експериментальні дослідження алгоритмів реляційної сегментації з метою визначення раціональних шляхів їх реалізації і встановлення особливостей практичного використання.

2 АНАЛІЗ ІНФОРМАЦІЙНОГО ЗАБЕЗПЕЧЕННЯ

2.1 Задачі та алгоритми

Перед системою інтелектуального розпізнавання стоїть декілька комплексних задач, для вирішення яких необхідне застосування різноманітних методів та алгоритмів (рис. 2.1).

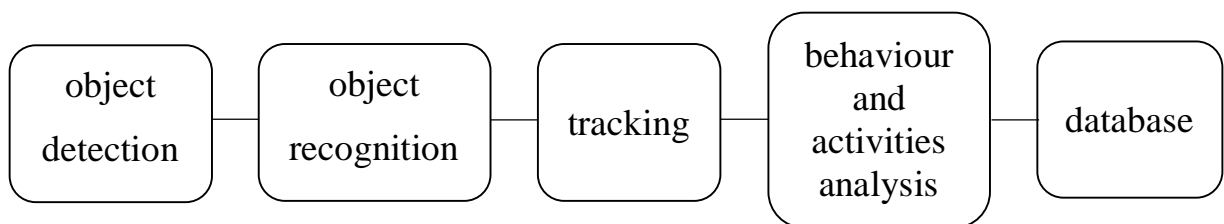


Рисунок 2.1 – Традиційний етапи обробки

Розпізнавання образів є однією з найбільш фундаментальних проблем теорії інтелектуальних систем. З іншого боку, роль розпізнавання образів має велике практичне значення. Замість терміна «ідентифікація» часто використовується інший – «класифікація». Ці два терміни в багатьох випадках вважаються подібними, але не повністю взаємозамінними. Кожне з них має свою сферу застосування, і тлумачення обох понять завжди залежить від деталей конкретного завдання.

Наведемо деякі типові постановки задач розпізнавання:

- задача ідентифікації, яка полягає в тому, щоб вирізнити певний конкретний об'єкт серед його подібних (наприклад, впізнати серед інших людей свою дружину);
- віднесення об'єкта до того чи іншого класу або класифікація. Це може бути, наприклад, задача розпізнавання літер або прийняття рішення про наявність дефекту у деякій технічній деталі. Віднесення об'єкта до певного класу відображає найтипівшу проблему класифікації, і, коли говорять про

розпізнавання образів, найчастіше мають на увазі саме цю проблему. Тому ми розглядатимемо її першочергово;

– кластерний аналіз, який полягає в поділі заданого набору об'єктів на класи – групи об'єктів, схожі між собою за тим чи іншим критерієм. Цю задачу часто називають класифікацією без учителя, оскільки, на відміну від задачі 2, класи апріорно не задані.

Розпізнавання зображень (об'єктів, сигналів, ситуацій, подій чи процесів) – завдання розпізнавання об'єкта або визначення однієї з його властивостей за його зображенням (оптичне розпізнавання) чи звукозаписом (акустичне розпізнавання)) та іншими характеристиками.

Одним з основних понять є усталена концепція. У комп'ютері набір представлений набором унікальних елементів одного типу. Слово «унікальний» означає, що елемент множини присутній або відсутній. Універсальний набір містить усі можливі елементи для вирішення проблеми, зміст марний.

Образ – класифікаційне угруповання в системі класифікації, що об'єднує (виділяє) певну групу об'єктів за деякою ознакою. Образи володіють характерною властивістю, що виявляється в тому, що ознайомлення з кінцевим числом явищ з однієї і тієї ж множини дає можливість дізнаватися про як завгодно велике число його представників. Образи мають характерні об'єктивні властивості в тому сенсі, що різні люди, що навчаються на різному матеріалі спостережень, здебільшого однаково і незалежно один від одного класифікують одні й ті ж об'єкти. У класичній постановці завдання розпізнавання універсальна множина розбивається на частини-образи. Кожне відображення якого-небудь об'єкта на сприймаючі органи системи, що розпізнає, незалежно від його положення щодо цих органів, прийнято називати зображенням об'єкта, а множини таких зображень, об'єднані якими-небудь загальними властивостями, являють собою образи [23].

Метод присвоєння елементу зображенню називається правилом рішення. Ще одне важливе поняття – метрика, спосіб визначення відстані між

елементами універсального набору. Чим менша відстань, тим більше подібних об'єктів (символів, звуків тощо) ми можемо розпізнати. Елементи найчастіше представлені у вигляді набору чисел, а показники – у вигляді функції. Ефективність програми залежить від вибору відображення зображення та реалізації метрики. Один і той самий алгоритм розпізнавання з різними показниками виходить з ладу на різних частотах.

Навчанням часто називають процес, за допомогою якого система реагує на групи зовнішніх ідентичних сигналів, багаторазово піддаючи систему зовнішньому адаптації. Такі зовнішні зміни в освіті називаються «заохоченнями» та «штрафами».

Механізм здійснення цього налаштування майже повністю визначає алгоритм навчання. Навіть навчання відрізняється від навчання, оскільки не повідомляється додаткова інформація про точність відповіді системи.

Адаптація – це процес зміни параметрів і структури системи, а можливо і управляючих впливів, на основі поточної інформації з метою досягнення певного стану системи при початковій невизначеності в умовах роботи, що змінюються [24].

Навчання – це процес, в результаті якого система поступово набуває здатність відповідати потрібними реакціями на певні сукупності зовнішніх впливів, а адаптація – це підстроювання параметрів і структури системи з метою досягнення необхідної якості управління в умовах безперервних змін зовнішніх умов.

Проблеми розпізнавання легко вирішуються людьми, як правило, підсвідомо. Спроби ж побудувати штучні системи розпізнавання не настільки переконливі. Основна проблема полягає в тому, що часто неможливо адекватно визначити ознаки, на основі яких слід здійснювати розпізнавання. Системи штучного розпізнавання набули широкого поширення для завдань, де такі елементи можуть бути ізольовані.

Одне з найбільш вдалих формулювань ключової парадигми в теорії розпізнавання полягає в тому, що все в природі унікальне, все стандартизовано.

Значення цієї парадигми полягає в наступному. Кожен об'єкт описується властивістю. Наявність або відсутність таких властивостей, а також якісні та кількісні характеристики цих властивостей вважаються характеристиками об'єкта. Особливість всього означає, що в природі не існує двох різних речей, де всі функції однакові, і це дозволяє, принаймні теоретично, відрізнити одне від іншого. Але деякі частини різних об'єктів можуть збігатися, і це говорить про те, що ці об'єкти належать до одного типу або класу.

2.2 Перетворення Радона, перетворення Хафа, робастні оцінювачі

В математиці перетворення Радона – це інтегральне перетворення, яке переводить функцію f , визначену на площині, в функцію Rf , визначену на (двовимірному) просторі з прямих на площині, значення якої у певній прямій дорівнює криволінійному інтегралу від цієї функції над цією прямою.

Зміна була введена в 1917 році Йоганном Радоном, який також надав формулу зворотної зміни. Крім того, Радон визначив формули для тривимірного перетворення, при якому збіжність має місце на площині (пов'язана з лінією, відомою як рентгенівське перетворення. Пізніше це було узагальнено Евклідом). або інтегральна геометрія. Комплексний аналог перетворення радону, відомий як перетворення Пенроуза, перетворення Радона широко використовується для томографії, де дані проєкції зображення пов'язані зі скануванням поперечного перерізу об'єкта. Якщо функція представляє невизначену щільність, то зміна радону представляє дані проєкції, отримані як вихідні дані томографічного сканування.

Таким чином, обернене перетворення радону може бути використане для відновлення початкової щільності за проєкційними даними i , таким чином,

становить математичну основу томографічної реконструкції, яку також називають ітераційною реконструкцією.

Дані перетворення радону часто називають синограмами, оскільки перетворення радону в центрі точкового джерела є синусоїдною хвилею. Таким чином, перетворення радону для багатьох невеликих графічних об'єктів можна розглядати як кількість розмитих синусоїд з різними амплітудами і фазами. Перетворення Радона корисне при комп'ютерній томографії (КВТ-сканування), також застосовується у сканерах штрих-кодів, електронній мікроскопії, макромолекулярних зборах, таких як віруси та білкові комплекси, рефлекторній сейсмології та при вирішенні гіперболічних часткових диференціальних рівнянь.

За десятиліття свого існування алгоритм Хафа зарекомендував себе як цінний інструмент фахівця з аналізу зображень та технічного зору, що засвідчено в оглядових роботах.

В даний час безліч наукових колективів публікують роботи про те, як використовувати ПХ при вирішенні різних дослідницьких і інженерних задач. Очевидно, що в цьому випадку є можливість замінити СПХ на БПХ, значно зменшивши час обчислення відповідного етапу. Однак в силу того, що у діадичних патернів є деяка специфічна помилка апроксимації прямої на зображенні, точність рішення задачі може змінитися. Як показують дослідження, викривлення Хаф-образу на розмірах зображення незначні, що дозволяє зробити якісний висновок про допустимість заміни СПХ на БПХ. Таким чином, область застосування БПХ не вичерпується тільки роботами, де безпосередньо використовувався цей алгоритм, але і містить роботи, які експлуатують СПХ, а також деякі його модифікації. Основним способом застосування, для якого використовується ПХ – це детекція ліній на зображенні, часто прямих.

Однією з формалізацій даного завдання є завдання робастної ортогональної лінійної регресії. Детальніше це питання описано в наступному розділі. Крім цього, існує безліч робіт, що пропонують використання ПХ для

детекції ліній на зображенні.

ПХ застосовується для аналізу зображення документа, зокрема для визначення меж документа.

Визначивши орієнтацію документа на зображенні, можна проективно виправити зображення так, ніби воно знято вертикально, а межі зображення паралельні кордонам фотографічної матриці. Рішення завдання віртуального повороту документа до цього виду дозволяє істотно спростити подальший аналіз документа і реалізацію процедур розпізнавання його структури і полів. У контексті аналізу зображень документів цікаво згадати роботу в якій з використанням БПХ пропонується визначати параметри нахилу шрифту в тексті. Зараз вже можна конструювати системи технічного зору індустриальної якості, використовуючи БПХ [25].

Іншою великою областю застосування ПХ є технічне зір автономних систем, що функціонують на антропогенних територіях (що містять велику кількість прямолінійних образів). Наприклад, ПХ часто використовується для детекції дорожнього полотна і нанесеною на ньому дорожньою розміткою та для детекції дорожньої інфраструктури міста по знімках з безпілотної (який високо летить). ПХ часто використовується при вирішенні різних завдань відстеження об'єктів на зображенні. Так, наприклад, в одній з перших робіт запропоновано, використовуючи покадрову різницю зображень, відстежувати рівномірно рухомі об'єкти, які на таких треках, очевидно, залишають сильно виражені прямі лінії.

Ще одним критично важливим завданням є детекція точки сходу на зображенні. Точка сходу – це точка на площині зображення в якій перетинаються проекції паралельних в тривимірній сцені прямих. Цікавим способом застосування БПХ стало його використання для реалізації алгоритмів сліпого калібрування радіальної дисторсії. Автори пропонують новий функціонал Хаф-образу зображення, оптимізуючи який підбирають параметри компенсації радіальної дисторсії. Основне і єдине припущення, в рамках якого метод функціонує – наявність прямолінійних об'єктів в сцені.

Ще одним дуже популярним і одним з найстаріших способів застосування перетворення Хафа і перетворення Радону, є завдання томографічного відновлення. Процедуру формування проєкцій зображення аналізованого об'єкта з використанням рентгенівських променів прийнято моделювати перетвореннями Радона і Хафа. В результаті задача відновлення тривимірної структури об'єкта на практиці часто зводиться до задачі обернення перетворення Хафа або перетворення Радону, – як кажуть, до розв'язання зворотної задачі. Даній науковій сфері присвячена велика кількість робіт, в тому числі вітчизняних. В останні кілька десятиліть варто відзначити значне підвищення інтересу до аналізу медичних зображень, докладніше про це написано в роботі.

Взагалі кажучи, список різноманітних застосувань ПХ і БПХ для аналізу зображень і в якості одного з етапів функціонування системи технічного зору дуже великий. Перерахуємо деякі ще не згадані, але цікаві й оригінальні способи застосування ПХ.

ПХ пропонували використовувати для розпізнавання штрих-кодів аналізу деформацій підземного профілю, визначення ступеня подібності форм, детекції треків супутників по зображеннях з довгою витримкою, для виявлення радіолокаційних цілей, для розпізнавання рукописного тексту, для аналізу структури топології мікросхем по фотографіях, для підрахунку кількості осей транспортного засобу по треках детектора коліс зображень, отриманих з камери, знятої збоку, для тривимірної реконструкції плоских граней прозорих мінералів по набору зображень, для векторизації тривимірних зображень, для аналізу SAR знімків і так далі.

Перейдемо безпосередньо до опису алгоритму БПХ. Входом алгоритму є двовимірне зображення, а виходом – співрозмірний двовимірний Хаф-образ, в кожному пікселі якого міститься значення суми за відповідним координатам цього пікселя діадичного паттерну.

Код функції *fht2* для обчислення БПХ наведено нижче на мові MATLAB.

Лістинг 2.1 Алгоритм обчислення швидкого перетворення Хафа, записаний на мові MATLAB:

```

function h = fht2 (I)
n = size (I, 2);
if n < 2
h = m;
5 return
end
h = mergeHT ( fht2 (I(:, 1 : n / 2, :)), fht2 (I(:, (n / 2 + 1) :
n, :)));
end
10 function h = mergeHT (h0 , h1)
[m, n0 , o] = size (h0);
n = 2 * n0;
h = zeros (m, n, o);
r = (n0 - 1) / (n - 1);
15 for i = 1 : n
t = i - 1;
t0 = round (t * r);
s = t - t1;
h(:, i, :) = h0 (:, t0 + 1, :) + [h1(s + 1 : m, t0 + 1, :);
h1 (1 : s, t0 + 1, :) ];
20 end
End

```

Вхідне зображення I_n^2 рекурсивно розбивається на два рівних зображення вертикальної прямої, поки ширина його не дорівнюватиме 2.

Потім кожен стовпець підсумовується з парним поелементно, після чого правий стовпець циклічно зсувається вниз на один піксель і виконується друге поелементне підсумовування [26].

В результаті обчислюються суми по паттернам довжини 2 з нахилами: $t = 0$ і $t = 1$. Отримані суми застосовуються на наступному кроці рекурсії для обчислення сум по паттернам висоти 4 в кожній четвірці рядків. Так, щоб обчислити всі суми по паттернам довжини 4 і нахилом $t = 0$, досить скласти суми по паттернам довжини 2 у сусідніх пар рядків.

Процедура триває, поки ширина зображення, отриманого в результаті,

не буде складати n .

У своїй роботі П. Брейді показав вірність наступного твердження.

Алгоритм БПХ обчислює суми за всіма паттернами основної системи БПХ зображення $I_{(n',n)}$, де $n = 2^k$, причому $n' = 2a \cdot n$, де $a \in \mathbb{N}$ – довільно.

Оцінимо обчислювальну складність БПХ.

Сума по всім паттернам висоти 2 для пари рядків вимагає $2n$ операцій, і таких пар $-\frac{n}{2}$, для наступного рівня рекурсії потрібно $4n$ операцій, але четвірок на зображенні в два рази менше, ніж пар $-\frac{n}{4}$, продовжуючи, отримаємо наступну суму: $2n \cdot \frac{n}{2} + 4n \cdot \frac{n}{4} + \dots + n$. В силу того, що таких доданків $\log_2 n$ (по числу рівнів рекурсії), то отримуємо оцінку складності $n^2 \log n$.

Відповідно, для всіх типів прямих складність становитиме $4n^2 \log_2 n$. Легко проілюструвати за рахунок чого досягається прискорення (рис. 2.2). На основі принципу динамічного програмування, експлуатуючи структуру діадичного паттерну вдається збільшити кількість перевикористання проміжних обчислень: одна сума по подпаттерну використовується більшу кількість разів.

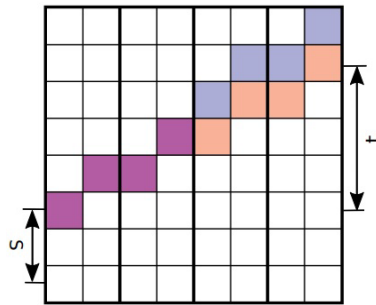


Рисунок 2.2 – Два діадичних паттерни з нахилами, що відрізняються на одиницю

Обчислення БПХ може бути прискорене за рахунок паралелізації обчислень. Оскільки на кожному кроці алгоритму виконується в точності n^2 додавань, то з використанням n^2 процесорів дані обчислення можна провести за константний час.

А так як глибина рекурсії дорівнює $\log_2 n$, то час обчислення може бути

в цьому випадку оцінено як $\Theta(\log_2 n)$ при тій же кількості використовуваної пам'яті.

Якщо дані в параметричних формулюваннях занадто суворі – їх функції розподілу мають належати до параметричного сімейства, у непараметричних вони дуже слабкі – слід продовжити роботу з функціями розподілу. Він ігнорує апіорну інформацію про «оцінювальну форму» розповсюдження. Можна очікувати, що розрахунок у цій «оціночній формі» апіорі покращить статистичні методи. Розвитком цієї ідеї є теорія стійкості (стійкості) статистичних методів, яка вважає, що розподіл вихідних даних сильно відрізняється від деяких параметричних сімейств. Ця теорія була розроблена П. Хубером, Ф. Гампелем та багатьма іншими з 1960-х років. З монографій російською мовою, що трактують про робастності і стійкість статистичних процедур, найпершою і найзагальнішою була книга, що впливає – монографія. Окремими випадками реалізації ідеї робастності (стійкості) статистичних процедур є розглянуті нижче статистика об'єктів нечислової природи та інтервальна статистика.

Існує велика розмаїтість моделей робастності залежно від того, які саме відхилення від заданого параметричного сімейства допускаються. Найпопулярнішою виявилася модель викидів, у якій вихідна вибірка «засмічується» малим числом «викидів», що мають принципово інший розподіл. Однак ця модель представляється «тупиковою», оскільки в більшості випадків великі викиди або неможливі через обмеженість шкали приладу, або від них можна позбутися, застосовуючи лише статистики, побудовані по центральній частині варіаційного ряду. Крім того, в таких моделях частота блокування, як правило, вважається відомою, що разом із зазначеним вище робить їх непридатними для практичного використання. Модель Ю. Більш імовірним видається Н. Благовіщенського, де відстань між розподілом кожного елемента у вибірці та основним розподілом не перевищує задану невелику величину [27].

2.3 Вивчення об'єктів за допомогою приблизного медіального фільтру та методу оптичного потоку

Медіанний фільтр (МФ) замінює центральний елемент маски медіаною впорядкованої вибірки, сформованої зі всіх амплітуд відділків, що покриваються маскою фільтра. При застосуванні МФ відбувається послідовна обробка кожної точки кадру, у результаті чого утворюється послідовність оцінок. При медіанній фільтрації використовується ковзне двовимірне вікно. У принципі, для кожного відліку виконується незалежна оцінка медіани у вікні. З метою прискорення оцінки доцільно алгоритмічно на кожному кроці використовувати раніше виконані обчислення. Розмір вікна встановлюється непарним і рівним $m \times n$. Відліки зображення, що знаходяться в межах вікна, утворюють робочу вибірку поточного відліку. Якщо впорядкувати послідовність $\{f_i, i = [1, mn]\}$ по зростанню, то її медіаною буде той елемент вибірки, що займає центральне положення в цій упорядкованій послідовності. Цей елемент є $\frac{(mn+1)}{2}$ -найбільшим і $\frac{(mn+1)}{2}$ -найменшим значенням у вибірці й визначає результат медіанної фільтрації для поточної точки кадру. Введемо позначення описаної процедури у вигляді:

$$g_{med} = med(f_1, f_2, \dots, f_k). \quad (0.1)$$

Розглянемо приклад. Припустимо, що впорядкована послідовність Y у вікні розміром 3×3 має вигляд: $Y = \{76, 100, 69, 120, 210, 143, 87, 130, 155\}$ де елемент 210 відповідає центру вікна (x, y) . Велике значення яскравості в цій точці кадру є результатом впливу імпульсної перешкоди. Упорядкована по зростанню вибірка має вигляд: $\{69, 76, 87, 100, 120, 130, 143, 155, 210\}$, отже, відповідно до розглянутої вище процедури, на виході медіанного фільтра одержуємо $g_{med} = 120$. Бачимо, що врахування яскравостей елементів околиці при фільтрації в поточній точці призвів до придушення імпульсної завади. Якщо імпульсна завада не є точковою, а займає деяку область, то вона

також може бути подавлена, якщо розмір цієї локальної області буде менше, ніж половина розміру апертури МФ. Тому для придушення імпульсних завад, що вражають локальні ділянки зображення, варто збільшувати розміри апертури МФ.

З цього слідує, що дія МФ полягає в «ігноруванні» як позитивних, так і негативних викидів значень вхідної вибірки. Такий принцип придушення завад може бути застосований і для ослаблення шуму на зображенні.

Однак дослідження придушення шуму за допомогою медіанної фільтрації показує, що її ефективність при рішенні цього завдання нижче, ніж у лінійної фільтрації. Медіанна фільтрація краще зберігає границі зображення, ніж будь-яка лінійна фільтрація.

Медіанні фільтри придушують імпульсні шуми. До таких шумів відноситься шум типу «сіль і перець», відліки якого мають значення, що відповідають максимальному («сіль») і мінімальному («перець») рівням квантування в сигналі зображення. Різкі зміни амплітуди зберігаються медіанним фільтром, а імпульсна завада, розмір якої $\leq \frac{mn}{2}$ таким фільтром придушується.

Однак при збільшенні маски фільтра можна втратити інформацію про малорозмірні області зображення та призвести до спотворення границь областей, особливо в кутових положеннях.

Оскільки застосування МФ призводить до придушення високих частот зображення, викликаючи розмивання країв і текстур, все більший розвиток отримують схеми адаптивної фільтрації, які дозволяють змінити імпульсну характеристику фільтра залежно від локального значення сигналу зображення.

Один з алгоритмів адаптивної медіанної фільтрації (АМФ) виконується таким способом. У вікні фільтрації оцінюються мінімальне значення сигналу f_{min} , максимальне значення f_{max} і медіана f_{med} .

Фільтрації піддається тільки той центральний елемент вікна $f(x, y)$, для якого виконується умова (2.2): значення медіани більше мінімального й менше

максимального значень у вікні й не виконується умова (2.3): значення сигналу в центрі вікна більше мінімального й менше максимального значень у вікні.

$$A_1 = f_{med} - f_{min}; A_2 = f_{med} - f_{max}; A_1 > 0, A_2 < 0; \quad (0.2)$$

$$B_1 = f(x, y) - f_{min}; B_2 = f(x, y) - f_{max}; B_1 > 0, B_2 < 0. \quad (2.3)$$

Застосування такого фільтра дозволяє видалити біполярну імпульсну заваду, забезпечити згладжування шумів і зменшити придушення високих частот у зображенні.

На рисунку нижче наведено приклад усунення шуму за допомогою медіанного фільтра.

У системі Matlab (Image Processing Toolbox) існує можливість формування й накладення на зображення трьох типів шумів. Для цього використовується вбудована функція `imnoise`, що призначена, в основному, для створення тестових зображень, що використовуються при виборі й дослідженні методів фільтрації шуму. В даному випадку, на зображення був накладений імпульсний шум за допомогою команди «`J = imnoise(I, 'salt & pepper', 0.02)`».

Для наочного порівняння приведемо три зображення разом: вихідне, зашумлене та відновлене. Як видно з рисунку вихідне та відновлене зображення майже не відрізняються один від одного.



Рисунок 2.3 – Приклад застосування медіанної фільтрації

Оптичний струм – це діаграма видимого руху об’єктів, поверхонь і поверхонь у видимій сцені внаслідок відносного руху глядача та сцени. Оптичний потік також можна визначити як розподіл видимої швидкості світла у візерунку зображення. Поняття оптичного потоку було введено американським психологом Джеймсом Дж. Гібсоном у 1940-х роках для опису зорових подразників, які виникають у тварин, які рухаються по всьому світу. Джеймс Гібсон наголошував на важливості оптичного струму для сприйняття можливостей, здатності розрізняти можливості дії в навколишньому середовищі. Послідовники Гібсона та його екологічного підходу до психології згодом показали роль стимулу оптичного потоку для сприймання руху спостерігачем у світі; сприймання форми, відстані й руху об’єктів у світі та контролю пересування [28].

Термін «оптичний потік» також використовують робототехніки, охоплюючи споріднені методики з обробки зображень та керування навігацією, включно з виявленням руху, сегментуванням об’єктів, інформацією про час до контакту, обчисленнями фокусу розширення, освітленістю, кодуванням компенсації руху, та вимірюванням стереодиспаритету.

Послідовності впорядкованих зображень уможливають оцінювання руху або як миттєвих швидкостей зображення, або як дискретних зміщень зображень. Фліт та Вейсс пропонують навчальний посібник з оптичного потоку на основі градієнта. Джон Л. Баррон, Девід Дж. Фліт і Стівен Бечеман надають аналіз продуктивності різноманітних методів оптичного потоку з акцентом на точність і щільність вимірювання.

2.4 Розпізнавання та відслідковування об’єктів

Технологію відстеження можна розділити на два основних методи: 2-D моделі з чіткими формами або без них і 3-D моделі. Метод моделі чітко

використовує апріорні геометричні знання об'єктів, що підлягають спостереженню.

На сучасному етапі розвитку інформаційних технологій велика увага приділяється системам, які використовують машинний зір як основне джерело інформації. Це пов'язано з необхідністю автоматизації та робототехніки в промислових процесах, науці та побуті. Ці процеси часто пов'язані з монотонною роботою, є небезпечними для людей або вимагають швидкого прийняття рішень. Прикладами таких систем є: біометрична ідентифікація, відеоспостереження, автономні транспортні системи, індексація зображень і відео в базах даних за їх змістом тощо. Також поширення мобільних телефонів з фотоапаратами та відеокамерами, наявність комп'ютерів зі значною обчислювальною потужністю призвели до швидкого розвитку комп'ютерного зору.

Комп'ютерний зір – це теорія і технологія створення програм і машин, які можуть виявляти, відстежувати та класифікувати об'єкти. Основна функція комп'ютерного зору – витягувати інформацію із зображень у вигляді їх послідовності. Тому актуальна тема роботи орієнтована на ідентифікацію та розпізнавання об'єктів. Слід зазначити, що розвиток комп'ютерного зору прискорили вдосконалення алгоритмів машинного зору, зокрема впровадження штучних нейронних мереж. Значний внесок у розробку таких алгоритмів зробив проект OpenCV, започаткований ініціативою Intel у 1999 р.

Зараз проект OpenCV підтримується некомерційною організацією OpenCV.org. Результатом проекту стало створення однойменної бібліотеки.

OpenCV (Open Source Computer Vision Library) – бібліотека з відкритим кодом функцій і алгоритмів для обробки зображень і бачення. Бібліотека має понад 2500 оптимізованих алгоритмів, які включають повний набір класичних та сучасних алгоритмів для комп'ютерного зору та машинного навчання. Він може бути вільно доступним для академічних та комерційних цілей. Відомі компанії, такі як Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota, а також багато стартапів використовують бібліотеку у своїх продуктах.

Для ідентифікації об'єктів на зображенні використовуються дві стратегії: моделювання фону та моделювання об'єктів. Вибір стратегії залежить від умов запису. Фонове моделювання можна використовувати лише в «хороших» умовах зйомки. Об'єктне моделювання – це більш загальний підхід.

Існує кілька способів моделювання об'єкта та пошуку об'єкта на зображенні. Не можна сказати, що один метод ефективніший за інший. Вибір того чи іншого методу залежить від багатьох умов. Найпростіший спосіб вибрати об'єкт на зображенні – застосувати кольорові фільтри. Ці методи використовуються, коли об'єкт видно у фоновому режимі.

Якщо об'єкт трохи складний, але виглядає добре, бажано розмітити край і проаналізувати контур. Це дозволяє перейти від роботи над зображенням до роботи з об'єктами на цьому зображенні. Потім ви можете перевірити наявність деяких геометричних фігур на зображенні. Спосіб використання полягає в пошуку на зображенні місць, які відповідають зображенню об'єкта, який ви шукаєте.

Якщо зображення об'єкта повертається або масштабується відносно шаблону, цей метод неефективний. Для таких випадків перевага віддається методам, заснованим на так званих спеціальних пунктах.

Особливі точки – це особливі властивості чого-небудь. Вони дозволяють відобразити об'єкт собі або подібним класам об'єктів. Існує кілька способів виділяти особливі точки. Деякі способи виділяють особливі точки на сусідніх кадрах, деякі – через великі проміжки часу та при різному освітленні, деякі дозволяють знайти особливі точки, навіть при повороті зображення.

Найскладнішими випадками розпізнавання є пошук об'єктів певного класу. В таких випадках задачу виявлення і розпізнавання можна вирішити за допомогою побудови класифікатора на основі машинного навчання, який складається з метода виділення особливостей (feature extractor) та власне класифікатора. Методи виділення особливостей залежать від поставленої задачі.

Для одного класу задач це може бути навчання на позитивних і негативних наборах зображень, для інших – це виділення кластерів дескрипторів особливих точок і створення, так би мовити, словника дескрипторів.

Не зважаючи на велику кількість відомих методів розпізнавання об'єктів не існує універсального набору для всіх умов розпізнавання. Задача розпізнавання може бути вирішена лише за конкретних умов [29].

2.5 Аналітика поведінки

Однією з найскладніших проблем у сфері комп'ютерного зору та штучного інтелекту є семантичне навчання поведінці та розуміння з боку спостереження за діяльністю візуального спостереження.

Дослідження в цій галузі зосереджуються головним чином на розробці методів аналізу візуальних даних з метою вилучення та обробки інформації про поведінку фізичних об'єктів (наприклад, людей) на сцені. В автоматизованих системах візуального спостереження важливе практичне значення має надійне виявлення підозрілих чи загрожуючи поведінці людей.

Автоматизована система візуального спостереження, як правило, вимагає надійного поєднання зображень та методів штучного інтелекту. Методи обробки зображень використовуються для забезпечення функцій зображення низького рівня. Для надання професійних рішень використовуються методи штучного інтелекту.

Було проведено широке дослідження низькорівневих методів обробки зображень, таких як виявлення, розпізнавання та відстеження об'єктів; однак відносно невелика кількість досліджень повідомила про надійну класифікацію та розуміння людської діяльності на основі послідовностей зображень.

Специфікація програми – це опис завдання, яке програма повинна виконати. Тому постає проблема формування словника абстрактної системи,

необхідно знайти відповідні терміни, які є найважливішим інструментом опису складних систем. З цією метою створено чотири головні книги: поняття, відносини, процеси та дії.

Суттєва відмінність організації понятійних засобів при розробці специфікації від завдань програмування полягає у тому, що опис складної системи вимагає від понять виразності і гнучкості, але, на відміну від завдань програмування, не є обов'язковою їх безпосередня машинна реалізація. Опис складної системи може бути дуже великим за обсягом. Тому сучасні методи опису складних систем практично повністю орієнтовані на використання інструментальних засобів у вигляді пакетів прикладних програм, що функціонують на ЕОМ.

Складною системою вважається набір мета-класів, класів, їх екземплярів та об'єктів. Стани завдань не дозволяють перенесення під час їх виконання. При об'єктно-орієнтованому підході природа системи задається не у вигляді процесів, що розвиваються з часом, а у формі надсилання повідомлень, викликаних станом самих об'єктів, які згодом викликають дії, серед іншого.

Об'єкти можуть отримувати повідомлення, інтерпретувати, активувати відповідні методи, змінювати свій статус, надсилати власні повідомлення тощо. Якщо зміна відбувається з часом, її можна розглядати як кілька подій, одну до зміни стану, а іншу – після зміни стану.

Якщо ви хочете розглянути умову, використовуйте кілька подій з умовою та результатом умови. Одночасні події (наприклад, чернетки) можна згрупувати в класи подій, де можна створити зв'язки узагальнення та агрегації.

2.6 Обґрунтування вибору програмних засобів

AI-проекти відрізняються від традиційних програмних проектів. Відмінності полягають у технологічному стеку, навичках, необхідних для проекту на базі AI, та необхідності глибоких досліджень. Для реалізації ваших

прагнень AI слід використовувати стабільну, гнучку та доступну інструментальну мову програмування. Python пропонує все це, саме тому ми сьогодні бачимо багато проектів AI які його використовують.

Від розробки до розгортання та обслуговування, Python допомагає розробникам бути продуктивними та впевненими у створенні програмного забезпечення. Переваги, завдяки яким Python найкраще підходить для машинного навчання та проектів на основі AI, включають простоту та послідовність, доступ до великих бібліотек та рамок для AI та машинного навчання (ML), гнучкість, незалежність платформи та широке співтовариство. Вони сприяють загальній популярності мови.

Простий і послідовний

Python пропонує стислий і читабельний код. Хоча складні алгоритми та універсальні робочі процеси стоять за машинним навчанням та AI, простота Python дозволяє розробникам писати надійні системи. Розробники змушені докладати усіх зусиль для вирішення проблеми ML, а не зосереджуватися на технічних нюансах мови.

Широкий вибір бібліотек та рамок

Впровадження алгоритмів AI та ML може бути складним і вимагає багато часу. Важливо створити добре структуровану та перевірену обстановку, щоб розробники могли розробити найкращі рішення щодо кодування.

Щоб скоротити час розробки, програмісти звертаються до ряду фреймворків і бібліотек Python. Бібліотека програмного забезпечення – це попередньо написаний код, який розробники використовують для вирішення загальних завдань програмування. Python, має багатий набір технологій, має широкий набір бібліотек для штучного інтелекту та машинного навчання. Ось деякі з них:

- Keras, TensorFlow та Scikit-learning для машинного навчання;
- NumPy для високоефективних наукових обчислень та аналізу даних;
- SciPy для комплексних обчислень;

- Pandas для аналізу даних загального призначення;
- Seaborn для візуалізації даних.

Платформонезалежність

Залежність від платформи відноситься до мови програмування або фреймворку, що дозволяє розробникам реалізовувати об'єкти на одному комп'ютері та використовувати їх на іншій машині без (або дуже мало) змін.

Одним із ключових чинників популярності Python є те, що мова є платформою незалежною. Python підтримується багатьма платформами, включаючи Linux, Windows та macOS. Код Python може використовуватися для створення автономних виконуваних програм для найбільш поширених операційних систем, а це означає, що програмне забезпечення Python можна легко поширювати та використовувати в цих операційних системах без інтерпретатора Python.

Більше того, розробники зазвичай використовують такі сервіси, як Google або Amazon для своїх обчислювальних потреб. Однак ви часто можете знайти компаній та науковців, які використовують власні машини з потужними графічними процесорами (GPU) для навчання своїх моделей ML. А той факт, що Python не залежить від платформи, робить це навчання набагато дешевшим та простішим.

TensorFlow є популярною основою для машинного навчання та глибокого навчання. Це безкоштовна бібліотека з відкритим кодом, випущена 9 листопада 2015 року та розроблена командою Google Brain. Він повністю заснований на мові програмування Python і використовується для чисельних обчислень і потоку даних, що робить машинне навчання швидшим і простішим.

TensorFlow може навчати та запускати глибокі нейронні мережі для розпізнавання зображень, класифікації чисел рукописного тексту, періодичних нейронних мереж, вставки слів, обробки природної мови, виявлення відео тощо. TensorFlow працює на більшій кількості процесорів, ніж графічних процесорів, таких як мобільні операційні системи.

Вперше він був оприлюднений наприкінці 2015 року, тоді як перша стабільна версія з'явилася у 2017 році. Він є відкритим кодом за ліцензією Apache Open Source. Ви можете користуватися ним, змінювати його та розповсюджувати модифіковану версію за окрему плату, не платячи нічого Google.

Він називається Tensorflow, оскільки він приймає вхідні дані як багатовимірні масиви, також відомі як тензори. Ви можете створити будь-яку блок-схему (так звана графік), яку ви хочете створити з цим записом. Вхідний отвір знаходиться з одного кінця, а потім проходить через цю систему за кілька операцій, а інший кінець залишається вихідним.

Тензор може виникнути з вхідних даних або з результату обчислення. За допомогою TensorFlow всі операції виконуються за розкладом.

3 РОЗРОБКА ІНФОРМАЦІЙНОЇ СИСТЕМИ, АЛГОРИТМИЧНОГО ТА ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

3.1 Сучасні вимоги та потреби у класифікації об'єктів на зображеннях

Давайте розглянемо етапи усунення несправностей. У цій статті розглядаються два способи вирішення проблеми ізоляції об'єкта в зображенні: за допомогою каскадного методу нейронних мереж.

Першим кроком є створення образу, тобто очищення зображення від звуку відбувається плавно. Для цього можна використовувати бібліотеку OpenCV. Далі йде крок ідентифікації об'єкта. Існує кілька способів досягти цього шляхом пошуку циклів, нейронних мереж або каскадної класифікації (рис. 3.1).

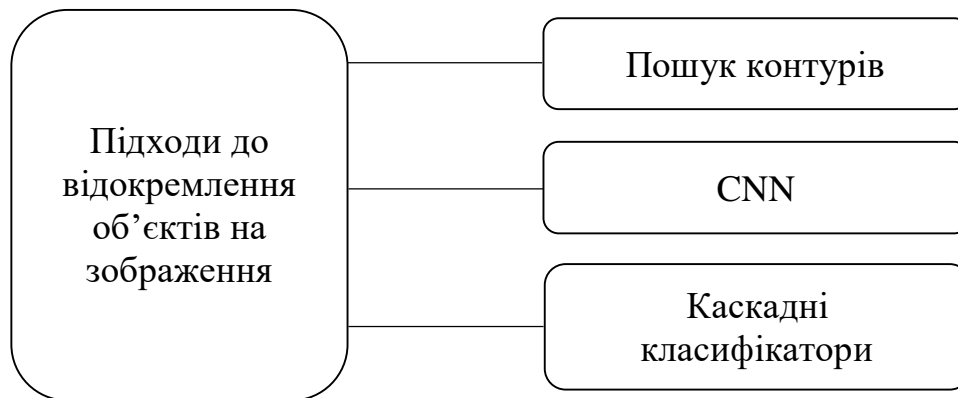


Рисунок 3.1 – Підходи виокремлення об'єктів на зображенні

Бібліотека OpenCV використовується для програмування програмного продукту за допомогою каскадних методів. Це звичайний інструмент обробки зображень і забезпечує ефективну та досить точну ідентифікацію об'єктів за допомогою каскадів волосся.

Ми визначаємо фази програмного забезпечення на основі бібліотеки OpenCV.

Першим кроком є завантаження відео на сервер. Потім необхідно створити зображення для подальшої обробки. Наступним етапом є розпізнавання об'єкта на зображенні. Потім необхідно обробити накопичені дані та візуалізувати у вигляді таблиць та графіків.

Для того, щоб методи виокремлення об'єктів на зображенні давали кращі результати, необхідно обробити та очистити зображення від різного виду пошкоджень.

Існує багато типів шумів, такі як гауссовський, «сіль та перець», пуассон, спекл-шум та інші.

Наприклад, розглянемо як впливає шум на якість зображення та як виглядає зображення після очищення від шуму (рис. 3.2).



Рисунок 3.2 – Зашумлене зображення до та після очищення

Шум присутній майже в будь-якому зображенні, часто з'являється при зборі або зберіганні даних.

Існує багато різних типів шуму, таких як гауссовий, який має добавки і незалежний від інтенсивності сигналу і Пуассона, який залежить від яскравості.

Розуміння теми зменшення шуму важливо для моделювання та його усунення. У 1992 році в різних структурах була представлена семінарська робота з ROF.

Це важливо для її здатності видаляти шум при цьому зберігаючи краї, на відміну від попередніх підходів, що використовували L^2 регуляризацію градієнта. Це досягається на прикладі регуляризації TV.

Надається наступна формула:

$$m\{\alpha \int |\nabla u| dx + \int |u - z|^2 dx\}, \quad (0.3)$$

$u \qquad \Omega \qquad 2 \Omega$

де α – позитивний параметр, що управляє рівнем видалення шуму.

Важливо враховувати різний тип пошкодження зображення. Хоча процес денационалізації входить до багатьох сучасних методів поділу, якщо зображення буде розмитим, воно завжди буде виходити за рамки звичайних методів поділу.

Розмиття зображення поширене в багатьох областях, таких як астрономічні або медичні зображення, кольорова ангиографія очного дна для візуалізації сітківки. Процес захоплення (або деконволюції) зображення – це завдання реконструкції реального зображення сприйманого звуку зображення та відновлення важливих функцій.

Математично згортка – це операція над двома функціями f і g , що виробляє функцію, яка розглядає кількість перекриття однієї функції при зміщенні її на іншу функцію.

При обробці зображення ядро складається з оператора зображення, що змінює значення пікселів зображення x залежно від значень сусідніх пікселів [30].

Оскільки зображення зберігається як набір дискретних пікселів, нам потрібна дискретна згорнута форма.

Існує три основні проблеми, що усунення розмиття:

– несліпа деконволюція: незважаючи на те, що функція поширення точок вважається відомою, це є рідкістю у реальних програмах.

Мета - відновити справжній образ, який широко вивчається;

– напівсліпа деконволюція: інформація про ядро розмитості передбачається або відома. Завдання полягає в оцінці правильних параметрів розмиття;

– деконволюція сліпого зображення: ядро зображення та розмиття є невідомими та завдання відновлення обох дуже складне.

Реєстрація зображень – це проблема, щоб об'єднати два зображення, щоб вони відповідали їхнім частинам. Модерсіцький класифікує інтенсивність методів реєстрації на два визначення: непараметричні та параметричні.

Він заснований на варіаційному методі, метою якого є знайти поле деформації для фізичних процесів, таких як викривлення.

До альтернативних методів належать основні методи, при яких певні ділянки визначають поле деформації. Можна включити нелінійні методи реєстрації в моделі змінної сегментації, але часто ці методи є параметричними (тобто методи, засновані на не змінних). Параметрична реєстрація зображення – це рядкова або афінна варіація, яка залежить від трьох і шести параметрів. Оскільки цей метод особливий, ми коротко опишемо деякі деталі, пов'язані з реєстрацією спорідненості.

Необхідно використовувати спеціальні прийоми, щоб ізолювати об'єкт від відеозображень або зображень.

Існує багато готових бібліотек і фреймворків, які можуть спростити це завдання для розробників і заощадити багато часу. Однак, щоб краще зрозуміти, як вони працюють, розглянемо, з чого починається обробка зображень. Розбиття зображення – це розбиття зображення на кілька областей із спільними властивостями. Різні методи для цього включають зниження споживання енергії, що призводить до диференціального рівняння в частинних похідних на основі розрахунку варіацій. Тоді розв'язок цього рівняння відповідає змістовному відображенню зображення.

3.2 Вимоги до програмного забезпечення

OpenCV (Open Source Computer Vision Library) – бібліотека програмного забезпечення з відкритим кодом та програмного забезпечення

для машинного навчання. OpenCV була побудований для забезпечення загальної інфраструктури для програм комп'ютерного зору та для прискорення використання машинного сприйняття в комерційних продуктах.

Бібліотека має понад 2500 оптимізованих алгоритмів, що включає в себе вичерпний набір як класичних, так і сучасних алгоритмів комп'ютерного зору та машинного навчання. Ці алгоритми можна використовувати для виявлення та розпізнавання облич, ідентифікації об'єктів, класифікації дій людини у відео, відстеження рухів камери, відстеження рухомих об'єктів, вилучення 3D-моделей об'єктів, отримання 3D-хмарних точок із стереокамер, скріплення зображень разом для отримання високої роздільної здатності зображення цілої сцени, пошук подібних зображень з баз даних зображень, виправлення червоних очей на зображеннях, зроблених за допомогою спалаху, відслідковування руху очей, розпізнавання краєвидів тощо.

У OpenCV є понад 47 тисячна користувацька спільнота та приблизно 18 мільйонів завантажень. Бібліотека широко використовується компаніями, науковими групами та державними установами.

Поряд з налагодженими компаніями, такими як Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota, які використовують бібліотеку, існує безліч стартапів, таких як Applied Minds, VideoSurf і Zeitera, які широко використовують OpenCV.

Розгорнутий OpenCV використовується для виявлення вторгнень у відео спостереження в Ізраїлі, моніторингу шахтного обладнання в Китаї, допомоги роботам у навігації та підборі предметів у Willow Garage, виявлення аварій, запуску інтерактивного мистецтва в Іспанії та Нью-Йорку, перевірка злітно-посадкових смуг на сміття в Туреччині, огляд етикетки на продуктах на заводах по всьому світу для швидкого виявлення обличчя в Японії.

Він має інтерфейси C ++, Python, Java та MATLAB та підтримує Windows, Linux, Android та Mac OS. OpenCV в основному схиляється до додатків зору в режимі реального часу та використовує інструкції MMX та SSE, коли вони доступні. Зараз активно розробляються повнофункціональні інтерфейси CUDA та OpenCL.

Існує понад 500 алгоритмів і приблизно в 10 разів більше функцій, які складають або підтримують ці алгоритми. OpenCV написано на мові C ++ і має шаблонний інтерфейс, який легко працює з контейнерами STL.

Каскад Хаара – алгоритм виявлення об’єктів на зображенні чи відео, який використовується для ідентифікації об’єктів у зображенні чи відео та заснований на концепції особливостей, запропонованих Полом Віолою та Майклом Джонсом у їхній роботі «Швидке виявлення об’єктів за допомогою посиленого каскаду простих функцій» в 2001 році.

Це підхід на основі машинного навчання, де функція каскаду навчається з безлічі позитивних і негативних образів. Потім він використовується для виявлення об’єктів на інших зображеннях.

Алгоритм має чотири етапи: вибір функції Хаара, створення інтегральних образів, навчання Adaboost, каскадні класифікатори.

Він добре відомий тим, що вміє виявляти обличчя та частини тіла на зображенні, але може бути навчений ідентифікувати практично будь-який об’єкт. Розглянемо обличчя як приклад.

Спочатку алгоритм вимагає класифікації більше позитивних зображень обличчя та негативних безликих зображень. Тоді нам потрібно отримати його властивості.

Перший крок – зібрати властивості волосся. Функція Hair враховує сусідні прямокутні області у певному місці вікна розпізнавання, підсумовує інтенсивність пікселів у кожній області та обчислює різницю між цими значеннями (рис.3.3).

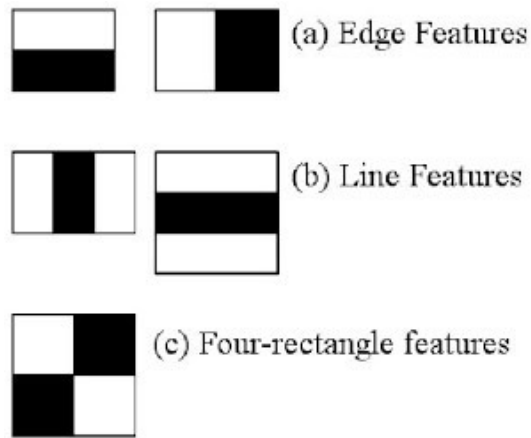


Рисунок 3.3 – Ознаки Хаара

Для цього використовуються комбіновані зображення.

Але найбільше ці малюнки не мають значення. Дивіться, наприклад, зображення нижче. Верхня лінія показує дві сторони. Перша обрана частина ніби зосередилася на тому, що область очей завжди темніше носа та щік. Другий обраний елемент спирається на те, що очі темніше, ніж перенісся. Але такі ж вікна використовуються на гранях чи де-небудь ще (рис. 3.4).

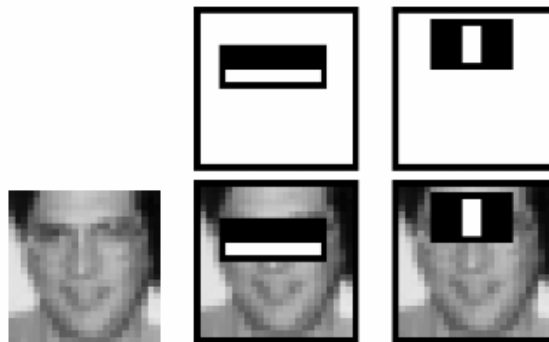


Рисунок 3.4 – Приклад розпізнавання очей за характеристиками Хаара

Тож як з 160000+ функцій вибрати найкращі функції? Це здійснюється за допомогою концепції під назвою Adaboost, яка обирає найкращі функції та навчає класифікатори, які ними користуються. Цей

алгоритм будує «сильний» класифікатор як лінійну комбінацію зважених простих «слабких» класифікаторів. Процес полягає в наступному.

Під час фази виявлення велике цільове вікно переміщується над вхідним зображенням, і властивості волосся обчислюються для кожного підрозділу зображення. Оскільки кожна ознака волосся є лише «поганою класифікацією» (якість виявлення трохи краща, ніж випадкове припущення), для опису одного об'єкта з достатньою точністю необхідна велика кількість ознак волосся, і тому вони організовані в каскадні класифікації. . надійна класифікація.

Каскадна класифікація складається з набору етапів, кожна з яких є групою вразливих студентів. Слабкі учні – це просто класифікації, які називаються пнями. Кожна фаза тренується за допомогою техніки, яка називається прискоренням. Заохочення дають можливість створити дуже точну класифікацію за допомогою середньозваженого рішення, прийнятого вразливими студентами.

Кожен рівень класифікації визначає область, визначену поточним розташуванням ковзного вікна, позитивним чи негативним. Позитивне означає, що щось знайдено, а негативне означає, що жодних об'єктів не знайдено. Якщо мітка негативна, класифікація цієї області завершується, і детектор переміщує вікно в наступне місце. Якщо мітка позитивна, класифікація переміщує область до наступного кроку. Коли остання фаза позитивно класифікує область, детектор повідомляє, що він знайде об'єкт у поточному положенні вікна.

Ці дії призначені для якнайшвидшого видалення негативних зразків. Передбачається, що більшість вікон не містять елемента, який ви шукаєте.

Щоб він працював належним чином, кожна фаза каскаду повинна мати невелику кількість помилково негативних ефектів. Якщо неправильний крок позначає елемент як негативний, класифікація буде припинена, і помилку більше не можна буде виправити. Однак кожна фаза може мати високий рівень хибнопозитивних результатів. Навіть якщо

детектор неправильно позначає непомітний об'єкт, ви можете виправити помилку, виконавши наведені нижче дії. Додавши додаткові кроки, ви можете зменшити загальну хибнопомітну ставку, але також зменшити загальну справді помітну ставку.

Навчання каскадного класифікатора вимагає набору позитивних зразків та набору негативних зображень. Ви повинні надати набір позитивних зображень із зацікавленими регіонами, які використовуються як позитивні зразки. Ви можете використовувати Image Labeler для позначення об'єктів, що цікавлять, обмежуючи полями. Image Labeler виводить таблицю для використання для позитивних зразків. Щоб досягти прийнятної точності детектора, необхідно встановити кількість етапів, тип функції та інші параметри функції (рис. 3.5).

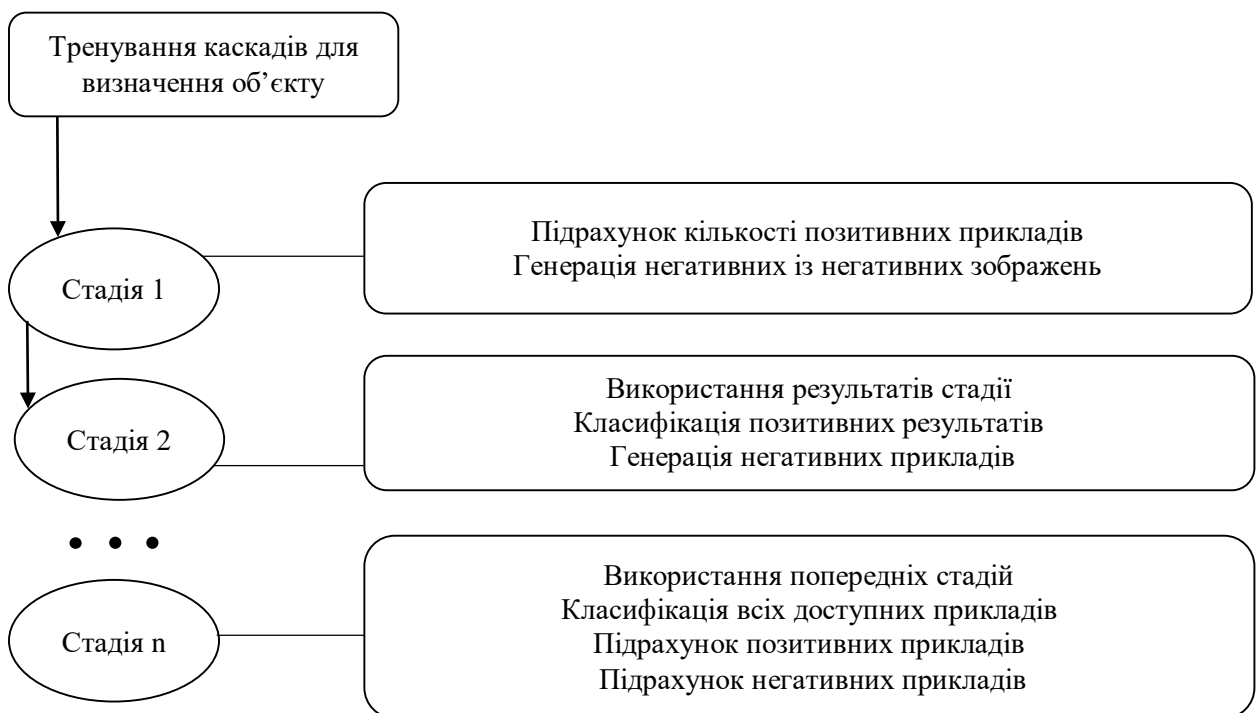


Рисунок 3.5 – Схема тренування каскадів

3.3 Програмні ресурси для створення інформаційної системи

Бібліотека OpenCV і фреймворк Tensorflow з використанням нейронних мереж були використані для створення блок-схеми для вирішення проблеми поділу об'єктів на зображення.

Розроблені сценічні проекти застосовуються до програмного забезпечення, створеного за цими технологіями. Розглянуто основні модулі бібліотеки OpenCV та схеми усунення несправностей Tensorflow.

Розглянемо схему розв'язання задачі з використанням бібліотеки OpenCV (рис. 3.6).

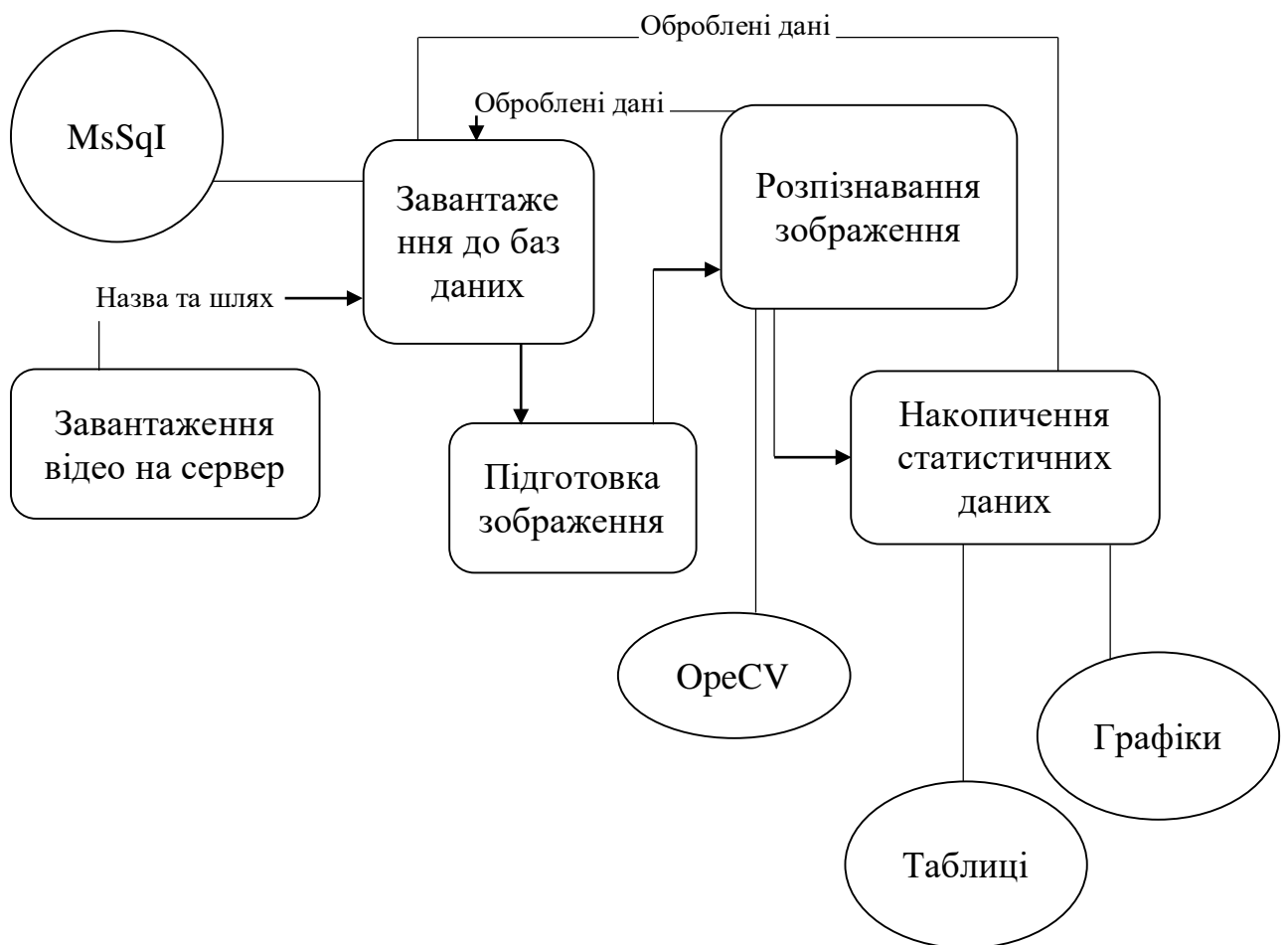


Рисунок 3.6 – Схема розв'язання задачі з використанням бібліотеки OpenCV

Спочатку потрібно завантажити зображення на сервер. MsSql використовується для зберігання даних. Потім потрібно підготувати зображення, тобто прибрати звук.

На виході ми обробляємо та відображаємо дані у вигляді графіків або таблиць. Основні модулі OpenCV описані нижче в таблиці 3.1.

Таблиця 3.1 – Модулі бібліотеки OpenCV

№	Назва	Призначення функції
1.	opencv_core	Базові структури, обчислення (математичні функції, генерація псевдовипадкових чисел, DFT, DCT, ввід/вивід XML і т.д.)
2.	opencv_highgui	Простий UI, завантаження/збереження зображень та відео
3.	opencv_ml	Методи та моделі машинного навчання (SVM, дерева прийняття рішень і т.п.)
4.	opencv_features2d	Різні дескриптори (SURF)
5.	opencv_video	Аналіз пересування та відстежування об'єктів (оптичний потік, шаблони пересування, усунення фону).
6.	opencv_objdetect	Виявлення об'єктів на зображенні (вейвлети Хаара, HOG і т.п.)
7.	opencv_calib3d	Калібровка камери, пошук стереовідповідності та елементи обробки трьохмірних даних
8.	opencv_flann	Бібліотека швидкого пошуку найближчих сусідів (FLANN).

Фреймворк Tensorflow є дуже популярним та доступним серед розробників. Він базується на використанні нейронних мереж.

В даному випадку використовувалися регіонально-згорткові нейронні

мережі (рис. А1-А3).

Розглянемо схему розв'язання задачі з використанням фреймворку Tensorflow та бібліотеки OpenCV (рис. 3.7).

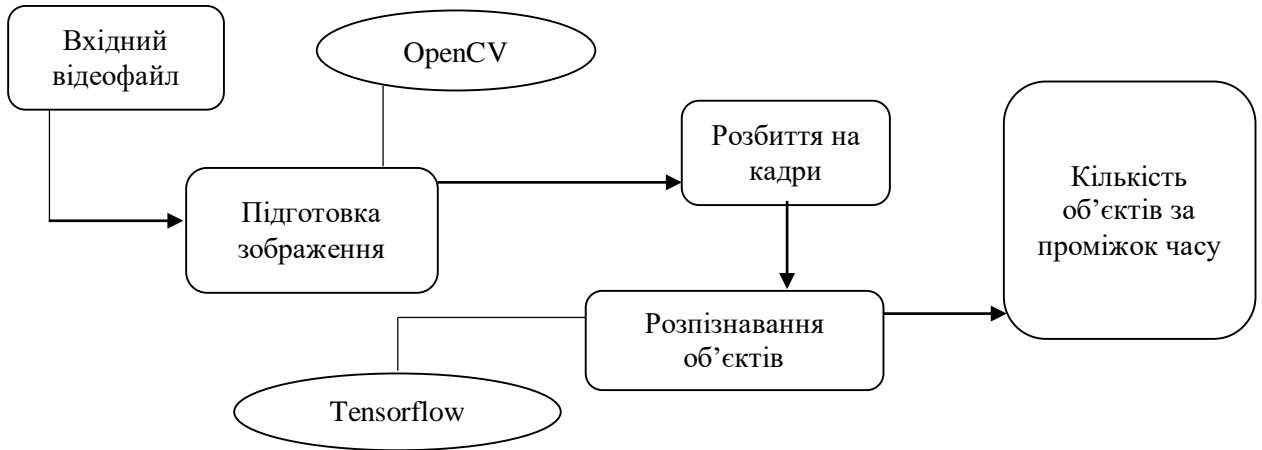


Рисунок 3.7 – Схема розв'язання задачі з використанням фреймворку Tensorflow

Першим завантажується вхідний файл. Зображення підготовлено за допомогою бібліотеки OpenCV. Фреймворк Tensorflow і нейронні мережі розпізнають об'єкти. Для навчання нейронної мережі використовувався підготовлений набір даних. Розглянемо етапи розв'язання задачі за допомогою Tensorflow (рис. 3.8).

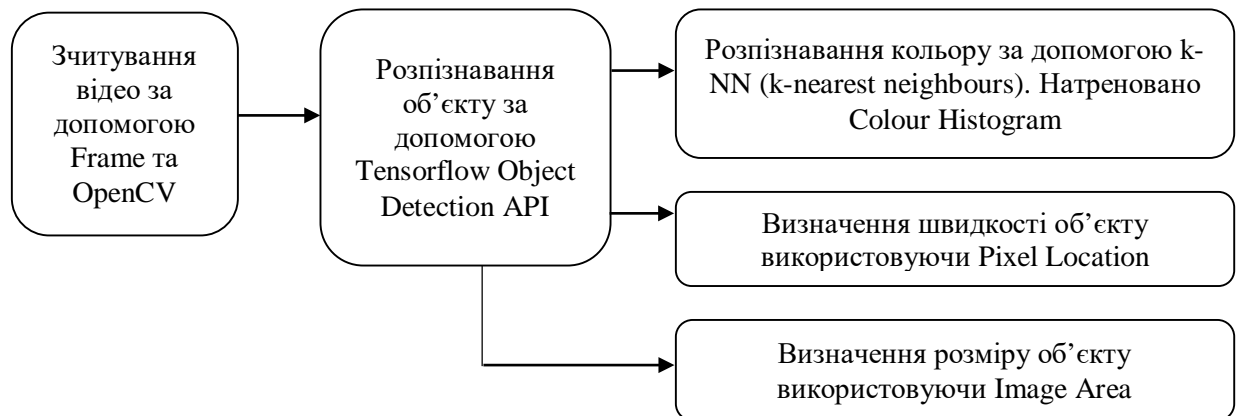


Рисунок 3.8 – Етапи рішення задачі з використанням Tensorflow

Результатом роботи програми є найкраще збіг обличчя, що відображається у відеопотоку, з обличчям з бази даних та копіювання інформації про цю людину на відео за допомогою доповненої реальності.

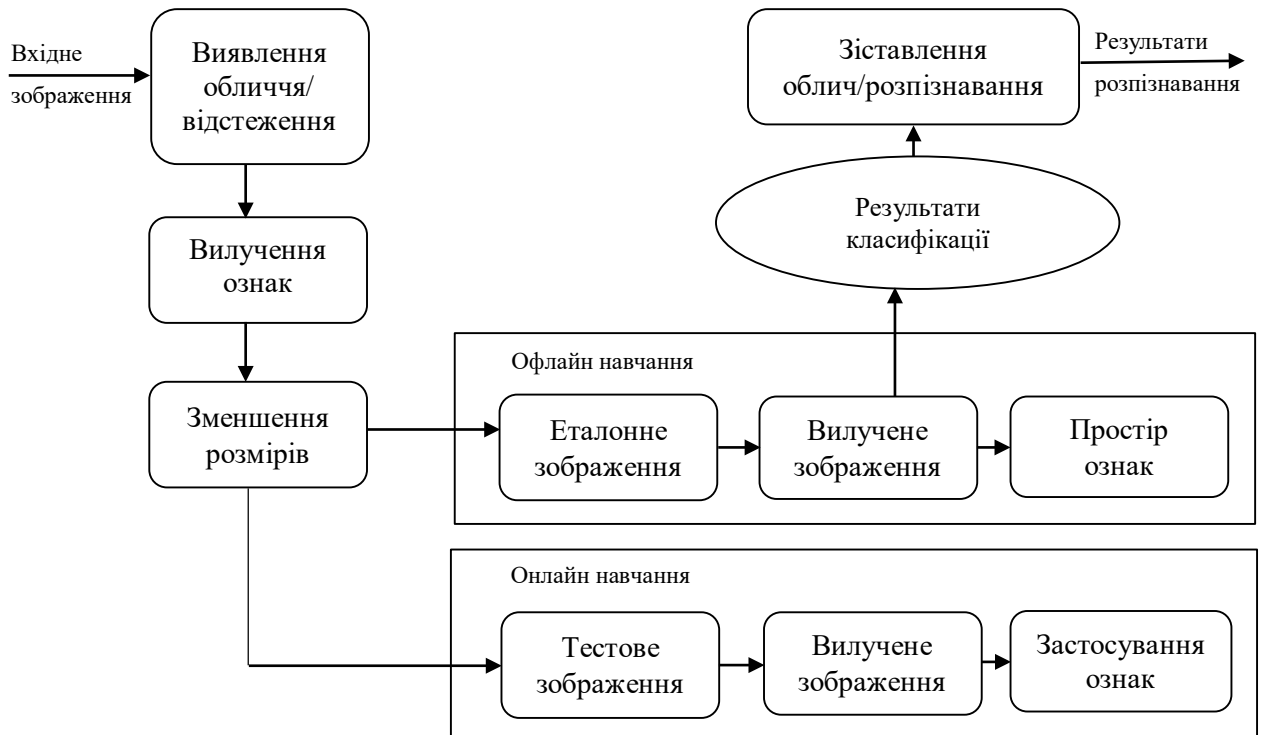


Рисунок 3.9 – Загальний процес розпізнавання обличчя людини

Для розпізнавання обличчя у програмі використовується наступний алгоритм (рис. 3.9):

- знайти всі обличчя на зображенні;
- розпізнати кожне обличчя, навіть якщо воно знаходиться під кутом або під поганим освітленням;
- визначити унікальні риси обличчя, що відрізняють одну людину від іншої, наприклад форма обличчя, розмір очей тощо;
- порівняти знайдені особливості цього обличчя з усіма людьми, обличчя яких система вже знає, щоб зрозуміти, хто зараз зображений на картинці.

Розглянемо кожний крок детальніше:

Крок 1. Пошук усіх облич.

На першому етапі алгоритму потрібно знайти всі обличчя. Для цього було використано гистограму направлених градієнтів (від англ. «Histogram of Oriented Gradients», HOG). Спочатку зображення приводиться до чорно-білої гами, тому що для пошуку обличчя дані про колір є надлишковими. Потім кожний піксель розглядається окремо і порівнюється з сусідніми. Мета цих дій – виявити, наскільки темним є поточний піксель порівняно з сусідніми. Потім малюється стрілка, що вказує, в якому напрямку зображення стає темнішим (градієнт). Таким чином, алгоритм отримує градієнтне зображення обличчя. Для кожної людини напрямок зміни кольору на обличчі від темного до світлого буде однаковим при будь-якому освітленні.

Крок 2. Оцінювання орієнтирів обличчя Після знаходження обличчя необхідно розібратися з проблемою, коли обличчя, повернуті у різні сторони, здаються комп'ютеру різними людьми.

Для розв'язання цієї задачі кожне зображення перетворюється так, щоб очі і губи завжди знаходилися в якомусь визначеному місці. Це спрощує задачу порівняння облич на наступних кроках. Для цього використовується алгоритм оцінювання орієнтирів обличчя (від англ. – «face landmarks estimation»).

Основна ідея полягає в тому, що ми помічаємо 68 спеціальних точок (так званих орієнтирів), які знаходяться на кожному обличчі - верхня частина брів, зовнішня точка кожного ока, внутрішня точка кожної брови і так далі. Потім алгоритм машинного навчання знайшов ці 68 точок на кожній поверхні.

Якщо ми знаємо, де знаходяться очі, а де рот, ми можемо обертати, вимірювати та переміщувати зображення так, щоб очі та рот були якомога центральнішими. Для цього використовувалися афінні модифікації.

Крок 3. Кодування обличчя.

Потрібно знайти спосіб зняти кілька основних вимірів з кожного обличчя, які можна було б порівняти з найближчими відомими вимірами та знайти найбільш схоже обличчя.

Наприклад, можна змінити розмір кожного вуха, відстань між очима, довжину носа тощо. Проблема в тому, що вимірювання, які здаються очевидними для людей (наприклад, колір очей), насправді не мають сенсу для комп'ютера, який розглядає окремі пікселі зображення.

Один із способів – дати комп'ютеру те, що йому потрібно. Алгоритми глибокого навчання самостійно визначають, які частини поверхні слід вимірювати, а які ні.

Рішення полягає у створенні згорткової нейронної мережі, яка може створити 128 вимірів для кожної поверхні. 128 вимірів кожної грані називають картою. Ідея перетворення сеансу необроблених даних, наприклад зображення, у список згенерованих комп'ютером чисел є важливою для машинного навчання.

Процес навчання згорткової нейронної мережі для отримання карти обличчя (кодування обличчя) вимагає великої кількості даних і великої обчислювальної потужності. Тому використовується модель на основі кількох наборів даних за різні тижні. Залишилося лише пройти зображення обличчя закінченої нейронної мережі та отримати 128 вимірювань для кожного обличчя.

Крок 4. Пошук інформації людини за кодуванням.

На останньому етапі потрібно знайти людину в базі відомих для розробленої програми людей, виміри обличчя якої найближчі до отриманих даних у реальному часі. Це робиться за допомогою класифікатора, що вимірює відстань між значеннями і визначає, наскільки вони схожі.

Для розпізнавання обличчя було обрано згорткову нейронну мережу архітектури ResNet34 [Residual Network (дослівно — «залишкова мережа»)]. ResNet має наступні переваги над базовими нейронними мережами, цим самим усуваючи їх основні недоліки:

ResNet – глибинна нейронна мережа.

Згідно з загальною теоремою апроксимації ми знаємо, що мережа прямого поширення з одним шаром може відтворити будь-яку функцію. Проте

шар має бути великим, через що нейронна мережа може бути схильною до перенавчання, що дасть погані результати. Для розв'язання цієї проблеми мережева архітектура має бути більш глибокою. ResNet може бути побудована як зі 100, так і з 1000 шарів.

ResNet долає проблему зникнення градієнта.

Глибокі мережі важко тренувати через проблему втрати градієнта - оскільки градієнт повертається до попередніх шарів, зміна множника може зменшити градієнт. Як результат, у міру поглиблення мережі її продуктивність може швидко насититися або знизитися.

Через наявність залишкових блоків ResNet мережа може ігнорувати градієнт деяких підмереж і просто переносити градієнт з верхніх шарів на нижні без будь-яких змін. Це означає, що ResNet може лише перемістити перехід з останнього шару, наприклад шару 151, безпосередньо до першого шару.

Для реалізації алгоритму обрано високорівневу мову програмування загального призначення Python. Для розпізнавання облич людей було використано бібліотеку `face_recognition`. Вона надає зручний API, який, у свою чергу, використовує бібліотеки `OpenFace`, `dlib`, `OpenCV`. Також ця бібліотека представляє методи роботи з якісно навченими моделями для визначення `face landmarks` і нейронною мережею ResNet34.

Розглянемо мережу ResNet34 (нейромережа на архітектурі ResNet з 34 шарами).

У першому шарі ResNet використовує згортку 7×7 з кроком 2 для зменшення вхідного сигналу у два рази (від англ. «downs-?mpling»), подібно шару агрегування (від англ. «pooling»). Потім йдуть три блоки ідентичності (або залишкові шари), після чого кількість пікселів вхідного зображення знову зменшується у два рази. Шар семплінгу (зменшення зображення) також є згортковим шаром, але не з'єднується з блоком ідентичності. Ці блоки архітектури повторюються аналогічним чином кілька разів.

Останній шар – це шар усередненого агрегування (від англ. «average

pooling»), який створює 1000 карт властивостей (для даних ImageNet) і середнє значення для кожної карти об'єктів. Результатом буде 1000-мірний вектор, який потім подається безпосередньо в шар Softmax, тому він повнозв'язний (від англ. «fully connected»). Основне призначення функції softmax полягає в перетворенні (ненормалізованого) виводу одиниць K (який, наприклад, представлений у вигляді вектора K елементів) повнозв'язного шару у розподіл ймовірностей (нормалізований вихід), який часто представлений як вектор K елементів, кожен з яких знаходиться між 0 та 1 (ймовірність), а сума усіх цих елементів дорівнює 1 (розподіл ймовірності).

У випадку задачі класифікацією i -й елемент вектора, що створюється функцією softmax, відповідає ймовірності входу мережі, що належить до i -го класу.

ВИСНОВКИ

У рамках кваліфікаційної роботи було розглянуто задачі та алгоритми для розпізнавання та відслідковування об'єктів і реалізовано інформаційну систему для вирішення проблеми класифікації цих об'єктів.

Застосування нейромережі на архітектурі ResNet з відповідними налаштуваннями для детектування та розпізнавання обличчя людини у відеопотоці з доповненою реальністю є хорошим вибором – цей метод інваріантний до масштабування, зміни сцени, поворотів голови, зміни освітленості, аксесуарів та зміни емоцій.

Експерименти показують стабільні результати. За перешкод, а саме, коли закривається частина обличчя, яка є визначальною для розпізнавання, програма не може визначити, кого зображено, тому що не може порівняти особливі ознаки наявного зображення обличчя та обличчя з бази даних.

Система є платформою для подальших розробок. Зокрема, планується проведення експериментальних досліджень з використанням інших методів розпізнавання обличчя у відеопотоці та виконання порівняльного аналізу отриманих результатів, а також створення більш зручного графічного інтерфейсу програми та адаптація для мобільної версії.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Bilonoh, B., Bodyanskiy, Y., Kolchygin, B., & Mashtalir, S. (2021, May). Tunable Activation Functions for Deep Neural Networks. In *International Scientific Conference “Intellectual Systems of Decision Making and Problem of Computational Intelligence”* (pp. 624-633). Springer, Cham.
2. Bilonoh, B., & Mashtalir, S. (2020, August). Parallel multi-head dot product attention for video summarization. In *2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP)* (pp. 158-162). IEEE.
3. Bodyanskiy, Y. V., Tyshchenko, O. K., & Mashtalir, S. V. (2019, June). Fuzzy clustering high-dimensional data using information weighting. In *International Conference on Artificial Intelligence and Soft Computing* (pp. 385-395). Springer, Cham.
4. Mashtalir, S., & Mashtalir, V. (2020). Spatio-temporal video segmentation. In *Advances in Spatio-Temporal Segmentation of Visual Data* (pp. 161-210). Springer, Cham.
5. Kinoshenko, D., Kobylin, O., Mashtalir, S., & Stolbovyi, M. (2019, March). Metric video retrieval speedup by irrelevant data elimination. In *Eleventh International Conference on Machine Vision (ICMV 2018)* (Vol. 11041, pp. 176-183). SPIE.
6. Mashtalir, S. V., Stolbovyi, M. I., & Yakovlev, S. V. (2019). Clustering video sequences by the method of harmonic k-means. *Cybernetics and Systems Analysis*, 55(2), 200-206.
7. Машталир, С. В., Столбовой, М. И., & Яковлев, С. В. (2019). Кластеризация последовательностей видеоданных на основе гармонических k-средних. *Кибернетика и системный анализ*.
8. Mashtalir, S. V., Stolbovoi, M. I., & Yakovlev, S. V. (2019). Hybrid Approach to Clustering Various Lengths Video. *Journal of Automation and Information Sciences*, 51(3).

9. Mashtalir, S., Mikhnova, O., & Stolbovyi, M. (2018, August). Sequence matching for content-based video retrieval. In *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)* (pp. 549-553). IEEE.
10. Mashtalir, S., Stolbovyi, M., & Mikhnova, O. (2019). Multidimensional sequence clustering with adaptive iterative dynamic time warping. *International Journal of Computing*, 18(1), 53-59.
11. Mashtalir, S., Mashtalir, V., & Stolbovyi, M. (2018, August). Representative based clustering of long multivariate sequences with different lengths. In *2018 IEEE second international conference on Data Stream Mining & Processing (DSMP)* (pp. 545-548). IEEE.
12. Kinoshenko, D., Mashtalir, S., Shlyakhov, V., & Stolbovyi, M. (2018, January). Video shots retrieval with use of pivot points. In *International Conference on Computer Science, Engineering and Education Applications* (pp. 102-111). Springer, Cham.
13. Information Resources Management Association (Ed.). (2018). *Computer Vision: Concepts, Methodologies, Tools, and Applications*. IGI Global.
14. Машталир, С. В., & Столбовой, М. И. (2018). Адаптивные матричные модели в задаче контроля потоков видео. *Радіоелектроніка, інформатика, управління*, (4 (47)), 188-194.
15. Богучарский, С. И., Машталир, С. В., & Столбовой, М. И. (2018). Быстрое обнаружение изменения свойств многомерных временных рядов на основе идентификационного подхода к ансамблю моделей. *Системи обробки інформації*, (3), 74-78.
16. Mashtalir, S. V., & Stolbovyi, M. I. (2018). ADAPTIVE MATRIX MODELS IN THE VIDEO STREAMS CONTROL PROBLEM. *RADIO ELECTRONICS COMPUTER SCIENCE CONTROL*, (4), 188-194.
17. Hu, Z., Mashtalir, S. V., Tyshchenko, O. K., & Stolbovyi, M. I. (2018). Clustering matrix sequences based on the iterative dynamic time deformation

procedure. *International Journal of Intelligent Systems and Applications*, 10(7), 66-73.

18. Hu, Z., Mashtalir, S. V., Tyshchenko, O. K., & Stolbovyi, M. I. (2017). Video shots' matching via various length of multidimensional time sequences. *International Journal of Intelligent Systems and Applications*, 9(11), 10.

19. Oleg, K., Sergii, M., & Mykhailo, S. (2017, October). Video clustering via multidimensional time-series analysis. In *Proceedings of the 9th International Conference on Information Management and Engineering* (pp. 60-63).

20. Yarahmadi, A., Creemers, M., Qabbaah, H., Vanhoof, K., Palagin, O., Romanov, V., ... & Kurkchiyan, V. (2017). Unraveling Bi-Lingual Multi-Feature Based Text Classification: A Case Study. *International Journal of Information Theories and Applications*, 24(4), 3-18.

21. Mashtalir, S., Mashtalir, V., & Stolbovyi, M. (2017). Video shot boundary detection via sequential clustering. *International Journal "Information Theories and Applications*, 24(1), 50-59.

22. Mashtalir, S., & Mikhnova, O. (2017). Detecting significant changes in image sequences. In *Multimedia Forensics and Security* (pp. 161-191). Springer, Cham.

23. Богучарский, С. И., & Машталир, С. В. (2015). Иерархическая агломеративная кластеризация в базах данных мультимедиа. *Electrotechnic and Computer Systems*, (19 (95)), 239-242.

24. Mashtalir, S., & Mikhnova, O. (2014). Key frame extraction from video: framework and advances. *International Journal of Computer Vision and Image Processing (IJCVIP)*, 4(2), 68-79.

25. Богучарский, С. И., Каграманян, А. Г., & Машталир, С. В. (2014). Иерархическая агломеративная кластеризация изображений в больших базах данных. *Системи обробки інформації*, (8), 93-97.

26. Calic, J., & Izquierdo, E. (2002, April). Efficient key-frame extraction and video analysis. In *Proceedings. International Conference on Information Technology: Coding and Computing* (pp. 28-33). IEEE.

27. Bodyanskiy, Y., Kinoshenko, D., Mashtalir, S., & Mikhnova, O. (2012). On-line video segmentation using methods of fault detection in multidimensional time sequences. *International Journal of Electronic Commerce Studies*, 3(1), 1-20.

28. Гончаренко, М. О., & Машталир, С. В. (2012). Детектирование изменений сцены в потоке видеоданных. *Электротехнические и компьютерные системы*, (7), 143-147.

29. Ведмедь, А. Г., Машталир, С. В., & Сакало, Е. С. (2010). Восстановление изображений с использованием анализа главных и независимых компонент. *Системы обработки информации*, (6), 66-72.

30. Bodyanskiy, Y., Grimm, P., Mashtalir, S., & Vinarski, V. (2010, July). Fast training of neural networks for image compression. In *Industrial Conference on Data Mining* (pp. 165-173). Springer, Berlin, Heidelberg.