Міністерство освіти і науки України

Національний університет «Львівська політехніка»

Кафедра електронних обчислювальних машин



**ЗВІТ**

З перед дипломної практики

Студента IV курсу групи КІ-406

Ярмола Юрія Юрійовича

**Тема:** “Методи та комп'ютерні засоби аналізу зображень у відеопотоці”

**Термін практики** з «2» вересня 2024р. по «20» вересня 2024р.

**Керівники практики:**

від кафедри \_\_\_\_\_\_\_\_\_\_\_\_\_\_ асистент каф. ЕОМ Гузинець Н. В.

(підпис) (наук.ст, вч.звання, прізвище, ім’я, по-батькові)

**ОЦІНКА** \_\_\_\_\_\_\_\_\_

**Дата** \_\_\_\_\_\_\_\_\_\_\_\_\_

**ЗМІСТ**

[1. ОГЛЯД ТА ОБГРУНТУВАННЯ ПРЕДМЕТНОЇ ОБЛАСТІ 3](#_Toc177585831)

[1.1 Первинна обробка відеозображень 4](#_Toc177585832)

[1.2 Відстеження об'єктів 6](#_Toc177585833)

[1.3 Машинне навчання для розпізнавання об'єктів 10](#_Toc177585834)

[1.4 Актуальність аналізу відеопотоків 14](#_Toc177585835)

[1.5 Сучасні технології 18](#_Toc177585836)

[2. ВИСНОВКИ 25](#_Toc177585837)

[СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 26](#_Toc177585838)

# 1. ОГЛЯД ТА ОБГРУНТУВАННЯ ПРЕДМЕТНОЇ ОБЛАСТІ

Сучасне суспільство стикається з великим потоком відео-даних, обсяг яких постійно зростає. Це пояснюється розвитком цифрових технологій, широким використанням мобільних пристроїв, камер спостереження, а також соціальних мереж, де щоденно генерується велика кількість відеоконтенту. У зв'язку з цим зростає потреба в автоматизованих засобах обробки та аналізу відеоданих для ефективного управління інформацією та її використання в різних сферах діяльності.

Сьогодні автоматизована обробка відео набуває все більшого значення в різних галузях, починаючи від систем безпеки і відеоспостереження, закінчуючи медичними дослідженнями та розважальними технологіями. Наприклад, у системах відеоспостереження автоматичний аналіз відеопотоків допомагає ідентифікувати підозрілу поведінку, розпізнавати обличчя та виявляти потенційні загрози в режимі реального часу. У медицині технології аналізу відео використовуються для автоматизованого виявлення патологій на основі відеоданих, таких як медичні сканування або результати ендоскопічних процедур. У розважальній індустрії, особливо в кіноіндустрії та відеоіграх, відеоаналіз допомагає створювати складні анімації та симуляції, що відображають реальну поведінку об'єктів та людей.

Ключовими аспектами у процесі аналізу відеопотоків є обробка відеозображень, відстеження об'єктів у кадрі, а також розпізнавання та класифікація об'єктів. Обробка відеозображень передбачає покращення якості зображень, фільтрацію шумів, виділення контурів об'єктів та інші методи, що допомагають підготувати дані до подальшого аналізу. Ці операції особливо важливі для забезпечення точності та надійності подальшої обробки даних. Відстеження об'єктів є важливим етапом аналізу, який дозволяє ідентифікувати рухомі об'єкти в кадрі та відслідковувати їх траєкторію протягом відеопотоку. Ця задача має ключове значення у багатьох застосуваннях, включаючи автономні транспортні системи, де відстеження об'єктів забезпечує безпеку руху.

## 1.1 Первинна обробка відеозображень

Первинна обробка відеозображень є основою для подальшого аналізу та обробки даних комп'ютерними системами. Її метою є підготовка зображень до наступних етапів обробки, забезпечуючи точність і надійність аналізу. Завдяки первинній обробці відеозображень можна отримати необхідні характеристики зображень, такі як контури, кольорові переходи, а також структурні особливості, які необхідні для виконання складніших алгоритмів. Основні типові завдання для первинної обробки включають в себе такі важливі етапи:

* Геометричні перетворення – це набір операцій, які дозволяють змінювати просторову структуру зображення. Сюди входять такі операції, як обертання зображення на певний кут, зміна масштабу (збільшення або зменшення), а також переворот зображення по одній з осей. Ці методи часто використовуються для приведення зображень до однакового формату або для їх корекції в разі спотворення.
* Колірна корекція – це процес зміни кольорових характеристик зображення з метою поліпшення його якості або адаптації до подальшого аналізу. Типовими операціями є зміна яскравості і контрасту, що дозволяє виділяти важливі деталі на зображенні. Перетворення між різними колірними просторами, наприклад, з RGB у YCbCr або HSV, дозволяє краще аналізувати різні аспекти зображення, такі як насиченість або яскравість.
* Порівняння зображень – це важливий етап для знаходження подібностей і відмінностей між двома і більше зображеннями. Це може використовуватись для виявлення змін у серії відеокадрів або для ідентифікації конкретних об'єктів на зображенні. Як окремий випадок можна згадати знаходження кореляції між зображенням і зразком, що застосовується в різних системах безпеки та аутентифікації, наприклад, у детекторах банкнот або системах розпізнавання облич.
* Комбінування зображень – це процес поєднання кількох зображень для отримання нового зображення з покращеними характеристиками або для створення візуальних ефектів. Це використовується в обробці відео для покращення якості кадрів або створення складних візуалізацій, наприклад, для графічного дизайну або кіновиробництва.
* Інтерполяція і згладжування – використовуються для покращення якості зображення, зменшення шумів та артефактів. Інтерполяція дозволяє підвищити роздільну здатність зображення, додаючи нові пікселі на основі існуючих, тоді як згладжування допомагає усунути різкі переходи і шуми, що часто зустрічаються при зйомці в умовах низької освітленості або при компресії відеоданих.
* Сегментація зображень – це поділ зображення на окремі області або сегменти, які мають певні спільні характеристики. Це важливий крок для подальшого аналізу, оскільки дозволяє спростити візуальні дані, виділивши найбільш важливі частини зображення, такі як об'єкти або фони. Сегментація також використовується для зменшення обсягу даних, які передаються по каналах зв'язку, що є важливим при передачі відео в реальному часі.
* Редагування та ретушування – ці методи дозволяють вручну або автоматично виправляти певні частини зображення. Вони широко застосовуються у фото редагуванні , відео монтажі та медіа індустрії для покращення естетичної якості зображень або виправлення дефектів, таких як небажані об'єкти або спотворення кольорів.
* Розширення динамічного діапазону (HDR) – це техніка, яка полягає у поєднанні кількох зображень, знятих з різною експозицією, для створення одного зображення з розширеним динамічним діапазоном. Це дозволяє зберегти більше деталей як в світлих, так і в темних ділянках зображення, що особливо важливо для обробки відео з високою контрастністю.
* Компенсація втрати різкості – включає використання методів, таких як нерізке маскування, для підвищення чіткості зображення. Це може бути корисним при обробці зображень з поганою якістю фокусування або при відновленні деталей на зображеннях, які були дещо розмиті внаслідок руху або оптичних ефектів.

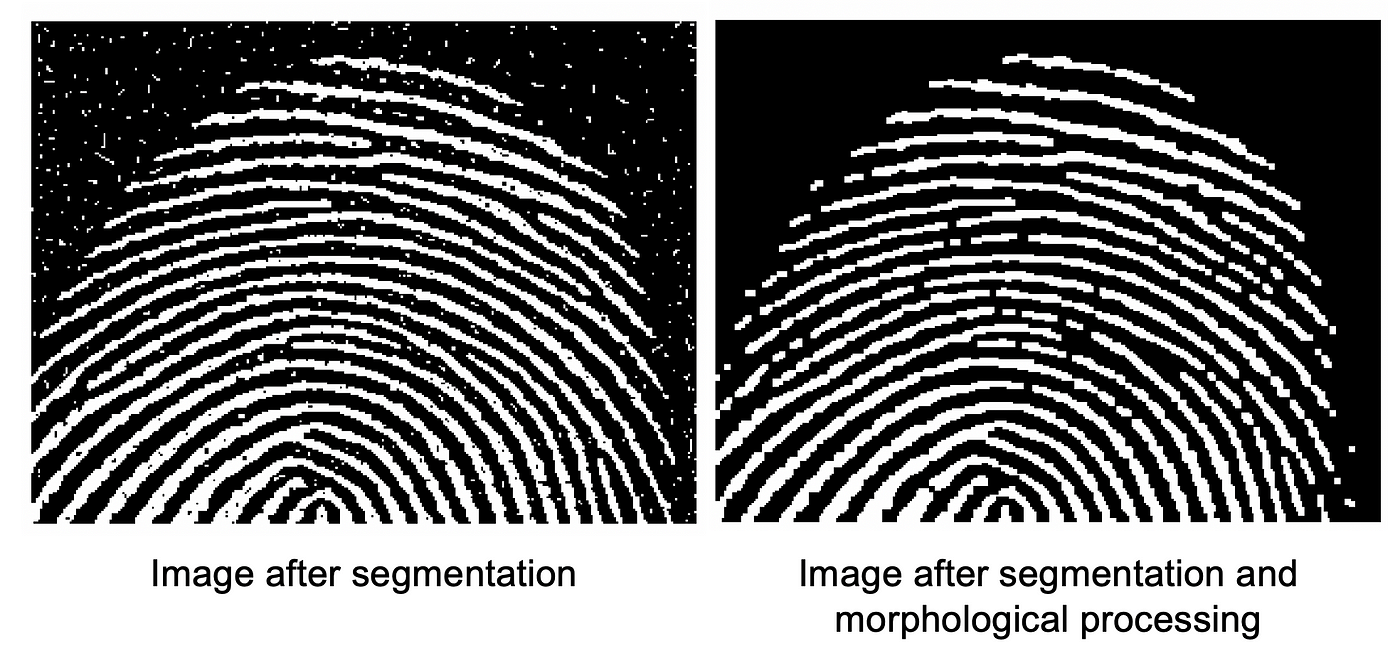


Рис. 1.1.1 – Приклад застосування алгоритмів для покращення фото

## 1.2 Відстеження об'єктів

Відстеження об'єктів у відеопотоці є однією з ключових задач у багатьох сферах, таких як системи безпеки, автономні транспортні засоби, спортивні аналітичні системи та інші області, де необхідно відстежувати рухомі об'єкти в кадрі. Для реалізації цієї задачі використовуються різні алгоритми відстеження, які базуються на різних математичних і комп'ютерних підходах. Розглянемо основні алгоритми відстеження об'єктів, їх принципи роботи та ключові характеристики.

* **Алгоритми на основі оптичного потоку (Optical Flow)**

Оптичний потік є одним із найбільш популярних підходів для відстеження об'єктів. Оптичний потік визначає вектор зміщення кожного пікселя між двома послідовними кадрами, що дозволяє зрозуміти, як рухаються об'єкти в просторі. Основною ідеєю оптичного потоку є аналіз зміни інтенсивності пікселів у кадрах для визначення напрямку і швидкості руху.

* Принцип роботи: Оптичний потік ґрунтується на припущенні, що яскравість пікселя залишається постійною під час руху, тобто інтенсивність світла в пікселі не змінюється між двома кадрами, а змінюється лише його положення. Це дозволяє визначати вектор переміщення пікселів та, відповідно, траєкторію руху об'єкта.
* Приклади методів: Метод Лукаса-Канаде (Lucas-Kanade), метод Фарнебека (Farneback).

Оптичний потік забезпечує точне відстеження руху, але має певні обмеження, зокрема в умовах швидкого руху об'єктів або за наявності великої кількості шумів у кадрі.

* **Кореляційні фільтри (Correlation Filters)**

Кореляційні фільтри використовуються для відстеження об'єктів шляхом порівняння зображення з шаблоном, який відповідає об'єкту. Цей метод зручний для відстеження об'єктів в умовах реального часу завдяки своїй швидкості та ефективності.

* Принцип роботи: Алгоритм створює фільтр (шаблон), що відповідає об'єкту в початковому кадрі, після чого використовує його для пошуку об'єкта в наступних кадрах шляхом кореляції між шаблоном і новими зображеннями. Кореляційний фільтр працює за допомогою порівняння схожості між об'єктом і фоном.
* Приклади методів: MOSSE (Minimum Output Sum of Squared Error), KCF (Kernelized Correlation Filter).

Кореляційні фільтри є швидкими і ефективними для реального часу, але можуть мати проблеми при зміні масштабу або форми об'єкта.

* **Алгоритми на основі баєсового підходу: Kalman Filter і Particle Filter**

Ці алгоритми використовують баєсову статистику для прогнозування положення об'єкта на основі його попереднього стану. Їх основною особливістю є здатність передбачати наступне положення об'єкта навіть у випадках, коли його частково або повністю закривають інші об'єкти.

* Kalman Filter: Це один із найпоширеніших алгоритмів для відстеження об'єктів, який базується на лінійних гауссових моделях. Він використовує прогноз і корекцію для оцінки стану об'єкта (положення, швидкість) в кожному кадрі.
  + Принцип роботи: Калманів фільтр використовує два кроки: прогнозування та оновлення. Спочатку він прогнозує наступне положення об'єкта на основі попередніх вимірювань, після чого коригує це прогнозоване положення на основі нових даних із відео.
  + Недоліки: Добре працює для лінійних систем, але може бути неточним у складних сценаріях з нелінійними рухами.
* Particle Filter: Використовує множину частинок для представлення ймовірного стану об'єкта. На відміну від Калмана, цей метод добре працює з нелінійними та складними моделями руху.
  + Принцип роботи: Particle Filter відстежує об'єкт, генеруючи багато можливих траєкторій руху (частинок), які відповідають різним можливим сценаріям переміщення об'єкта. Частинки, які найкраще відповідають новим даним із відео, зберігаються, а інші видаляються.
  + Недоліки: Висока обчислювальна складність, особливо при великій кількості об'єктів або складному фоні.
* **Методи на основі машинного навчання та глибокого навчання (Deep Learning)**

Сучасні підходи до відстеження об'єктів активно використовують глибокі нейронні мережі (Deep Neural Networks, DNN). Ці методи забезпечують високу точність та здатні враховувати складні зміни в структурі та текстурі об'єктів під час їх руху.

* Принцип роботи: Методи машинного навчання, такі як Convolutional Neural Networks (CNN), використовуються для виділення ознак об'єкта в кожному кадрі. Для відстеження об'єкта використовуються Recurrent Neural Networks (RNN) або Long Short-Term Memory (LSTM), які дозволяють запам'ятовувати попередні стани об'єкта і прогнозувати його положення в майбутньому.
* Приклади методів: GOTURN (Generic Object Tracking Using Regression Networks), DeepSORT (Deep Simple Online and Realtime Tracking), Siamese Networks.

Основною перевагою методів на основі глибокого навчання є їх здатність працювати з великими обсягами даних та адаптуватися до складних сценаріїв, таких як зміна освітлення, швидкий рух, часткове або повне закриття об'єкта. Однак вони потребують великих обчислювальних ресурсів і часу на тренування моделей.

## 1.3 Машинне навчання для розпізнавання об'єктів

Машинне навчання є ключовою технологією в сучасних системах аналізу зображень, особливо коли мова йде про розпізнавання об'єктів у відеопотоці. Зі зростанням обсягів даних та складністю завдань, традиційні методи обробки зображень часто виявляються недостатніми. Сучасні підходи, засновані на алгоритмах машинного навчання, дозволяють створювати системи, здатні автоматично навчатися та вдосконалювати свої результати без необхідності ручного програмування. У цій галузі особливу роль відіграють методи глибокого навчання (Deep Learning), які дозволяють обробляти великі масиви даних і забезпечують високу точність розпізнавання об'єктів.

**Основи машинного навчання в аналізі зображень**

Машинне навчання (Machine Learning) — це підхід, який дозволяє комп'ютерним системам навчатися на основі даних та робити прогнози або приймати рішення без явного програмування. У контексті аналізу зображень, машинне навчання дозволяє створювати моделі, які можуть "бачити" та "розуміти" візуальну інформацію, подібно до того, як це робить людина. Алгоритми машинного навчання працюють шляхом обробки великої кількості зображень або відео для виявлення патернів та регулярностей, які потім використовуються для класифікації або розпізнавання об'єктів у нових відеопотоціях.

Алгоритми машинного навчання, які застосовуються для аналізу зображень у відеопотоці, можна поділити на кілька категорій:

* **Класифікаційні алгоритми**: використовуються для визначення, до якого класу належить об'єкт на зображенні (наприклад, автомобіль, пішохід, тварина).
* **Алгоритми сегментації зображень**: дозволяють виділити окремі об'єкти або області на зображенні (наприклад, сегментація дорожніх знаків або транспортних засобів).
* **Алгоритми відстеження об'єктів**: використовуються для визначення траєкторії руху об'єкта у відеопотоці.
* **Регресійні алгоритми**: застосовуються для передбачення розмірів або положення об'єктів у кадрі (наприклад, прогнозування координат об'єкта).

Глибоке навчання та його роль у розпізнаванні об'єктів

Глибоке навчання (Deep Learning) є підмножиною машинного навчання і базується на використанні глибоких нейронних мереж, які складаються з багатьох шарів (звідси і назва "глибокі"). Ці мережі можуть автоматично виділяти ознаки зображень або відео, що робить їх надзвичайно потужними у задачах розпізнавання об'єктів.

Одним з основних інструментів глибокого навчання у сфері аналізу зображень є згорткові нейронні мережі (Convolutional Neural Networks, CNN). CNN використовуються для автоматичної обробки зображень, виділення ознак та розпізнавання об'єктів.

Згорткові нейронні мережі (CNN)

CNN — це клас глибоких нейронних мереж, які спеціально розроблені для обробки візуальних даних. Вони стали стандартом для обробки зображень та відео завдяки здатності автоматично виявляти важливі ознаки зображень, такі як контури, текстури, форми та об'єкти.

Основні компоненти CNN:

* Згорткові шари (Convolutional layers): ці шари використовуються для виділення локальних ознак на зображенні. Кожен згортковий шар містить фільтри, які застосовуються до вхідного зображення для виділення певних патернів.
* Шари підсумовування (Pooling layers): зменшують розмір вихідних даних, видаляючи несуттєву інформацію і зберігаючи важливі ознаки.
* Повнозв'язні шари (Fully connected layers): об'єднують всі виділені ознаки для того, щоб класифікувати об'єкт або виконати інше завдання.

Згорткові нейронні мережі стали надзвичайно успішними завдяки своїй здатності автоматично адаптуватися до різноманітних завдань, пов'язаних з аналізом зображень. Вони застосовуються для розпізнавання облич, автомобілів, пішоходів, різних об'єктів на вулиці тощо.

Рекурентні нейронні мережі (RNN) та аналіз тимчасових залежностей

Для обробки відеопотоці, де важливу роль відіграють тимчасові залежності між кадрами, часто використовуються рекурентні нейронні мережі (Recurrent Neural Networks, RNN). RNN дозволяють моделі зберігати інформацію про попередні кадри, що особливо корисно для відстеження об'єктів, коли їх рух є ключовим фактором.

RNN мають унікальну властивість — здатність "пам'ятати" попередні стани системи, що дозволяє їм обробляти послідовності даних. Однак стандартні RNN мають проблеми із "забуванням" інформації на довгих відрізках часу, тому для покращення роботи часто використовуються довго-короткочасні пам'яті (LSTM, Long Short-Term Memory) та GRU (Gated Recurrent Units). Ці моделі дозволяють ефективніше працювати з тимчасовими залежностями у відео.

Сімейства алгоритмів для розпізнавання об'єктів

У сучасних системах аналізу зображень на основі машинного навчання активно використовуються різні алгоритми для розпізнавання об'єктів. Деякі з найбільш поширених алгоритмів включають:

1. YOLO (You Only Look Once): алгоритм розпізнавання об'єктів у реальному часі, який виконує прогноз для всього зображення одним проходом, що забезпечує високу швидкість роботи.
2. Faster R-CNN (Region-based Convolutional Neural Networks): поєднує CNN для виділення ознак зображень із регіональними пропозиціями для розпізнавання об'єктів у різних областях зображення.
3. SSD (Single Shot Multibox Detector): алгоритм, який поєднує високу швидкість і точність, що робить його популярним для розпізнавання об'єктів у відеопотоці.
4. RetinaNet: використовує особливу функцію втрат під назвою "Focal Loss", що дозволяє краще працювати з незбалансованими даними, де на більшості зображень присутні лише один або два об'єкти.

Ці алгоритми забезпечують різні підходи до розпізнавання об'єктів, при цьому їх вибір залежить від конкретних вимог до продуктивності та точності.

Штучний інтелект у відеоаналітиці

Застосування штучного інтелекту (ШІ) у відеоаналітиці відкриває нові горизонти для автоматизації процесів аналізу відео. Завдяки поєднанню методів машинного навчання та ШІ, системи можуть не лише розпізнавати об'єкти, але й приймати рішення на основі аналізу поведінки об'єктів у часі.

ШІ використовується для:

* Аналізу поведінки об'єктів: виявлення аномальних дій або підозрілої поведінки.
* Прогнозування подій: на основі аналізу відеоданих системи можуть передбачити певні події, наприклад, прогнозувати можливість зіткнення транспортних засобів.
* Виявлення аномалій: системи ШІ можуть автоматично виявляти аномальні об'єкти або дії на основі аналізу великої кількості відео.

Машинне навчання та штучний інтелект стали фундаментальними технологіями у розпізнаванні об'єктів у відеопотоці. Завдяки сучасним алгоритмам глибокого навчання, таким як CNN, RNN, YOLO та Faster R-CNN, системи можуть автоматично обробляти великі обсяги даних, забезпечуючи високу точність та ефективність розпізнавання. Ці інструменти відіграють вирішальну роль у багатьох галузях, таких як безпека, транспорт, медицина та інші сфери, де потрібен автоматизований аналіз відеоданих.

## 1.4 Актуальність аналізу відеопотоків

Системи обробки та аналізу відеозображень знаходять широке застосування у багатьох галузях, завдяки швидкому розвитку алгоритмів машинного навчання, штучного інтелекту та обчислювальних технологій. Ці системи дозволяють автоматизувати процеси аналізу відеоданих, забезпечуючи високу точність, швидкість і надійність. У різних сферах, від безпеки до медицини, такі системи відіграють важливу роль у підвищенні ефективності, зниженні людських помилок та поліпшенні якості послуг.

**Галузі застосування систем обробки та аналізу відео**

* Відеоспостереження та безпека

Системи відеоспостереження є однією з найважливіших областей використання технологій аналізу відео. У цій сфері автоматизовані системи здатні:

* + Виявляти підозрілу поведінку, аномальні ситуації або незаконні дії.
  + Розпізнавати обличчя в реальному часі для ідентифікації осіб, що дозволяє підвищити безпеку об'єктів та громадських місць.
  + Відстежувати рух транспортних засобів або людей, наприклад, на вулицях, парковках або в торгових центрах.

Сучасні системи безпеки використовують такі алгоритми, як YOLO або Faster R-CNN, для швидкого розпізнавання об'єктів, а також системи розпізнавання осіб, які застосовуються на прикордонних переходах та у великих громадських просторах.

* Медицина

Системи аналізу відео знаходять широке застосування в медицині, де вони використовуються для:

* + Аналізу відеозаписів ендоскопічних процедур або МРТ зображень для виявлення патологій, таких як пухлини або пошкодження органів.
  + Підтримки хірургічних операцій, де відеоаналіз допомагає визначати точне положення інструментів або забезпечує візуальний контроль за процесом операції.
  + Діагностики шкірних захворювань або очних патологій через відеоаналітику з високою роздільною здатністю.

Наприклад, штучні нейронні мережі, такі як Convolutional Neural Networks (CNN), дозволяють автоматизовано виявляти ракові клітини на зображеннях, що значно підвищує точність діагностики та зменшує час обробки.

* Автономні транспортні засоби

Однією з найпередовіших сфер застосування систем обробки відео є безпілотні автомобілі. Для таких транспортних засобів життєво важливо "бачити" та аналізувати оточення у режимі реального часу. Завдяки системам комп'ютерного зору, автомобілі можуть:

* + Розпізнавати дорожні знаки та сигнали світлофорів.
  + Виявляти пішоходів, інші транспортні засоби або перешкоди на дорозі.
  + Визначати дорожню розмітку для утримання на смузі руху.

Такі компанії, як Tesla та Waymo, використовують системи глибокого навчання для розпізнавання та відстеження об'єктів навколо транспортного засобу, що дозволяє забезпечити високий рівень безпеки.

* Спорт та розваги

В аналізі спортивних подій системи відеоаналізу дозволяють отримувати інформацію про дії гравців у реальному часі. У спорті такі технології використовуються для:

* + Аналізу ігор, відстеження руху м'яча, гравців та інших об'єктів на полі.
  + Надання інформації тренерам для покращення тактики або фізичної підготовки спортсменів.
  + Автоматизації відеоповторів та визначення ключових моментів гри, що робить спортивні трансляції більш інтерактивними.

Крім того, у сфері розваг технології відеоаналізу використовуються в кіноіндустрії для створення візуальних ефектів, автоматизованого монтажу відео та аналізу глядацької поведінки.

* Сільське господарство

Сучасні технології аналізу відеозображень відіграють важливу роль у сільському господарстві. Зокрема, дрони з відеокамерами, оснащені алгоритмами машинного навчання, можуть:

* + Визначати стан рослин, виявляти хвороби або шкідників.
  + Оцінювати якість ґрунтів, моніторити посіви та прогнозувати врожайність.
  + Автоматично управляти процесом зрошення та внесення добрив на основі відеоаналізу.

Наприклад, технології комп'ютерного зору використовуються для виявлення аномалій у рості рослин, що дозволяє аграріям швидко реагувати на зміни та підвищувати ефективність виробництва.

* Виробництво та контроль якості

На виробничих підприємствах системи аналізу відео використовуються для контролю якості продукції. Камери високої роздільної здатності, оснащені алгоритмами машинного навчання, можуть:

* + Виявляти дефекти продукції на конвеєрних лініях.
  + Моніторити процеси складання або пакування товарів.
  + Забезпечувати безпеку на виробництві, відстежуючи рух людей і машин.

Наприклад, на автомобільних заводах відеоаналітика дозволяє автоматично виявляти пошкоджені деталі або помилки під час складання автомобілів.

* Маркетинг та ритейл

У сфері ритейлу системи аналізу відеозображень використовуються для моніторингу поведінки клієнтів у магазинах. Завдяки таким системам маркетологи можуть:

* + Відстежувати переміщення покупців по торгових площах, аналізувати їхню поведінку та визначати, які товари привертають найбільшу увагу.
  + Аналізувати ефективність розташування товарів на полицях.
  + Автоматизувати облік клієнтів та оцінювати їхню лояльність за допомогою технологій розпізнавання обличчя.

Таким чином, маркетологи отримують важливі дані для покращення продажів та створення індивідуальних пропозицій для клієнтів.

* Розумні міста та інфраструктура

У контексті розвитку концепції "розумних міст" (Smart Cities) відеоаналіз допомагає автоматизувати безліч процесів, які раніше потребували участі людини. У містах такі системи можуть:

* + Моніторити трафік та оптимізувати транспортні потоки, запобігаючи заторам.
  + Виявляти порушення правил дорожнього руху та автоматично виписувати штрафи.
  + Відстежувати екологічні показники (забруднення повітря, шум, освітлення).
  + Підвищувати безпеку громадян шляхом інтеграції систем відеоспостереження з автоматизованими алгоритмами виявлення небезпечних ситуацій.
* Робототехніка

У сфері робототехніки відеоаналіз відіграє вирішальну роль у здатності роботів взаємодіяти з навколишнім середовищем. Роботи з комп'ютерним зором можуть:

* + Виконувати складні маніпуляції з предметами.
  + Орієнтуватися у просторі та взаємодіяти з людьми.
  + Використовувати відеоаналіз для визначення перешкод та уникнення зіткнень.

**Майбутні перспективи**

З розвитком технологій штучного інтелекту та комп'ютерного зору можливості систем обробки та аналізу відео продовжуватимуть зростати. Очікується, що у найближчі роки такі системи стануть ще точнішими, швидшими та доступнішими, відкриваючи нові можливості для застосування у різних сферах. Автоматизація процесів аналізу відеоданих дозволить значно покращити якість життя, безпеку та ефективність бізнес-процесів у багатьох галузях.

## 1.5 Сучасні технології

Сучасні технології та програмні засоби для аналізу зображень і відеопотоків грають ключову роль у розробці систем комп'ютерного зору та машинного навчання. Існує низка популярних бібліотек і фреймворків, які активно використовуються в дослідженнях і прикладних рішеннях для обробки та аналізу зображень і відео. Серед найвідоміших і найчастіше використовуваних варто виділити OpenCV, TensorFlow і PyTorch. Кожен із цих інструментів має свої переваги, недоліки та особливості застосування, що робить їх придатними для різних завдань. У цьому розділі розглянемо ці інструменти, їхні особливості, а також порівняємо їх за різними критеріями, такими як швидкість роботи, простота використання та можливості.

OpenCV (Open Source Computer Vision Library)

Загальний огляд

OpenCV — це відкрита бібліотека, яка була розроблена для обробки зображень і відео в реальному часі. Вона містить величезну кількість функцій для обробки зображень, зокрема алгоритми для виявлення об’єктів, відстеження руху, сегментації, вирівнювання, калібрування камери тощо. Спочатку OpenCV була написана на C++, але сьогодні вона також підтримує інтерфейси для мов Python, Java і MATLAB, що робить її доступною для широкого кола розробників.

Основні можливості

1. Обробка зображень: OpenCV включає численні алгоритми для фільтрації, масштабування, сегментації зображень, корекції кольорів, а також перетворення зображень із одного колірного простору в інший.
2. Обробка відео: Бібліотека забезпечує можливість читання та обробки відеопотоків у реальному часі. Вона дозволяє здійснювати відстеження об'єктів, детектування руху та побудову систем відеоспостереження.
3. Машинне навчання: В OpenCV є модуль для машинного навчання, який підтримує класифікатори на основі дерев рішень, SVM, нейронних мереж та інших моделей, що полегшує роботу з розпізнаванням об'єктів.
4. Інтеграція з іншими інструментами: OpenCV легко інтегрується з іншими бібліотеками та фреймворками, такими як TensorFlow або PyTorch, що робить її універсальним інструментом для обробки зображень і подальшої аналітики.

Переваги

1. Висока продуктивність: Завдяки реалізації на C++ OpenCV забезпечує високу швидкість обробки даних, що є критичним для обробки відео в реальному часі.
2. Широкий набір алгоритмів: Бібліотека включає великий набір готових функцій для обробки зображень і відео, що робить її універсальним інструментом для різних завдань.
3. Відкрите джерело: Вільний доступ до коду дозволяє розробникам змінювати та адаптувати бібліотеку під свої потреби.

Недоліки

1. Обмежені можливості машинного навчання: Хоча OpenCV включає базові алгоритми машинного навчання, вона не надає таких потужних інструментів для побудови нейронних мереж, як TensorFlow або PyTorch.
2. Менш гнучка для складних моделей: OpenCV добре підходить для базових завдань комп'ютерного зору, але для глибокого навчання може вимагати інтеграції з іншими бібліотеками.

TensorFlow

Загальний огляд

TensorFlow — це фреймворк для машинного навчання, розроблений компанією Google. Він є одним із найпопулярніших інструментів для побудови та тренування нейронних мереж. TensorFlow відомий своєю здатністю працювати з великими наборами даних і забезпечувати високу продуктивність при тренуванні глибоких моделей. Однією з основних переваг TensorFlow є його універсальність: він може використовуватися як для створення простих моделей машинного навчання, так і для складних систем із мільйонами параметрів.

Основні можливості

1. Підтримка нейронних мереж: TensorFlow підтримує багатошарові нейронні мережі, згорткові нейронні мережі (CNN), рекурентні нейронні мережі (RNN), а також інші типи моделей глибокого навчання.
2. Тренування на GPU: TensorFlow оптимізований для роботи з графічними процесорами (GPU), що дозволяє суттєво прискорити процес тренування моделей.
3. TensorBoard: Інструмент для візуалізації процесу тренування нейронної мережі. Він дозволяє слідкувати за параметрами моделі в реальному часі, що полегшує налагодження і аналіз результатів.
4. TensorFlow Lite та TensorFlow.js: Полегшені версії TensorFlow для мобільних пристроїв і веб-додатків відповідно, що робить можливим запуск моделей навіть на обмежених платформах.

Переваги

1. Підтримка великих нейронних мереж: TensorFlow спеціалізується на роботі з великими моделями, що дозволяє використовувати його для складних завдань, таких як розпізнавання об'єктів у відеопотоці.
2. Широка спільнота та документація: Завдяки великій кількості користувачів TensorFlow, існує багато ресурсів, навчальних матеріалів і прикладів використання.
3. Гнучкість: TensorFlow можна використовувати для побудови як простих, так і складних моделей, що робить його універсальним інструментом для різних завдань.

Недоліки

1. Складність у використанні: Через складну архітектуру TensorFlow може бути важким для початківців у порівнянні з більш простими бібліотеками.
2. Вимогливість до ресурсів: Для тренування великих моделей на TensorFlow потрібні потужні апаратні ресурси, що може бути проблемою для деяких користувачів.

PyTorch

Загальний огляд

PyTorch — це ще один популярний фреймворк для глибокого навчання, розроблений компанією Facebook. PyTorch став дуже популярним завдяки простоті використання та інтеграції з Python. Основною перевагою PyTorch є його динамічна обчислювальна графіка, що дозволяє легко налагоджувати та модифікувати моделі під час тренування. Це робить PyTorch особливо привабливим для дослідників і розробників, які працюють над експериментальними проектами в галузі машинного навчання.

Основні можливості

1. Динамічна графіка обчислень: PyTorch дозволяє змінювати модель під час її тренування, що забезпечує більшу гнучкість при експериментах з різними архітектурами нейронних мереж.
2. Простота використання: Інтерфейс PyTorch дуже схожий на Python, що робить його зручним для розробників із різним рівнем досвіду.
3. Підтримка GPU: Як і TensorFlow, PyTorch підтримує обчислення на GPU, що дозволяє швидко тренувати великі моделі.

Переваги

1. Гнучкість та зручність у використанні: PyTorch надає зручний і гнучкий інтерфейс, що робить його придатним як для досліджень, так і для комерційних застосувань.
2. Інтеграція з Python: PyTorch тісно інтегрований з мовою Python, що дозволяє легко використовувати його разом із іншими бібліотеками, такими як NumPy чи SciPy.

Недоліки

1. Менша продуктивність при великих обчисленнях: У деяких випадках TensorFlow може бути швидшим за PyTorch при тренуванні великих моделей через кращу оптимізацію.
2. Відносно нова спільнота: Хоча PyTorch стає все більш популярним, його спільнота ще не така велика, як у TensorFlow, що може обмежити кількість доступних ресурсів і прикладів.

Порівняння OpenCV, TensorFlow і PyTorch

Швидкість роботи

* OpenCV забезпечує високу швидкість при обробці зображень і відео, оскільки багато її функцій оптимізовані для роботи в реальному часі.
* TensorFlow і PyTorch мають схожу продуктивність при тренуванні великих моделей, але TensorFlow зазвичай виявляється швидшим при роботі з великими нейронними мережами завдяки кращій оптимізації для GPU.

Простота використання

* PyTorch вважається більш простим у використанні завдяки динамічній обчислювальній графіці та зручному інтерфейсу.
* TensorFlow може бути складнішим для початківців через свою складну архітектуру.
* OpenCV є простим для задач обробки зображень, але може бути складнішим, якщо його використовувати разом із глибоким навчанням.

Можливості

* TensorFlow і PyTorch забезпечують потужні можливості для побудови та тренування нейронних мереж, включаючи роботу з великими наборами даних.
* OpenCV більше орієнтована на обробку зображень і відео в реальному часі, але її можна використовувати разом із іншими фреймворками для машинного навчання.

# 2. ВИСНОВКИ

Тема "Методи та комп'ютерні засоби аналізу зображень у відеопотоці" є надзвичайно актуальною в умовах швидкого розвитку технологій комп'ютерного зору і штучного інтелекту. Сучасні методи обробки відеозображень, такі як геометричні перетворення, колірна корекція та сегментація, є основою для подальшого аналізу. Відстеження об'єктів у відеопотоці, завдяки алгоритмам, таким як оптичний потік, кореляційні фільтри і глибокі нейронні мережі, дозволяє точно визначати положення об'єктів і їх траєкторії у часі.

Машинне навчання, зокрема згорткові (CNN) та рекурентні нейронні мережі (RNN), а також моделі, як YOLO і Faster R-CNN, забезпечують потужні інструменти для розпізнавання та класифікації об'єктів у відеопотоці, підвищуючи точність і швидкість обробки.

Застосування таких технологій охоплює безліч галузей, включаючи відеоспостереження, медицину, автономні транспортні засоби, сільське господарство, виробництво, ритейл та інші сфери. Це дозволяє автоматизувати процеси, підвищувати ефективність, забезпечувати безпеку і поліпшувати якість життя.

З огляду на ці досягнення, можна очікувати подальший розвиток і вдосконалення технологій аналізу відеопотоків, що забезпечить нові можливості для їх застосування та значний внесок у різні аспекти суспільного життя.

# СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

* + 1. Garg, & Verma, G. (2017). Operating Systems: An Introduction. Mercury Learning & Information.
    2. Panek. (2020). Windows operating system fundamentals (1st edition). Sybex. – 419 p.
    3. David A. Solomon (2017). “Windows Internals”.
    4. Габрусєв В.Ю. (2007). Основи операційних систем: ядро, процес, потік. – 96с.
    5. Бондаренко М.Ф. Операційні системи: навч. посібник / М.Ф. Бондаренко, О.Г. Качко. – Харків: Компанія СМІТ, 2008. – 432 с.
    6. Johnson M. Hart. Windows System Programming, 4th edition / Hart Johnson. – Addison-Wesley, 2010. – 656 p.
    7. Thomas W. Doeppner. Operating Systems In Depth: Design and Programming / W. Thomas. – John Wiley & Sons, 2010. – 462 p
    8. Goodfellow I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville. – MIT Press, 2016. – 800 p.
    9. Bishop C.M. Pattern Recognition and Machine Learning / C.M. Bishop. – Springer, 2006. – 738 p.
    10. Szegedy C. Going deeper with convolutions / C. Szegedy, W. Liu, Y. Jia et al. – IEEE Conference on Computer Vision and Pattern Recognition, 2015. – 12 p.
    11. Redmon J. You Only Look Once: Unified, Real-Time Object Detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi. – IEEE Conference on Computer Vision and Pattern Recognition, 2016. – 9 p.
    12. LeCun Y. Gradient-based learning applied to document recognition / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner. – Proceedings of the IEEE, 1998. – 2278 p.
    13. Krizhevsky A. ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G. Hinton. – Advances in Neural Information Processing Systems, 2012. – 7 p.
    14. He K. Deep Residual Learning for Image Recognition / K. He, X. Zhang, S. Ren, J. Sun. – IEEE Conference on Computer Vision and Pattern Recognition, 2016. – 9 p.