

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ

Кафедра програмної інженерії

**КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА**

на тему: «**ВИКОРИСТАННЯ МЕТОДІВ АНОТАЦІЇ  
ФОТО ТА ВІДЕО МАТЕРІАЛІВ ДЛЯ МАШИННОГО  
НАВЧАННЯ МОДЕЛЕЙ КОМП'ЮТЕРНОГО ЗОРУ**»

Виконав: студент 4 курсу, групи 6.1210-2пі  
спеціальності 121 інженерія програмного забезпечення  
(шифр і назва спеціальності)

освітньої програми програмна інженерія  
(назва освітньої програми)

К.А. Кошель

(ініціали та прізвище)

Керівник доцент кафедри програмної інженерії,  
к.ф.-м.н., Мильцев О.М.  
(посада, вчене звання, науковий ступінь, прізвище та ініціали)

Рецензент завідувач кафедри фундаментальної та прикладної  
математики, професор, д.т.н. Гребенюк С.М.  
(посада, вчене звання, науковий ступінь, прізвище та ініціали)

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

Факультет математичний

Кафедра програмної інженерії

Рівень вищої освіти бакалавр

Спеціальність 121 інженерія програмного забезпечення

(шифр і назва)

Освітня програма програмна інженерія

**ЗАТВЕРДЖУЮ**

Завідувач кафедри програмної  
інженерії, к.ф.-м.н., доцент

\_\_\_\_\_ Лісняк А.О.

(підпис)

“ \_\_\_\_\_ ” \_\_\_\_\_ 2023 р.

**З А В Д А Н Н Я**  
**НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТОВІ**

Кошелю Кірілу Андрійовичу

(прізвище, ім'я та по-батькові)

1. Тема роботи Використання методів анотації фото та відео матеріалів для  
машинного навчання моделей комп'ютерного зору

керівник роботи Мильцев Олександр Михайлович, к.ф.-м.н.

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

затверджені наказом ЗНУ від « 21 » грудня 2023 року № 2180-с

2. Строк подання студентом роботи 03.06.2024 р.

3. Вихідні дані до роботи 1. Постановка задачі.  
2. Перелік літератури.

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

1. Постановка задачі.

2. Основні теоретичні відомості.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень) \_\_\_\_\_  
презентація за темою доповіді

## 6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 25.12.2023 р.**КАЛЕНДАРНИЙ ПЛАН**

№	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1.	Розробка плану роботи.	15.01.2024	
2.	Збір вихідних даних.	05.02.2024	
3.	Обробка методичних та теоретичних джерел.	11.03.2024	
4.	Розробка першого та другого розділу.	22.04.2024	
5.	Розробка третього розділу.	20.05.2024	
6.	Оформлення та нормоконтроль кваліфікаційної роботи бакалавра.	27.05.2024	
7.	Захист кваліфікаційної роботи.	19.06.2024	

Студент \_\_\_\_\_  
(підпис)К.А. Кошель \_\_\_\_\_  
(ініціали та прізвище)Керівник роботи \_\_\_\_\_  
(підпис)О.М. Мильцев \_\_\_\_\_  
(ініціали та прізвище)**Нормоконтроль пройдено**Нормоконтролер \_\_\_\_\_  
(підпис)А.В. Столярова \_\_\_\_\_  
(ініціали та прізвище)

## РЕФЕРАТ

Кваліфікаційна робота бакалавра «Використання методів анотації фото та відео матеріалів для машинного навчання моделей комп'ютерного зору»: 44 с., 5 рис., 25 джерел.

АНОТАЦІЯ ЗОБРАЖЕНЬ, ГЕНЕРАЦІЯ ОПИСУ ЗОБРАЖЕНЬ, КОМП'ЮТЕРНИЙ ЗІР, ЛОКАЛІЗАЦІЯ ГРАФІЧНИХ ЕЛЕМЕНТІВ, ЛОКАЛІЗАЦІЯ ТЕКСТУ, МАШИННЕ НАВЧАННЯ, ОБРОБКА ФОТО ТА ВІДЕО МАТЕРІАЛІВ, РОЗПІЗНАВАННЯ ОБ'ЄКТІВ, РУЧНА АНОТАЦІЯ, СЕГМЕНТАЦІЯ ЗОБРАЖЕНЬ.

Об'єкт дослідження – модель комп'ютерного зору.

Мета роботи: дослідження існуючих методів навчання моделі комп'ютерного зору із використанням штучного інтелекту, аналізі методів і інструментів для обробки фото та відео матеріалів, вивченні типів анотацій та їх класифікації, а також описанні прикладів застосування таких систем.

Метод дослідження – аналіз літератури та існуючих досліджень для вивчення сучасних методів навчання моделей комп'ютерного зору із використанням штучного інтелекту, а також огляд методів і інструментів для обробки фото та відео матеріалів. Створити класифікацію різних типів анотацій, що використовуються для фото та відео матеріалів, з описом переваг і недоліків кожного типу анотацій.

## SUMMARY

Bachelor's qualifying paper "Utilization of Annotation Methods for Photo and Video Materials in Machine Learning of Computer Vision Models": 44 pages, 5 figures, 25 references.

IMAGE ANNOTATION, IMAGE DESCRIPTION GENERATION, COMPUTER VISION, GRAPHIC ELEMENTS LOCALIZATION, TEXT LOCALIZATION, MACHINE LEARNING, PHOTO AND VIDEO MATERIALS PROCESSING, OBJECT RECOGNITION, MANUAL ANNOTATION, IMAGE SEGMENTATION.

The object of research is a computer vision model.

The purpose of the work: the study of existing methods of training the computer vision model using artificial intelligence, the analysis of methods and tools for processing photo and video materials, the study of annotation types and their classification, as well as the description of examples of the use of such systems.

The research method is an analysis of literature and existing research to study modern methods of training computer vision models using artificial intelligence, as well as an overview of methods and tools for processing photo and video materials. Create a classification of different types of annotations used for photo and video materials, describing the advantages and disadvantages of each type of annotation.

## ЗМІСТ

Завдання на кваліфікаційну роботу.....	2
Реферат .....	4
Summary .....	5
Вступ.....	8
1 Методи анотації.....	10
1.1 Ручна анотація.....	10
1.1.1 Опис.....	10
1.1.2 Переваги.....	10
1.1.3 Недоліки.....	10
1.1.4 Приклади використання .....	11
1.1.5 Інструменти .....	11
1.2 Напів автоматична анотація.....	12
1.2.1 Опис.....	12
1.2.2 Переваги.....	13
1.2.3 Недоліки.....	13
1.2.4 Приклади використання .....	13
1.2.5 Інструменти .....	14
1.3 Автоматична анотація .....	14
1.3.1 Опис.....	14
1.3.2 Переваги.....	15
1.3.3 Недоліки.....	15
1.3.4 Приклади використання .....	16
1.3.5 Інструменти .....	16
2 Типи анотацій .....	17
2.1 Класифікація.....	17
2.2 Локалізація.....	18
2.3 Сегментація.....	20

2.4 Відстеження .....	21
2.5 Опис .....	23
3 Застосування .....	25
3.1 Розпізнавання облич .....	25
3.2 Розпізнавання об'єктів.....	27
3.3 Сегментація зображень.....	29
3.4 Відстеження об'єктів .....	32
3.5 Генерація опису зображень.....	36
Висновки .....	40
Перелік посилань.....	42

## ВСТУП

**Машинне навчання** – це підгалузь штучного інтелекту, яка дає змогу комп'ютерам навчатися без явного програмування. Це досягається за допомогою алгоритмів, які можуть аналізувати дані, виявляти закономірності та робити прогнози. Машинне навчання використовується в широкому спектрі застосувань, включаючи розпізнавання облич, мовний переклад, автопілот та медичну діагностику.

**Комп'ютерний зір** – це галузь штучного інтелекту, яка дає змогу комп'ютерам інтерпретувати та розуміти візуальну інформацію. Це досягається за допомогою алгоритмів, які можуть обробляти зображення та відео, виявляти об'єкти, класифікувати їх та відстежувати їх рух. Комп'ютерний зір використовується в широкому спектрі застосувань, включаючи розпізнавання облич, медичну візуалізацію, робототехніку та автономне водіння.

Алгоритми машинного навчання використовуються для навчання моделей комп'ютерного зору на великих наборах даних зображень та відео. Ці моделі можуть вивчати закономірності даних і використовувати їх для виконання завдань, таких як розпізнавання об'єктів, класифікація зображень та відстеження руху.

**Анотація фото та відео матеріалів** – це процес маркування даних для навчання моделей машинного навчання. Цей процес може бути виконаний вручну, напівавтоматично або автоматично. Ручна анотація є найточнішою, але вона також є найдорожчою та найчасоємнішою. Напівавтоматична анотація використовує інструменти для полегшення ручної анотації, тоді як автоматична анотація використовує алгоритми машинного навчання для автоматичного маркування даних.

Використання анотації фото та відео матеріалів є важливим для машинного навчання моделей комп'ютерного зору, оскільки воно дає змогу



моделям навчатися на високоякісних даних. Це, у свою чергу, допомагає моделям покращити свою точність та продуктивність.

Машинне навчання та комп'ютерний зір мають широкий спектр застосувань. Деякі з найпоширеніших застосувань включають:

- розпізнавання облич: розпізнавання облич використовується для ідентифікації та верифікації людей – це використовується в таких застосуваннях, як авторизація доступу, моніторинг безпеки та соціальні мережі;
- розпізнавання об'єктів: розпізнавання об'єктів використовується для виявлення та класифікації об'єктів у зображеннях та відео – це використовується в таких застосуваннях, як пошук зображень, візуальний контроль та каталогізація продуктів;
- сегментація зображень: сегментація зображень використовується для виділення окремих об'єктів або областей на зображеннях – це використовується в таких застосуваннях, як медичний візуалізація, редагування фотографій та автономне водіння;
- відстеження об'єктів: відстеження об'єктів використовується для відстеження руху об'єктів протягом зображення або відео – це використовується в таких застосуваннях, як моніторинг трафіку, аналіз поведінки та спортивні трансляції;
- генерація опису зображень: генерація опису зображень використовується для створення текстових описів зображень для людей з порушеннями зору – це використовується в таких застосуваннях, як доступність до інформації та покращення пошуку зображень.

# 1 МЕТОДИ АНОТАЦІЇ

## 1.1 Ручна анотація

### 1.1.1 Опис

Ручна анотація – це процес маркування даних для навчання моделей машинного навчання, який здійснюється людьми. Це найточніший метод анотації, але він також є найдорожчим та найчасоємнішим.

При ручній анотації експерти або фахівці переглядають зображення та відео та надають їм мітки. Ці мітки можуть бути у формі класів, описів, координат або інших форматів, залежно від завдання, яке виконується.

### 1.1.2 Переваги

**Висока точність:** ручна анотація дає найвищу точність, оскільки експерти можуть використовувати свої знання та судження для правильного маркування даних.

**Гнучкість:** ручна анотація може використовуватися для будь-якого типу даних, включаючи зображення, відео, текст та аудіо.

**Контроль:** люди мають повний контроль над процесом анотації, що дає змогу їм гарантувати, що дані анотовані правильно та послідовно.

### 1.1.3 Недоліки

**Висока вартість:** ручна анотація може бути дуже дорогою, оскільки вона потребує багато часу та ресурсів.

**Тривалість:** ручна анотація може бути дуже тривалою, особливо для великих наборів даних.

**Необ'єктивність:** ручна анотація може бути схильною до людської помилки та упереджень.

#### **1.1.4 Приклади використання**

**Розпізнавання облич:** ручна анотація використовується для створення наборів даних зображень обличчя, які позначені іменами людей на них. Ці набори даних використовуються для навчання моделей розпізнавання облич [1].

**Розпізнавання об'єктів:** ручна анотація використовується для створення наборів даних зображень, які позначені об'єктами, що на них містяться. Ці набори даних використовуються для навчання моделей розпізнавання об'єктів [2].

**Сегментація зображень:** ручна анотація використовується для створення наборів даних зображень, які сегментовані на різні об'єкти або області. Ці набори даних використовуються для навчання моделей сегментації зображень [3].

**Відстеження об'єктів:** ручна анотація використовується для створення наборів даних відео, які позначені об'єктами, що відстежуються протягом відео. Ці набори даних використовуються для навчання моделей відстеження об'єктів [2].

#### **1.1.5 Інструменти**

Існує багато інструментів, які можна використовувати для ручної анотації даних. Деякі з найпопулярніших інструментів включають:

- LabelImg: LabelImg – це безкоштовний інструмент з відкритим кодом для анотації зображень та відео (<https://labelbox.com/>);
- VGG Image Annotator: VGG Image Annotator – це безкоштовний онлайн-інструмент для анотації зображень (<https://www.image-net.org/>);
- COCO Annotator: COCO Annotator – це безкоштовний онлайн-інструмент для анотації зображень та відео (<https://cocodataset.org/>);
- MATLAB: MATLAB – це комерційне програмне забезпечення, яке можна використовувати для анотації даних (<https://www.mathworks.com/>).

## 1.2 Напів автоматична анотація

Напівавтоматична анотація поєднує в собі ручні та автоматичні методи для створення анотацій до тексту, зображень або інших даних. Цей процес використовує комп'ютерні алгоритми для полегшення роботи людини-анотатора, роблячи його більш ефективним.

### 1.2.1 Опис

Напівавтоматичні системи анотації включають наступні кроки:

- підготовка даних: дані готуються до анотації, очищаються від помилок та форматуються;
- автоматична анотація: комп'ютерна програма генерує початкові анотації для даних;
- перевірка людиною: людина-анотатор перевіряє та коригує автоматично згенеровані анотації;
- навчання: система навчається на основі виправлень людини-анотатора, щоб покращити свою точність у майбутньому.

### **1.2.2 Переваги**

**Збільшена швидкість:** напівавтоматична анотація може значно прискорити процес анотації порівняно з ручною анотацією.

**Покращена точність:** автоматичні алгоритми можуть допомогти виявити закономірності та зв'язки, які людина може пропустити, що може призвести до більш точних анотацій.

**Зменшення витрат:** напівавтоматична анотація може бути більш економічною, ніж ручна анотація, оскільки вона потребує менше людської праці.

**Масштабованість:** напівавтоматичні системи можна легко масштабувати для обробки великих обсягів даних.

### **1.2.3 Недоліки**

**Вартість:** розробка та впровадження напівавтоматичної системи анотації може бути дорогою.

**Потреба в навчальних даних:** для навчання напівавтоматичної системи потрібна значна кількість високоякісних навчальних даних.

**Потенційна упередженість:** автоматичні алгоритми можуть бути упередженими, що може призвести до упереджених анотацій.

**Потреба в людському контролі:** напівавтоматичні системи все ще потребують людського контролю та втручання, що може бути дорогим та трудомістким.

### **1.2.4 Приклади використання**

Напівавтоматична анотація використовується в різних галузях, включаючи:

- обробка природної мови (ОНМ): анотація тексту для таких завдань, як визначення частин мови, іменування сутностей та виявлення настроїв;
- комп'ютерний зір: анотація зображень для таких завдань, як розпізнавання об'єктів, сегментація зображень та ключових точок;
- біоінформатика: анотація біологічних даних, таких як послідовності ДНК та білків.

### 1.2.5 Інструменти

Існує багато інструментів для напівавтоматичної анотації, включаючи:

- **Labelbox:** <https://labelbox.com/>;
- **TagCrowd:** <https://tagcrowd.com/>;
- **BRAT:** <https://brat.nlplab.org/>;
- **LightTag:** <https://www.lighttag.io/help>;
- **DodoLab:** <https://www.dodolab.ca/>.

## 1.3 Автоматична анотація

Автоматична анотація використовує комп'ютерні алгоритми для автоматичного створення анотацій до тексту, зображень або інших даних без втручання людини. Цей процес може значно прискорити та масштабувати процес анотації, роблячи його більш економним та ефективним.

### 1.3.1 Опис

Системи автоматичної анотації включають наступні кроки:

- підготовка даних: дані готуються до анотації, очищаються від

помилки та форматується;

- навчання моделі: навчають модель машинного навчання на наборі даних з анотаціями, зробленими вручну;
- автоматична анотація: модель машинного навчання використовується для автоматичного створення анотацій для нових даних;
- оцінка: якість автоматичних анотацій оцінюється за допомогою порівняння з анотаціями, зробленими вручну.

### **1.3.2 Переваги**

**Швидкість:** автоматична анотація може значно прискорити процес анотації порівняно з ручною анотацією.

**Масштабованість:** автоматичні системи можна легко масштабувати для обробки великих обсягів даних.

**Економічність:** автоматична анотація може бути більш економічною, ніж ручна анотація, оскільки вона потребує менше людської праці.

**Зменшення упередженості:** автоматичні системи можуть бути менш упередженими, ніж люди-анотатори, які можуть бути схильні до суб'єктивності.

### **1.3.3 Недоліки**

**Точність:** автоматичні анотації не завжди такі ж точні, як анотації, зроблені вручну, особливо для складних або нечітких даних.

**Потреба в навчальних даних:** для навчання моделей машинного навчання потрібна значна кількість високоякісних навчальних даних з анотаціями, зробленими вручну.

Потенційна упередженість: навчальні дані можуть бути упередженими, що може призвести до упереджених автоматичних анотацій.

Обмежена гнучкість: автоматичні системи можуть бути не такими гнучкими, як люди-анотатори, і можуть не впоратися з нестандартними або складними випадками.

### 1.3.4 Приклади використання

Автоматична анотація використовується в різних галузях, включаючи:

- обробка природної мови (ОНМ): анотація тексту для таких завдань, як визначення частин мови, іменування сутностей та виявлення настроїв;
- комп'ютерний зір: анотація зображень для таких завдань, як розпізнавання об'єктів, сегментація зображень та визначення ключових точок;
- біоінформатика: анотація біологічних даних, таких як послідовності ДНК та білків.

### 1.3.5 Інструменти

Існує багато інструментів для автоматичної анотації, включаючи:

- **spaCy**: <https://spacy.io/>;
- **OpenNLP**: <https://opennlp.apache.org/>;
- **Labelbox**: <https://labelbox.com/>.



## 2 ТИПИ АНОТАЦІЙ

### 2.1 Класифікація

Класифікація є фундаментальним завданням в сфері обробки природної мови (ОНМ), яке полягає у присвоєнні категорій або міток текстам. Ці категорії можуть описувати тему тексту, його емоційну тональність, об'єктивність, або інші характеристики.

Існує безліч методів класифікації тексту, які можна поділити на дві основні категорії:

- методи машинного навчання: ці методи використовують алгоритми машинного навчання, які навчаються на наборі даних з анотованими текстами, щоб автоматично класифікувати нові тексти;
- методи на основі правил: ці методи використовують набір правил, які вручну розроблені експертами, для класифікації текстів.

**Методи машинного навчання.** Найпоширенішими методами машинного навчання для класифікації тексту є:

- наївний метод Байєса: цей метод ґрунтується на теоремі Байєса і припускає, що всі ознаки тексту є незалежними;
- метод k-найближчих сусідів (k-NN): цей метод класифікує новий текст на основі його схожості з k найближчими текстами з набору даних навчання;
- дерева рішень: цей метод використовує дерево рішень для представлення правил класифікації;
- підтримка векторів (SVM): цей метод використовує гіперплощину для розділення текстів на різні класи;
- нейронні мережі: цей метод використовує штучні нейронні мережі для навчання складних моделей класифікації.

**Методи на основі правил.** Методи на основі правил зазвичай використовуються для класифікації текстів, коли доступні знання про предметну область. Ці методи можуть бути більш точними, ніж методи машинного навчання, але вони потребують більше часу і ресурсів для розробки.

**Вибір методу класифікації.** Вибір методу класифікації тексту залежить від декількох факторів, таких як:

- розмір та складність набору даних: великі та складні набори даних зазвичай потребують використання методів машинного навчання;
- якість анотацій: якщо анотації неточні або неповні, то методи машинного навчання можуть давати неточні результати;
- доступні ресурси: розробка методів на основі правил потребує більше часу і ресурсів, ніж використання методів машинного навчання [4].

## 2.2 Локалізація

Локалізація є процесом адаптації продукту або послуги до певної мови та культури. Це може включати переклад тексту, адаптацію графічних елементів та зміну форматування даних. Локалізація важлива для того, щоб зробити продукт або послугу доступними та зручними для користувачів у всьому світі.

**Методи локалізації.** Існує два основних методи локалізації:

- машинний переклад: цей метод використовує програмне забезпечення для автоматичного перекладу тексту з однієї мови на іншу;
- людський переклад: цей метод використовує професійних перекладачів для перекладу тексту.

**Вибір методу локалізації.** Вибір методу локалізації залежить від декількох факторів, таких як:

- обсяг тексту: для великих обсягів тексту машиний переклад може бути більш економічним варіантом;
- якість перекладу: людський переклад зазвичай забезпечує більш високу якість перекладу, ніж машиний;
- терміни: машинний переклад зазвичай виконується швидше, ніж людський;
- бюджет: людський переклад зазвичай дорожчий, ніж машиний [5].

**Локалізація тексту.** Локалізація тексту включає переклад тексту з однієї мови на іншу. Це може бути складним завданням, оскільки необхідно враховувати не лише значення тексту, але й його стиль, тон та культурні нюанси.

**Локалізація графічних елементів.** Локалізація графічних елементів включає адаптацію зображень, ілюстрацій та інших візуальних елементів до певної культури. Це може включати зміну кольорів, символів та зображень.

**Локалізація форматування даних.** Локалізація форматування даних включає зміну форматування дат, чисел, валют та інших даних, щоб вони відповідали очікуванням користувачів у певній культурі [6].

**Інструменти локалізації.** Існує багато інструментів, які можуть допомогти з локалізацією продукту або послуги. Ці інструменти можуть включати програмне забезпечення для перекладу, інструменти управління термінологією та системи управління контентом.

### **Рекомендації.**

- планування: перед початком локалізації потрібно розробити план, який буде включати визначення ваших цілей, цільової аудиторії та бюджету;
- вибір правильних методів: вибір методів локалізації, які найкраще відповідають потребам;
- використання якісних інструментів: використання якісних

- інструментів локалізації, щоб полегшити роботу;
- перевірка та тестування: переконатися, що продукт або послуга локалізована правильно, перевіривши та протестувавши його з користувачами у цільовій аудиторії [7].

### 2.3 Сегментація

Сегментація ринку – це процес поділу ринку на групи споживачів з схожими потребами, характеристиками та поведінкою. Це дозволяє компаніям краще зрозуміти своїх клієнтів, розробити більш ефективні маркетингові стратегії та створити продукти та послуги, які краще відповідають потребам своїх цільових аудиторій [8].

**Методи сегментації.** Існує багато методів сегментації ринку, які можна використовувати. Деякі з найпоширеніших методів включають:

- демографічна сегментація: цей метод сегментує ринок на основі таких характеристик, як вік, стать, дохід, освіта, місце проживання та професія;
- психографічна сегментація: цей метод сегментує ринок на основі таких характеристик, як особистість, цінності, стиль життя та інтереси;
- поведінкова сегментація: цей метод сегментує ринок на основі таких характеристик, як знання про продукт, ступінь використання продукту та лояльність до продукту [9].

**Вибір методу сегментації.** Вибір методу сегментації залежить від декількох факторів, таких як:

- цілі сегментації: що потрібно досягти за допомогою сегментації;
- продукт або послуга: який продукт або послуга продається;
- ринок: який ринок буде обслуговуватися;
- доступні дані: які дані про цільових клієнтів.

**Переваги сегментації.** Сегментація ринку може мати багато переваг для компаній, таких як:

- покращена ефективність маркетингу: сегментація дозволяє компаніям краще target свої маркетингові зусилля на конкретні групи споживачів, що може призвести до кращої рентабельності інвестицій (ROI);
- розробка нових продуктів та послуг: сегментація може допомогти компаніям краще зрозуміти потреби своїх клієнтів, що може допомогти їм розробити нові продукти та послуги, які краще відповідають цим потребам;
- підвищення лояльності клієнтів: задоволення потреб конкретних груп споживачів може призвести до підвищення лояльності клієнтів;
- підвищення конкурентоспроможності: сегментація може допомогти компаніям краще зрозуміти своїх конкурентів та їхні цільові аудиторії, що може допомогти їм розробити більш конкурентні стратегії [10].

## 2.4 Відстеження

Відстеження – це процес збору та аналізу даних про поведінку користувачів. Ці дані можуть використовуватися для різних цілей, таких як:

- покращення досвіду користувачів: відстеження може допомогти компаніям краще зрозуміти, як користувачі взаємодіють з їхніми вебсайтами та мобільними додатками – цю інформацію можна використовувати для покращення дизайну та функціональності цих продуктів;
- персоналізація: відстеження може використовуватися для персоналізації контенту та реклами для окремих користувачів – це може призвести до кращого досвіду користувачів та підвищення

- рентабельності інвестицій (ROI) для компаній;
- вимірювання ефективності маркетингових кампаній: відстеження може використовуватися для вимірювання ефективності маркетингових кампаній – ця інформація може допомогти компаніям покращити свої маркетингові зусилля та отримати кращу рентабельність інвестицій;
- запобігання шахрайству: відстеження може використовуватися для виявлення та запобігання шахрайству [11].

**Методи відстеження.** Існує багато методів відстеження, які можна використовувати. Розглянемо деякі з найпоширеніших методів.

**Файли cookie:** файли cookie – це невеликі текстові файли, які зберігаються на комп’ютері користувача, коли він відвідує вебсайт. Ці файли cookie можуть використовуватися для відстеження того, які сторінки відвідував користувач, скільки часу він проводив на кожній сторінці та які дії він виконував.

**Пікселі відстеження:** пікселі відстеження – це невеликі зображення, які вбудовуються на вебсайти. Коли користувач переглядає сторінку з пікселем відстеження, його IP-адреса та інша інформація про його пристрій надсилаються на сервер, який відстежує піксель.

**Аналітика вебсайтів:** аналітика вебсайтів – це програмне забезпечення, яке використовується для відстеження трафіку на вебсайті. Це програмне забезпечення може надавати інформацію про те, звідки приходять користувачі, які сторінки вони відвідують і скільки часу вони проводять на сайті.

**Відстеження мобільних додатків:** відстеження мобільних додатків – це програмне забезпечення, яке використовується для відстеження того, як користувачі взаємодіють з мобільними додатками. Це програмне забезпечення може надавати інформацію про те, які екрани переглядають користувачі, скільки часу вони проводять у кожному екрані та які дії вони виконують [12].

**Конфіденційність.** Відстеження може викликати проблеми з конфіденційністю, оскільки воно може використовуватися для збору особистої інформації про користувачів без їхньої згоди. Важливо знати про методи відстеження, які використовуються на вебсайтах та в мобільних додатках, які ви відвідуєте, і вживати заходів для захисту своєї конфіденційності [13].

## 2.5 Опис

Опис – це процес представлення інформації у чіткій, лаконічний та зрозумілий спосіб. Він може бути усним, письмовим або за допомогою інших видів медіа, таких як фотографії, відео або графіки.

**Мета опису.** Мета опису може варіюватися залежно від контексту. Деякі з найпоширеніших цілей опису включають:

- інформування: опис може використовуватися для надання інформації про щось, наприклад, про об'єкт, місце, подію або людину;
- переконання: опис може використовуватися для того, щоб переконати когось у чомусь, наприклад, у тому, щоб купити продукт або підтримати певну точку зору;
- розвага: опис може використовуватися для того, щоб розважити когось, наприклад, створити у нього певний настрій або емоцію;
- навчання: опис може використовуватися для того, щоб навчити когось чомусь, наприклад, як виконати завдання або як використовувати продукт.

**Елементи опису.** Ефективний опис зазвичай включає такі елементи:

- тема: тема опису – це те, про що описується;
- деталі: деталі опису – це конкретні факти та інформація, яка використовується для опису теми;
- якість: якість опису – це те, як деталі описуються, може включати

використання образної мови, сенсорних деталей та прикладів;

- структура: структура опису – це те, як деталі опису організовані, може включати використання вступу, основної частини та висновку.

**Типи опису.** Існує багато типів опису, які можна використовувати залежно від мети опису та теми. Деякі з найпоширеніших типів опису включають:

- об’єктивний опис: об’єктивний опис представляє інформацію про тему нейтрально та неупереджено;
- суб’єктивний опис: суб’єктивний опис представляє інформацію про тему з точки зору автора;
- технічний опис: технічний опис використовується для надання детальної та точної інформації про тему, зазвичай для фахівців;
- неформальний опис: неформальний опис використовується для надання загальної інформації про тему, зазвичай для широкої аудиторії;
- літературний опис: літературний опис використовується для створення у читача певного настрою або емоції за допомогою образної мови та сенсорних деталей.

#### **Приклади опису.**

- опис продукту: опис продукту використовується для надання інформації про продукт, його характеристики та переваги;
- опис послуги: опис послуги використовується для надання інформації про послугу, як вона надається та які переваги вона пропонує;
- опис місця: опис місця використовується для надання інформації про місце, його характеристики та визначні пам’ятки;
- опис людини: опис людини використовується для надання інформації про зовнішність, особистість та характер людини;
- опис події: опис події використовується для надання інформації про те, що сталося під час події, коли вона сталася та хто був задіяний [14].



## 3 ЗАСТОСУВАННЯ

### 3.1 Розпізнавання облич

Розпізнавання облич – це процес ідентифікації або верифікації особи за її зображенням. Це складна задача комп'ютерного бачення, яка використовується в різних сферах, таких як безпека, правоохоронна діяльність та маркетинг.

**Методи розпізнавання облич.** Існує два основних методи розпізнавання облич:

- методи на основі рис обличчя: ці методи ідентифікують особу на основі її унікальних рис обличчя, таких як очі, ніс, рот та форма обличчя;
- методи на основі шаблонів: ці методи створюють шаблон зображення обличчя особи, а потім порівнюють цей шаблон з іншими зображеннями для ідентифікації особи.

**Застосування розпізнавання облич.** Розпізнавання облич використовується в різних сферах, таких як:

- безпека: розпізнавання облич використовується для контролю доступу до будівель, комп'ютерів та інших безпечних зон;
- правоохоронна діяльність: розпізнавання облич використовується для ідентифікації підозрюваних та злочинців;
- маркетинг: розпізнавання облич використовується для персоналізації реклами та визначення настроїв клієнтів;
- соціальні мережі: розпізнавання облич використовується для маркування людей на фотографіях та відео.

**Етичні міркування.** Використання розпізнавання облич викликає ряд етичних міркувань, таких як:

- конфіденційність: збір та зберігання даних про обличчя людей може

становити загрозу для конфіденційності;

- упередженість: системи розпізнавання облич можуть бути упередженими щодо певних груп людей, що може призвести до дискримінації;
- зловживання: системи розпізнавання облич можуть використовуватися для відстеження та контролю людей без їхньої згоди.

**Майбутнє розпізнавання облич.** Технологія розпізнавання облич постійно розвивається, і очікується, що вона стане ще більш точною та потужною в найближчі роки (див. рис. 3.1). Це призведе до нових застосувань розпізнавання облич, а також до нових етичних міркувань [15].



Рисунок 3.1 – Технологія розпізнавання облич

Суть технології полягає в використанні камер з функцією розпізнавання облич для виявлення узлових точок та вимірювання відстані між певними точками на обличчі. Щоб отримати так звані відбитки обличчя камері необхідно виявити близько 80 таких точок. Справжнім проривом стало усвідомлення того, що відбиток обличчя є унікальним кодом, який є у кожної

людини та який можна зчитати на відстані.

Після отримання відбитка обличчя здійснюється пошук в базі даних для співвіднесення обличчя з іменем. Вірогідність помилки при цьому мінімальна. З 1000 сканів тільки 8 можуть бути ідентифіковані з помилкою. З цієї причини технології розпізнавання обличчя має широкі перспективи для застосування [16].

### **3.2 Розпізнавання об'єктів**

Розпізнавання об'єктів – це процес ідентифікації об'єктів на зображеннях або відео. Це складна задача комп'ютерного бачення, яка використовується в різних сферах, таких як робототехніка, самостійне керування автомобілями та візуальний пошук.

**Методи розпізнавання об'єктів.** Існує багато різних методів розпізнавання об'єктів, але деякі з найпоширеніших включають:

- методи на основі рис: ці методи ідентифікують об'єкти на основі їх унікальних рис, таких як форма, колір та текстура;
- методи на основі шаблонів: ці методи створюють шаблон об'єкта, а потім порівнюють цей шаблон з іншими зображеннями для ідентифікації об'єкта;
- методи на основі глибокого навчання: ці методи використовують штучні нейронні мережі для вивчення складних патернів на зображеннях, що дозволяє їм ідентифікувати об'єкти з високою точністю.

**Застосування розпізнавання об'єктів.** Розпізнавання об'єктів використовується в різних сферах, таких як:

- робототехніка: розпізнавання об'єктів може використовуватися для того, щоб роботи могли орієнтуватися в навколишньому середовищі та взаємодіяти з об'єктами;

- самостійне керування автомобілями: розпізнавання об'єктів може використовуватися для того, щоб самокеровані автомобілі могли ідентифікувати інші транспортні засоби, пішоходів та велосипедистів;
- візуальний пошук: розпізнавання об'єктів може використовуватися для того, щоб люди могли знаходити зображення певних об'єктів в Інтернеті;
- безпека: розпізнавання об'єктів може використовуватися для того, щоб ідентифікувати підозрілі об'єкти або людей.

**Складнощі розпізнавання об'єктів.** Розпізнавання об'єктів – це складна задача, і існують деякі виклики, які необхідно подолати, перш ніж воно зможе бути широко використано. Деякі з цих викликів включають:

- варіативність: об'єкти можуть виглядати по-різному залежно від освітлення, кута огляду та інших факторів;
- затьмарення: об'єкти можуть бути частково або повністю затьмарені іншими об'єктами;
- складність фону: об'єкти можуть бути розташовані на складному фоні, що може ускладнити їх ідентифікацію [17].

Технологія розпізнавання об'єктів продовжує постійно розвиватися. Це призводить до нових застосувань розпізнавання об'єктів, а також до подолання деяких із поточних проблем (див. рис. 3.2) [18].

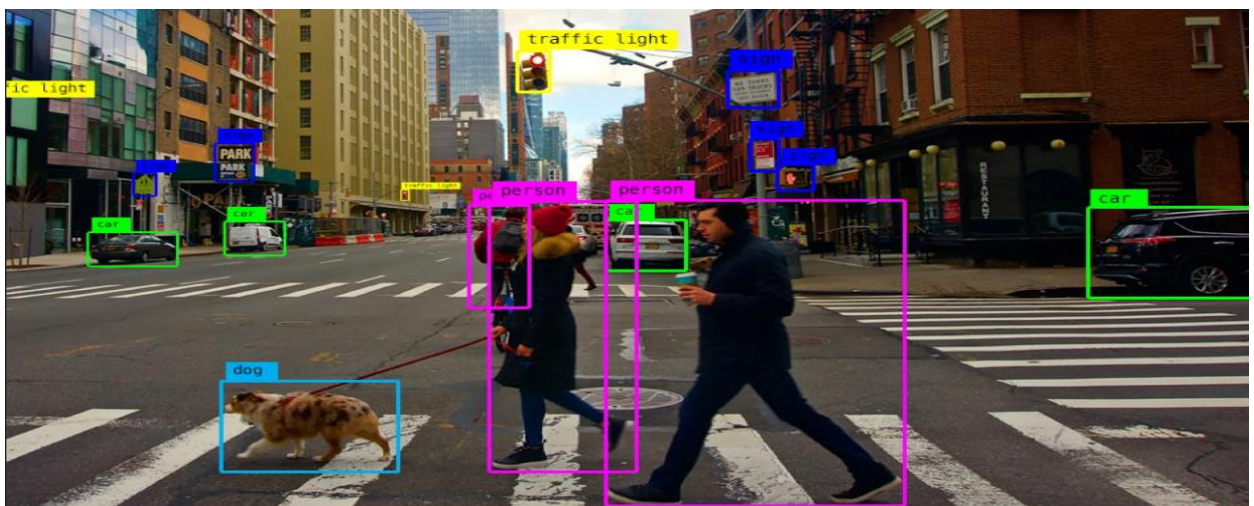


Рисунок 3.2 – Приклад розпізнавання об'єктів комп'ютерним зором

Впровадження виявлення об'єктів за допомогою глибокого навчання є складним завданням, але за допомогою правильних інструментів і методів воно здійсненне як для початківців, так і для досвідчених розробників.

Ключем до успіху є наявність добре анотованого набору даних, хороших методів попередньої обробки та ретельного налаштування гіперпараметрів моделі. Маючи їх на місці, можна навчити модель глибокого навчання для виявлення об'єктів, яку можна використовувати для різноманітних додатків, від відеоспостереження до безпілотних автомобілів.

Підводячи підсумок, реалізація виявлення об'єктів за допомогою глибокого навчання включає кілька кроків, включаючи анотування набору даних, попередню обробку зображень, навчання моделі, оцінку її продуктивності та використання для виявлення об'єктів. За допомогою правильних інструментів і методів можна створити надійну та точну систему виявлення об'єктів, яку можна використовувати для різноманітних застосувань [19].

### 3.3 Сегментація зображень

Сегментація зображень – це процес розбиття зображення на окремі сегменти, які відповідають певним об'єктам або регіонам інтересу. Це складна задача комп'ютерного бачення, яка використовується в різних сферах, таких як медична візуалізація, автономне керування та аналіз зображень.

**Методи сегментації зображень.** Існує багато різних методів сегментації зображень, але деякі з найпоширеніших включають:

- методи на основі порогового значення: ці методи сегментують зображення на основі його інтенсивності пікселів;
- методи на основі регіонів: ці методи сегментують зображення, групуючи пікселі з подібними властивостями, такими як колір або текстура;

- методи на основі навчання: ці методи використовують штучні нейронні мережі для вивчення того, як сегментувати зображення.

**Застосування сегментації зображень.** Сегментація зображень використовується в різних сферах, таких як:

- медична візуалізація: сегментація зображень може використовуватися для ідентифікації та вимірювання пухлин, органів та інших анатомічних структур;
- автономне керування: сегментація зображень може використовуватися для того, щоб самокеровані автомобілі могли ідентифікувати інші транспортні засоби, пішоходів та велосипедистів;
- аналіз зображень: сегментація зображень може використовуватися для виявлення та класифікації об'єктів на зображеннях;
- редагування зображень: сегментація зображень може використовуватися для видалення або зміни певних об'єктів на зображенні.

**Складнощі сегментації зображень.** Сегментація зображень – це складна задача, і існують деякі проблеми, які необхідно подолати, перш ніж вона зможе бути широко використано (див. рис. 3.3).

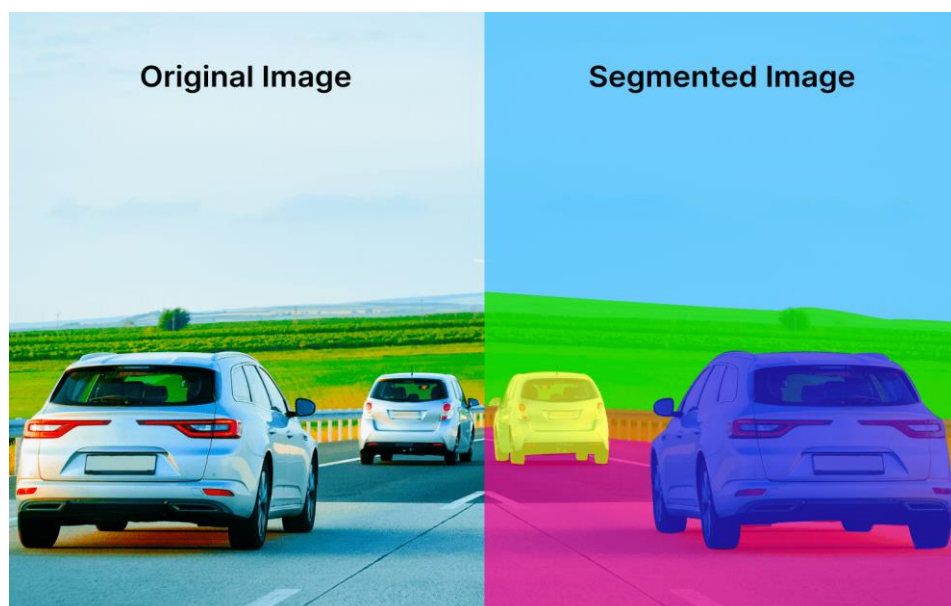


Рисунок 3.3 – Сегментація зображення

Деякі з них включають:

- шум: зображення можуть бути зашумлені, що може ускладнити сегментацію;
- складність фону: об'єкти можуть бути розташовані на складному фоні, що може ускладнити їх сегментацію;
- змінність: об'єкти можуть виглядати по-різному залежно від освітлення, кута огляду та інших факторів [20].

Які найпоширеніші прийоми сегментації зображення? Сегментація зображення є критично важливим завданням комп'ютерного зору та обробки зображень. Існують різні методи сегментації зображення.

**Сегментація на основі ребер.** Сегментація на основі країв – це популярна техніка, яка визначає краї об'єктів на зображенні. Ця техніка використовує алгоритми виявлення країв для визначення місцезнаходження пов'язаних об'єктів за допомогою інформації з країв. Алгоритми сегментації на основі країв можуть точно представляти межі об'єктів на зображенні за допомогою ланцюжків країв, що складаються з окремих країв.

**Сегментація на основі порогового значення.** Сегментація на основі порогу – це проста техніка, яка сегментує зображення на основі інтенсивності пікселів відносно заданого значення або порогу. Він може точно відокремити об'єкти з більшою інтенсивністю від інших об'єктів або фону. Порогове значення ділить зображення в градаціях сірого на два сегменти на основі їх співвідношення з пороговим значенням.

**Сегментація на основі регіону.** Сегментація на основі регіонів передбачає поділ зображення на регіони зі схожими характеристиками. Алгоритм знаходить вихідну точку та збільшує регіони, додаючи більше пікселів або звужуючи та об'єднуючи їх з іншими точками. Ця техніка корисна для сегментації об'єктів зі схожою текстурою або кольорами.

**Сегментація на основі кластерів.** Сегментація на основі кластерів використовує алгоритми кластеризації для групування пікселів зі схожими характеристиками в кластери. Ця техніка корисна для виявлення прихованої інформації на зображеннях, наприклад візерунків або структур. Алгоритм



розбиває зображення на кластери пікселів, відокремлюючи елементи даних і групуєчи подібні елементи в кластери.

**Сегментація вододілу.** Це техніка, яка розглядає зображення як топографічні карти, де яскравість пікселів визначає висоту. Ця техніка виявляє лінії, що утворюють хребти та западини, позначаючи ділянки між лініями вододілу. Він розділяє зображення на кілька областей на основі висоти пікселів, групуєчи пікселі з однаковим значенням сірого. Ця техніка корисна для обробки медичних зображень і може допомогти визначити відмінності між світлішими та темнішими ділянками МРТ-сканування, потенційно допомагаючи в діагностиці.

Сегментація зображень вже довела, що є цінним інструментом у різних сферах. Оскільки технологія продовжує розвиватися, ми можемо очікувати ще більше інноваційних застосувань цієї техніки в майбутньому. Одним із нових потенційних застосувань сегментації зображень є сфера безпеки. Сегментуючи зображення людей у громадських місцях, співробітники служби безпеки можуть швидко визначити потенційні загрози та вжити необхідних заходів. Це може допомогти запобігти злочинам і зберегти громадську безпеку.

Оскільки в галузях промисловості будь-якого розміру та галузей зростає інтерес до програм комп'ютерного зору, сегментація зображень є основною основою, яка впливає на більшість програм комп'ютерного зору. Зростання сегментації зображень у найближчі роки буде швидким, переважно в аналізі геопросторових даних, відеоспостереження, медичних зображень та аналізу супутникових зображень [21].

### **3.4 Відстеження об'єктів**

Відстеження об'єктів – це процес відстеження руху та ідентифікації об'єктів у послідовності зображень або відео (див. рис. 3.4). Це складна задача комп'ютерного зору, яка використовується в різних сферах, таких як



відеоспостереження, аналіз руху та автономне керування.

**Методи відстеження об'єктів.** Існує багато різних методів відстеження об'єктів, але деякі з найпоширеніших включають:

- методи на основі фільтрів: ці методи використовують фільтри, такі як фільтр Калмана, для прогнозування розташування об'єкта в наступному кадрі;
- методи на основі відповідності: ці методи зіставляють об'єкти в одному кадрі з об'єктами в наступному кадрі на основі їх характеристик, таких як колір або форма;
- методи на основі глибокого навчання: ці методи використовують штучні нейронні мережі для вивчення того, як відстежувати об'єкти в відео.

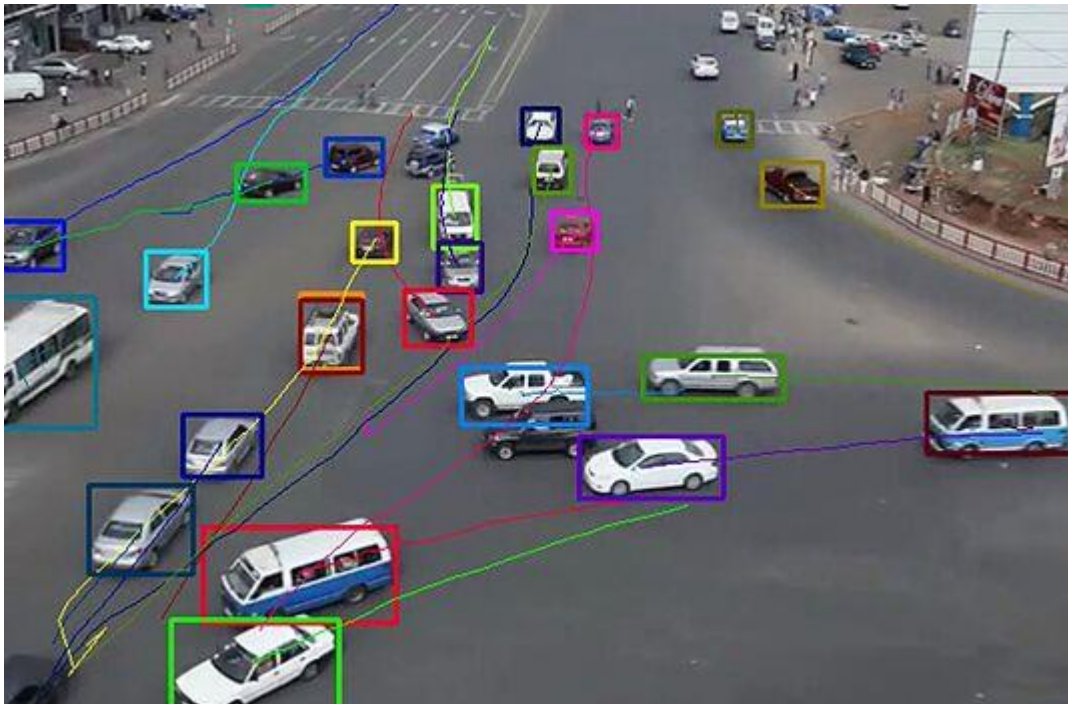


Рисунок 3.4 – Відстеження об'єктів комп'ютерним зором

**Застосування відстеження об'єктів.** Відстеження об'єктів використовується в різних сферах, таких як:

- відеоспостереження: відстеження об'єктів може використовуватися для відстеження підозрілих осіб або діяльності;

- аналіз руху: відстеження об'єктів може використовуватися для підрахунку кількості людей або транспортних засобів, які проїжджають через певну область;
- автономне керування: відстеження об'єктів може використовуватися для того, щоб самокеровані автомобілі могли відстежувати інші транспортні засоби, пішоходів та велосипедистів;
- інтерфейси людина-машина: відстеження об'єктів може використовуватися для керування комп'ютером або іншим пристроєм за допомогою рухів рук або тіла.

#### **Складнощі відстеження об'єктів:**

- затьмарення: об'єкти можуть бути частково або повністю затьмарені іншими об'єктами;
- зміна освітлення: умови освітлення можуть змінюватися протягом відео, що може ускладнити відстеження об'єктів;
- складність фону: об'єкти можуть бути розташовані на складному фоні, що може ускладнити їх відстеження [22].

Практичні приклади відстеження об'єктів.

**Відеоспостереження.** Відстеження об'єктів широко використовується в системах відеоспостереження для моніторингу та аналізу діяльності в режимі реального часу. Це дозволяє автоматично відстежувати осіб або об'єкти, що представляють інтерес, наприклад, підозрілих осіб у натовпі або неавторизованих транспортних засобів у заборонених зонах. Відстеження об'єктів підвищує ефективність роботи персоналу служби безпеки, надаючи їм точну та своєчасну інформацію.

**Керування трафіком.** Відстеження об'єктів використовується в системах керування дорожнім рухом для моніторингу та аналізу руху транспортних засобів на дорогах. Це дозволяє відстежувати транспортні засоби, щоб збирати дані про дорожній рух, виявляти затори та оптимізувати час сигналів світлофора. Відстежуючи транспортні засоби, системи управління дорожнім рухом можуть реалізовувати інтелектуальні алгоритми

для адаптивного контролю руху та ефективного маршрутизації.

**Доповнена реальність (AR).** Відстеження об'єктів є важливим компонентом додатків доповненої реальності. AR використовує відстеження об'єктів для накладення віртуальних об'єктів або інформації на сцену реального світу. Наприклад, в іграх із доповненою реальністю відстеження об'єктів дозволяє віртуальним об'єктам взаємодіяти з фізичними об'єктами в режимі реального часу, покращуючи відчуття занурення. Відстеження об'єктів також дозволяє накладати віртуальні анотації або інформацію на певні об'єкти чи місця.

**Спортивний аналіз.** Відстеження об'єктів знаходить застосування в аналізі спорту, де воно допомагає відстежувати спортсменів і об'єкти під час ігор або подій. Відстежуючи гравців, м'ячі чи інші ігрові елементи, спортивні аналітики можуть отримати цінну інформацію, як-от моделі руху гравців, траєкторію м'яча та статистику гри. Відстеження об'єктів полегшує поглиблений аналіз, оцінку ефективності та прийняття стратегічних рішень у різних видах спорту, включаючи футбол, баскетбол і теніс.

**Робототехніка.** Відстеження об'єктів відіграє вирішальну роль у роботизованих системах, дозволяючи роботам відстежувати об'єкти в своєму оточенні та взаємодіяти з ними. Наприклад, у промисловій автоматизації роботи, оснащені можливостями відстеження об'єктів, можуть точно знаходити об'єкти на складальних лініях і керувати ними. У сфері охорони здоров'я відстеження об'єктів дозволяє роботам-хірургам відстежувати рух хірургічних інструментів і стежити за ними, підвищуючи точність і безпеку.

**Автономні транспортні засоби.** Відстеження об'єктів є основним компонентом автономних транспортних засобів, що дозволяє їм сприймати та розуміти оточення. Безперервно відстежуючи пішоходів, транспортні засоби та дорожні знаки, автономні транспортні засоби можуть приймати обґрунтовані рішення, такі як дотримання смуги руху, адаптивний круїз-контроль і уникнення зіткнень. Відстеження об'єктів сприяє безпечній та ефективній роботі автономних транспортних засобів у різних середовищах.

Відстеження об'єктів є основним компонентом програм комп'ютерного зору, що дозволяє оцінювати та контролювати положення та траєкторії об'єктів. Ми обговорили відмінності між відстеженням і виявленням об'єктів, дослідили алгоритми KCF і CSRT з їхніми математичними основами та надали зразок коду Python для відстеження на основі KCF. Порівнюючи сильні та слабкі сторони KCF і CSRT, ми підкреслили потенціал методів глибокого навчання для підвищення точності. Розуміючи теоретичні основи та етапи впровадження, дослідники та розробники можуть зробити внесок у розвиток комп'ютерного зору. Моделі глибокого навчання пропонують переваги, але також мають обмеження. Відстеження об'єктів залишається динамічною сферою з постійним прогресом у алгоритмах і глибокому навчанні. З розвитком технологій відстеження об'єктів продовжує відігравати важливу роль у різних програмах, змінюючи спосіб сприйняття та взаємодії з візуальним світом [23].

### **3.5 Генерація опису зображень**

Генерація опису зображень (ГОЗ) – це процес автоматичного створення текстового опису зображення. Це складна задача штучного інтелекту, яка поєднує в собі комп'ютерний зір та обробку природної мови. ГОЗ має багато потенційних застосувань, таких як допомога людям з вадами зору, створення підписів до зображень у соціальних мережах та автоматизація завдань з описування зображень (див. рис. 3.5).

**Методи генерації опису зображень.** Існує два основних підходи до ГОЗ.

Підхід на основі шаблонів: цей підхід використовує заздалегідь визначені шаблони для генерування опису зображення. Шаблони зазвичай включають опис об'єктів, людей та сцен на зображенні, а також їх взаємозв'язків.

Підхід на основі глибокого навчання: цей підхід використовує штучні нейронні мережі для навчання на великих наборах даних зображень та відповідних описів. Нейронні мережі навчаються виявляти закономірності між зображеннями та їхніми описами, а потім використовують ці закономірності для генерування нових описів.

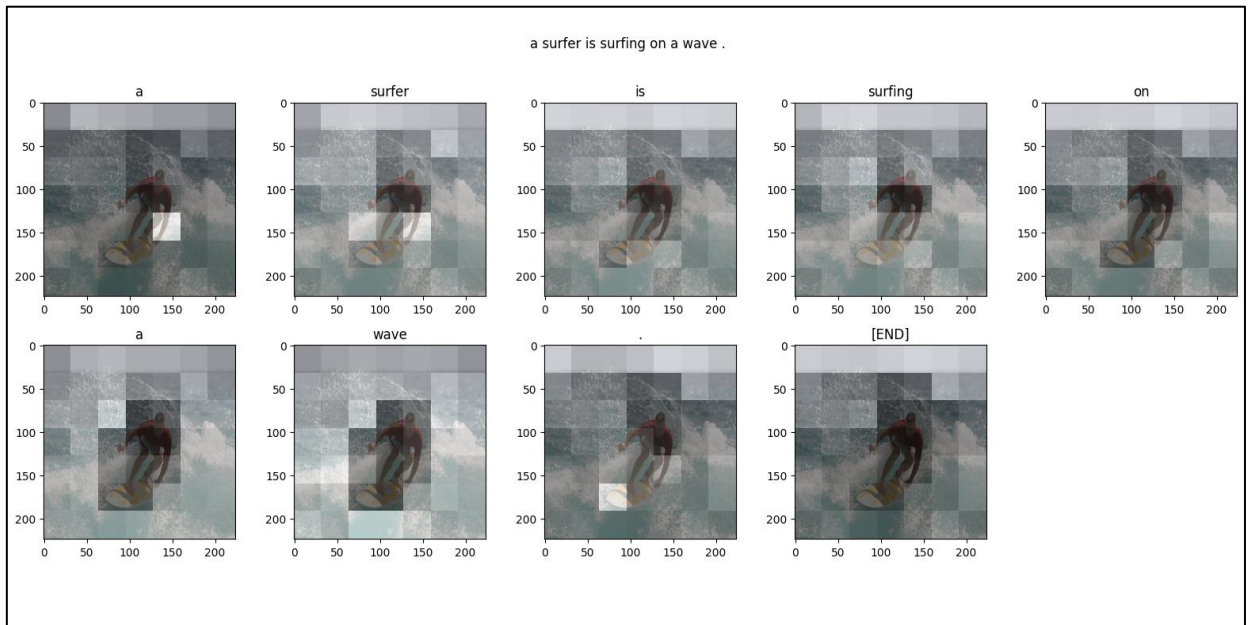


Рисунок 3.5 – Процес опису зображення методом обробки фото

**Застосування генерації опису зображень.** ГОЗ має багато потенційних застосувань, таких як:

- допомога людям з вадами зору: ГОЗ може використовуватися для створення текстових описів зображень, які можуть бути прочитані людьми з вадами зору за допомогою програм читання з екрану;
- створення підписів до зображень у соціальних мережах: ГОЗ може використовуватися для автоматичного створення підписів до зображень, які публікуються в соціальних мережах;
- автоматизація завдань з описування зображень: ГОЗ може використовуватися для автоматизації завдань з описування зображень, таких як створення підписів до фотографій у продуктових каталогах або описів сцен у медичних зображеннях.

### **Складнощі генерації опису зображень:**

- складність природної мови: природна мова складна і багатогранна, що може ускладнити генерування точних і граматично правильних описів зображень;
- неоднозначність: зображення можуть мати кілька можливих описів, і може бути складно визначити, який опис є найкращим;
- відсутність даних: для навчання моделей ГОЗ потрібні великі набори даних зображень та відповідних описів. Створення таких наборів даних може бути дорогим і трудомістким [24].

**Генератор підписів до зображень із використанням глибокого навчання.** Генератори підписів до зображень – це захоплююче застосування глибокого навчання, яке поєднує розрив між комп’ютерним зором і обробкою природної мови (НЛП). Ось розбивка того, як вони працюють.

**Згорткові нейронні мережі (CNN): це робочі конячки для розпізнавання зображень.** Вони приймають зображення як вхідні дані та обробляють його через кілька шарів, вилучаючи такі елементи, як форми, краї та кольори.

Мережі довготривалої короткочасної пам’яті (LSTM): це особливий тип повторюваних нейронних мереж (RNN), які вправно обробляють послідовні дані, такі як речення. У субтитрах до зображень LSTM беруть характеристики зображення, отримані CNN, і використовують їх для прогнозування по одному слову, формуючи підпис, який описує зміст зображення.

### **Процес:**

- підготовка даних: потрібен великий набір даних зображень із відповідними підписами – це навчає модель розпізнавати зв’язок між візуальними функціями та відповідними мовними описами;
- вилучення функцій: CNN робить зображення та витягує набір функцій, які представляють його візуальний зміст;
- прогнозування послідовності: мережа LSTM отримує витягнуті функції та починає генерувати підпис слово за словом, він враховує

раніше згенеровані слова та особливості зображення, щоб передбачити наступне слово в послідовності – це триває до тих пір, поки не буде сформований повний титр.

**Застосування.** Генератори підписів до зображень мають ряд застосувань, зокрема:

- покращення доступності зображень: генеруючи підписи до зображень, користувачі з вадами зору можуть зрозуміти вміст;
- автоматичне позначення тегами зображень: підписи можна використовувати для автоматичного позначення тегами зображень для кращої організації та пошуку;
- описи зображень у соціальних мережах: можна створювати підписи для публікацій у соціальних мережах, щоб покращити залучення [25].

## ВИСНОВКИ

Анотація даних є критично важливим елементом у сфері машинного навчання та комп'ютерного зору. Вона забезпечує підготовку високоякісних тренувальних наборів даних, що є необхідними для навчання моделей штучного інтелекту. Існують різні методи анотації, кожен з яких має свої переваги та недоліки.

Ручна анотація є найбільш точним методом, оскільки здійснюється безпосередньо людиною. Вона дозволяє забезпечити високий рівень деталізації та точності, проте є дуже трудомісткою і часо-витратною. Ручна анотація широко використовується в проєктах, де необхідна висока точність, наприклад, у медицині та наукових дослідженнях.

Напів-автоматична анотація комбінує переваги ручної та автоматичної анотації. Цей метод включає використання програмного забезпечення для попередньої обробки даних, після чого людина-коректор виконує остаточні правки. Такий підхід значно скорочує час, необхідний для анотації, одночасно підтримуючи високу якість результатів.

Автоматична анотація використовує алгоритми машинного навчання для повної автоматизації процесу анотації. Цей метод є найшвидшим і найменш затратним з точки зору людських ресурсів, проте може поступатися в точності ручній та напів автоматичній анотації. Автоматична анотація часто використовується для великих обсягів даних, де необхідна оперативність.

Типи анотацій включають класифікацію, локалізацію, сегментацію, відстеження та опис. Кожен тип має своє специфічне застосування залежно від завдань проєкту. Наприклад, класифікація використовується для розпізнавання об'єктів, локалізація для визначення їхнього розташування, а сегментація для детального виділення об'єктів на зображеннях.

Таким чином, правильний вибір методу анотації та типу анотації є критично важливими для успішного виконання проєктів у сфері машинного



навчання та комп'ютерного зору. Кожен метод і тип аотації має свої особливості, які визначають їх доцільність залежно від конкретних вимог та обмежень проєкту.

## ПЕРЕЛІК ПОСИЛАНЬ

1. Facial recognition. URL: [https://en.wikipedia.org/wiki/Facial\\_recognition](https://en.wikipedia.org/wiki/Facial_recognition) (дата звернення: 15.03.2024).
2. Object detection. URL: [https://en.wikipedia.org/wiki/Object\\_detection](https://en.wikipedia.org/wiki/Object_detection) (дата звернення: 15.03.2024).
3. Image segmentation. URL: [https://en.wikipedia.org/wiki/Image\\_segmentation](https://en.wikipedia.org/wiki/Image_segmentation) (дата звернення: 15.03.2024).
4. Text classification. URL: [https://en.wikipedia.org/wiki/Text\\_classification](https://en.wikipedia.org/wiki/Text_classification) (дата звернення: 17.03.2024).
5. Localization. URL: <https://en.wikipedia.org/wiki/Localization> (дата звернення: 26.03.2024).
6. Localization Strategy. URL: <https://www.contentstack.com/cms-guides/localization-strategy-global-content-operation> (дата звернення: 26.03.2024).
7. Best Localization Software. URL: <https://www.g2.com/categories/localization> (дата звернення: 26.03.2024).
8. Market Segmentation. URL: [https://en.wikipedia.org/wiki/Market\\_segmentation](https://en.wikipedia.org/wiki/Market_segmentation) (дата звернення: 28.03.2024).
9. How to Segment Your Market: A Step-by-Step Guide. URL: <https://academy.hubspot.com/lessons/understanding-segmentation-in-hubspot> (дата звернення: 28.03.2024).
10. 7 Effective Market Segmentation Strategies to Target Your Audience. URL: <https://support.sproutsocial.com/hc/en-us/articles/205047165-How-do-I-use-Audience-Targeting-in-Sprout> (дата звернення: 28.03.2024).
11. Tracking (computer science). URL: <https://en.wikipedia.org/wiki/Tracking> (дата звернення: 02.04.2024).
12. Electronic Frontier Foundation: Tracking. URL:

- <https://www.eff.org/press/releases/eff-report-exposes-explains-big-techs-personal-data-trackers-lurk-social-media> (дата звернення: 02.04.2024).
13. The Ultimate Guide to Online Privacy. URL: <https://www.vpn.com/privacy/faq/vpn-for-internet-privacy/> (дата звернення: 02.04.2024).
  14. Description. URL: <https://en.wikipedia.org/wiki/Description> (дата звернення: 02.04.2024).
  15. Face recognition URL: [https://en.wikipedia.org/wiki/Face\\_recognition](https://en.wikipedia.org/wiki/Face_recognition) (дата звернення: 30.04.2024).
  16. Як працює розпізнавання обличч? URL: [https://kristall-systems.net.ua/ua/novosti/kak\\_rabotaet\\_raspoznavanje\\_lits\\_perspektivyi\\_tehnologii\\_v\\_ukraine/](https://kristall-systems.net.ua/ua/novosti/kak_rabotaet_raspoznavanje_lits_perspektivyi_tehnologii_v_ukraine/) (дата звернення: 30.04.2024).
  17. Object detection. URL: [https://en.wikipedia.org/wiki/Object\\_detection](https://en.wikipedia.org/wiki/Object_detection) (дата звернення: 30.04.2024).
  18. The Future of Object Recognition Technology. URL: <https://www.gartner.com/en/information-technology/glossary/image-recognition> (дата звернення: 30.04.2024).
  19. How to Implement Object Detection Using Deep Learning: A Step-by-Step Guide. URL: <https://www.augmentedstartups.com/blog/how-to-implement-object-detection-using-deep-learning-a-step-by-step-guide> (дата звернення: 30.04.2024).
  20. Image segmentation. URL: [https://en.wikipedia.org/wiki/Image\\_segmentation](https://en.wikipedia.org/wiki/Image_segmentation) (дата звернення: 01.05.2024).
  21. Image Segmentation: The Most Interesting Applications. URL: <https://deeplobe.ai/image-segmentation-the-most-interesting-applications/> (дата звернення: 01.05.2024).
  22. Video tracking. URL: [https://en.wikipedia.org/wiki/Video\\_tracking](https://en.wikipedia.org/wiki/Video_tracking) (дата звернення: 01.05.2024).
  23. Object tracking. URL: <https://medium.com/@khwabkalra1/object-tracking-2fe4127e58bf> (дата звернення: 01.05.2024).

24. Image captioning. URL:  
[https://en.wikipedia.org/wiki/Natural\\_language\\_generation#Image\\_captioning](https://en.wikipedia.org/wiki/Natural_language_generation#Image_captioning)  
(дата звернення: 02.05.2024).
25. Image Caption Generator. URL:  
<https://www.analyticsvidhya.com/blog/2021/12/step-by-step-guide-to-build-image-caption-generator-using-deep-learning/> (дата звернення: 02.05.2024).