

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
КАФЕДРА КОМП'ЮТЕРИЗОВАНИХ СИСТЕМ ЗАХИСТУ ІНФОРМАЦІЇ**

ДОПУСТИТИ ДО ЗАХИСТУ

**Завідувач кафедри Комп'ютеризованих
систем захисту інформації**

_____ Михайло СТЕПАНОВ

« ____ » _____ 2023 р.

На правах рукопису
УДК 004.056.5:510.22(043.3)

**КВАЛІФІКАЦІЙНА РОБОТА
ЗДОБУВАЧА ВИЩОЇ ОСВІТИ
ОСВІТНЬОГО СТУПЕНЯ «МАГІСТР»**

Тема: Програмний застосунок виявлення недозволених предметів на підприємстві за допомогою нейронних мереж

Виконавець: Григорій ВОЛОШИН

Керівник: к.т.н. Олена ВИСОЦЬКА

**Консультант розділу «Охорона
навколишнього середовища»:** к.т.н., доцент Тетяна ДМИТРУХА

Нормоконтролер: к.т.н. Олена ВИСОЦЬКА

Київ 2023

НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Факультет: Кібербезпеки та програмної інженерії

Кафедра: Комп'ютеризованих систем захисту інформації

Освітній ступінь: Магістр

Спеціальність: 125 «Кібербезпека»

Освітньо-професійна програма: «Безпека інформаційних і комунікаційних систем»

ЗАТВЕРДЖУЮ

Завідувач кафедри

Комп'ютеризованих систем захисту
інформації

_____ Михайло СТЕПАНОВ

«__» _____ 2023 р.

ЗАВДАННЯ

на виконання кваліфікаційної роботи

здобувача вищої освіти Волошина Григорія Володимировича

1. *Тема: Програмний застосунок виявлення недозволених предметів на підприємстві за допомогою нейронних мереж*

затверджена наказом ректора від «15» вересня 2023 р. № 1814/ст.

2. Термін виконання: з 16.10.2023 р. по 31.12.2023 р.

3. Вихідні дані: проаналізувати технології аналізу та обробки зображень і методи розпізнавання об'єктів на зображенні, та екологічні рухи в світі, а також існуючі архітектури нейронних мереж у розпізнаванні об'єктів на зображенні; розробити власну архітектуру згорткової нейронної мережі та розробити програмний додаток, провести тестування розробленого програмного додатку.

4. Зміст пояснювальної записки: аналіз та обробка зображень; розробка алгоритму та програмної реалізації системи моніторингу та тестування та демонстрація роботи системи, аналіз екологічного руху в світі.

5. КАЛЕНДАРНИЙ ПЛАН виконання кваліфікаційної роботи

№ з/п	Етапи виконання кваліфікаційної роботи	Термін виконання етапів	Примітка
1.	Уточнення постановки задачі	16.10.2023	<i>Виконано</i>
2.	Аналіз літературних джерел	20.10.2023	<i>Виконано</i>
3.	Обґрунтування вибору рішення	24.10.2023	<i>Виконано</i>
4.	Збір та аналіз інформації про технології аналізу та обробки зображень	26.10.2023	<i>Виконано</i>
5.	Збір та аналіз інформації про методи розпізнавання об'єктів на зображенні	02.11.2023	<i>Виконано</i>
6.	Збір та аналіз інформації про існуючі аналоги	17.11.2023	<i>Виконано</i>
7.	Розробка додатку та моделі машинного навчання та проведення	22.11.2023	<i>Виконано</i>
8.	Проведення експериментів	25.11.2023	<i>Виконано</i>
9.	Оформлення і друк пояснювальної записки	07.12.2023	<i>Виконано</i>
10.	Перевірка на антиплагіат	08.12.2023	<i>Виконано</i>
11.	Оформлення презентації	10.12.2023	<i>Виконано</i>
12.	Отримання рецензій від рецензента	18.12.2023	<i>Виконано</i>

6. Консультанти з окремих розділів

Розділ	Консультант (посада, П.І.Б.)	Дата, підпис	
		Завдання видав	Завдання прийняв
Охорона навколишнього середовища	Дмитруха Т.І.		

7. Дата видачі завдання: «16» жовтня 2023 р.

Здобувач вищої освіти

(підпис, дата)

Керівник кваліфікаційної роботи

(підпис, дата)

Григорій ВОЛОШИН

Олена ВИСОЦЬКА

РЕФЕРАТ

Кваліфікаційна робота на тему: «Програмний застосунок виявлення недозволених предметів на підприємстві за допомогою нейронних мереж» складається з вступу, основної частини, яка включає п'ять розділів, висновків до кожного розділу, загального висновку, додатку та списку використаної літератури і має 96 сторінок основного тексту. Загальний обсяг роботи становить 107 сторінок, що включає 20 рисунків та 3 таблиці. Використано 15 джерел літератури.

Запропонована нова структура згорткової нейронної мережі, яка має покращену архітектуру для більш ефективного навчання та розпізнавання.

Метою роботи є розробка програмного застосунку, який використовує методи комп'ютерного зору та нейронні мережі для виявлення та ідентифікації недозволених предметів на підприємствах, забезпечуючи високий рівень безпеки та автоматизацію процесів контролю.

У кваліфікаційній роботі розглянуті питання щодо сучасних методів розпізнавання об'єктів та класифікації зображень.

Проведені дослідження базуються на сучасних методах виявлення об'єктів на зображенні. Основна увага приділяється розробці згорткової нейронної мережі на основі глибокого навчання, її інтеграції у програмний застосунок та оцінці ефективності застосування розробленого рішення.

Створений застосунок дозволяє забезпечити високу точність класифікації та завдяки цьому забезпечити надійний захист підприємств.

Ключові слова: нейронні мережі, розпізнавання об'єктів, безпека на підприємствах, обробка цифрових зображень, програмний застосунок.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ	
СКОРОЧЕНЬ.....	7
ВСТУП.....	8
1. ВИЯВЛЕННЯ НЕДОЗВОЛЕНИХ ПРЕДМЕТІВ НА	
ЗОБРАЖЕННЯХ.....	12
1.1. Аналіз актуальності теми.....	12
1.2 Роль технологій виявленні недозволених предметів.....	14
1.3 Категорія недозволених предметів.....	18
1.4 Технології аналізу та обробки зображень і методи розпізнавання об'єктів на зображенні.....	20
1.5 Висновки до першого розділу.....	29
2. АНАЛІЗ МЕТОДІВ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА ЗОБРАЖЕННІ.....	30
2.1. Аналіз методів аналізу зображень.....	30
2.2. Дослідження попередньої обробки зображень.....	34
2.3. Аналіз нейронних мереж в розпізнаванні об'єктів.....	43
2.4. Дослідження специфіки застосування нейронних мереж для виявлення недозволених предметів.....	51
2.5. Аналіз і порівняння існуючих аналогів.....	53
2.6 Висновки до другого розділу.....	55
3. РОЗРОБКА АРХІТЕКТУРИ ПРОГРАМНОГО РІШЕННЯ.....	57
3.1 Аналіз програмного рішення.....	57
3.2 Дослідження підготовки та обробки дани.....	58
3.3 Опис розробки та архітектура моделі.....	61
3.4 Навчання, оцінка та візуалізація результатів моделі.....	70
3.5 Висновки до третього розділу.....	72
4. ОПИС, ТЕСТУВАННЯ ТА АНАЛІЗ РОБОТИ ЗАСТОСУНКУ.....	74
4.1 Вступ до четвертого розділу.....	74

4.2 Аналіз тестування та проведених експериментів.....	74
4.3 Тестування роботи застосунку.....	82
4.5 Аналіз порівняння з аналогами.....	86
4.6 Висновки до четвертого розділу.....	88
5. ОХОРОНА НАВКОЛИШНЬОГО СЕРЕДОВИЩА	90
ВИСНОВКИ.....	93
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	94
Додаток А.....	97

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

CNN - convolutional neural network, згорткові нейронні мережі

RNN - recurrent neural networks, рекурентні нейронні мережі

НМ - нейронна мережа

ЄС - Європейський Союз

HOG - histograms of oriented gradients, гістограми направлених градієнтів

RPN - Region Proposal Networks, регіональні мережі пропозицій

SSD - Single Shot Detector, одиночний детектор

YOLO - You Only Look Once, дивишся лише раз

RGB - Red Green Blue , червоний зелений синій

СМЯК - Cyan Magenta Yellow, Key чи Black, блакитний пурпурний жовтий чорний

GANs - Generative adversarial networks, генеративні змагальні мережі

МРТ - Магнітно-резонансна томографія

КТ - Комп'ютерна томографія

IoT - internet of things, Інтернет речей

ReLU - rectified linear unit, віпрямлений лінійний вузол

FC шари - fully connected, повнозв'язні шари

CPU - central processing unit, центральний процесор

GPU - graphics processing unit, графічний процесор

VGG - Visual Geometry Group, група візуальної геометрії

API - application programming interface, прикладний програмний інтерфейс

AWS - Amazon Web Services

ВСТУП

Актуальність дослідження. У сучасному світі, який характеризується динамічним розвитком технологій та зростанням глобалізаційних процесів, питання безпеки набуває особливої актуальності. Особливо це стосується інцидентів, пов'язаних з недозволеними предметами, які мають серйозні соціально-економічні, юридичні та психологічні наслідки, як для окремих осіб, так і для суспільства в цілому. З точки зору безпеки праці, інциденти на підприємствах представляють собою значущу проблему. Вони можуть призвести до фізичних травм, смертельних наслідків, а також мати серйозні юридичні наслідки. Статистика інцидентів на робочих місцях в різних країнах свідчить про високий рівень ризику, який необхідно враховувати та знижувати. У сфері промислової безпеки існує потреба в постійному удосконаленні та адаптації технологій до змінюваних умов та викликів. Розвиток технологій виявлення недозволених предметів, таких як системи відеоспостереження з інтеграцією інтелектуальних алгоритмів, автоматизовані системи розпізнавання та інші, відкривають нові можливості для підвищення безпеки. З огляду на важливість цього питання, ця робота акцентує увагу на необхідності глибокого аналізу існуючих ризиків, пов'язаних з інцидентами на підприємствах, а також розглядає сучасні технології та методики, які можуть бути використані для їх ідентифікації, прогнозування та запобігання. Таким чином, актуальність даного дослідження полягає у вирішенні важливих завдань забезпечення безпеки на підприємствах, а також у розробці та впровадженні ефективних технологій для виявлення та запобігання інцидентам, пов'язаним з недозволеними предметами. Особлива увага приділяється ролі технологій у цьому процесі, зокрема, використанню штучного інтелекту та інтеграції різноманітних систем безпеки для створення більш надійних та ефективних рішень.

Відомі підходи до вирішення поставленої задачі. У сфері розпізнавання об'єктів на зображеннях використовуються як класичні, так і сучасні методи, засновані на глибокому навчанні. Класичні методи включають фільтр Гаусса для шумоподавлення, порогову обрізку та метод шаблонів для ідентифікації об'єктів, а також методи, які базуються на гістограмах направлених градієнтів (HOG) та системі Віоли-Джонса для ефективного розпізнавання облич та інших об'єктів. З іншого боку, методи на основі глибокого навчання, такі як Region Proposal Networks (RPN), Single Shot MultiBox Detector (SSD), згорткові нейронні мережі (CNN) та YOLO (You Only Look Once), використовуються для більш складних задач розпізнавання об'єктів, надаючи високу точність та ефективність навіть у складних умовах. Ці методи сприяють автоматичному витягу особливостей з зображень та ефективній класифікації, що робить їх незамінними у різноманітних застосуваннях від безпеки до нагляду.

Метою роботи є розробка програмного застосунку для виявлення недозволених предметів на підприємствах на основі нейронних мереж.

Для досягнення поставленої мети необхідно розв'язати наступні задачі:

1. Проаналізувати сучасні методи цифрової обробки зображень та алгоритми розпізнавання об'єктів, особливо з використанням глибокого навчання та на основі результату проведеного аналізу обрати оптимальні методи для вирішення задачі виявлення недозволених предметів.

2. Розробити програмний застосунок виявлення недозволених предметів на підприємстві, який буде інтегрувати згорткові нейронні мережі (CNN) та технології обробки зображень, що дасть змогу ефективно виявляти та класифікувати об'єкти в різних умовах.

3. Провести тестування розробленого програмного застосунку, що дасть змогу дослідити доцільність використання розробленого застосунку для вирішення поставленої задачі, з акцентом на точність виявлення та швидкість обробки даних.

Галузь застосування розробленого програмного застосунку для виявлення недозволених предметів із застосуванням нейронних мереж має широке поле застосування, особливо у сферах промислової та комерційної безпеки. Він може бути використаний на виробничих об'єктах, у транспортних вузлах, а також у місцях суспільного зібрання для забезпечення безпеки та моніторингу. Також цей застосунок є корисним для правоохоронних органів у попередженні та реагуванні на потенційні загрози.

Об'єкт дослідження - процес виявлення недозволених об'єктів за допомогою методів комп'ютерного зору та нейронних мереж, а саме аналіз зображень, виявлення об'єктів і їх класифікацію за допомогою алгоритмів глибокого навчання.

Предметом дослідження є методи виявлення та класифікації об'єктів на зображенні на основі згорткових нейронних мереж та методів препроцесингу зображень.

Методи дослідження базуються на основі порівняльного методу існуючих технологій виявлення об'єктів на зображенні (для розробки згорткової нейронної мережі), функціонального програмування (програмної реалізації згорткової нейронної мережі) та експериментального методу (для налаштування параметрів розробленої нейронної мережі).

Наукова новизна роботи полягає в розробці програмного застосунку для виявлення недозволених предметів на підприємствах на основі нейронних мереж, в якому за рахунок розробленої архітектури згорткової нейронної мережі та налаштованим гіперпараметрам досягається висока точність класифікації при обмежених обчислювальних ресурсах та малій кількості тренувальних даних.

Практична цінність отриманих результатів полягає у розробці програмного застосунку на основі алгоритмів глибокого навчання та цифрової обробки зображень для ефективного виявлення недозволених предметів, що підвищує рівень безпеки у багатьох сферах та забезпечує високу точність класифікації з оптимізованим використанням обчислювальних ресурсів.

Апробація. Основні положення роботи доповідалися та обговорювалися на міжнародній науково-практичній конференції «Живучість та резильєнтність – 2023»:

- Волошин Г.В., Виявлення недозволених предметів на зображеннях за допомогою нейронних мереж // Живучість та резильєнтність – 2023: міжнародна науково-практична конференція 19 жовтня 2023 р.: тези доповіді. – К., 2023. – С.134-137.

Розділ 1. Виявлення недозволених предметів на зображеннях

1.1. Аналіз актуальності теми

Інциденти, пов'язані з недозволеними предметами, викликають комплексні соціально-економічні та психологічні наслідки для суспільства. З фізіологічного погляду, прямий контакт або взаємодія з такими предметами може призвести до травматичних впливів на людське тіло, що, у критичних обставинах, може призвести до летального результату. Ці безпосередні ризики супроводжуються юридичними наслідками, які включають в себе штрафи, обмеження свободи та внесення до кримінального запису особи. В контексті соціальної структури, такі записи слугують стигматизації особи, ускладнюючи її соціальну інтеграцію, можливості працевлаштування та міжнародні поїздки.

Організації та державні установи можуть відчутти потребу перегляду норм безпеки, що може призвести до збільшення бюрократичних процедур і витрат. В узагальненому вигляді, інциденти з недозволеними предметами створюють мультидисциплінарний вплив на суспільство, який вимагає комплексного підходу до аналізу та вирішення.

Соціальна відповідальність підприємств перед спільнотою в контексті безпеки стосується зобов'язань компаній розглядати та мінімізувати можливі ризики, що можуть зашкодити фізичному благополуччю спільноти або її членів. Це охоплює стандарти безпеки продуктів, працевлаштування, здоров'я та гігієни на робочому місці, а також превентивні заходи проти аварій чи катастроф. Враховуючи ці ризики, підприємства мають взаємодіяти зі спільнотою, щоб забезпечити її захист, надаючи чітку інформацію про

діяльність компанії, рівні ризику та запобіжні заходи. Ігнорування цих відповідальностей може призвести до правових наслідків, збитків для репутації та фінансових втрат, тому для підприємств важливо активно взаємодіяти зі спільнотою, забезпечуючи її безпеку на всіх рівнях.

Інциденти на підприємствах є однією з найбільш актуальних тем в сфері безпеки праці, оскільки вони тісно пов'язані з життям і здоров'ям працівників.

У Сполучених Штатах Америки 2021 року було зафіксовано приблизно 2,8 млн інцидентів на робочому місці. З цієї кількості більше 5,000 випадків мали смертельний наслідок. Варто зазначити, що особливо небезпечною для працівників є будівельна галузь.

Щодо Європейського Союзу, тут щорічно відбувається близько 3,3 млн інцидентів на робочих місцях. Смертельні наслідки має приблизно 0,1% з них [1]. Зауважимо, що країни ЄС активно працюють над покращенням умов праці та безпеки своїх працівників.

Китай, як одна з найбільших промислових держав світу, має високий рівень ризику щодо робочих інцидентів. За даними 2021 року, кількість інцидентів тут досягла понад 6 млн [2, 3]. Особливо велика кількість інцидентів припадає на сферу виробництва.

Роблячи висновок, можна сказати, що інциденти на підприємствах є значущою проблемою на світовому рівні. Це вимагає від урядів, організацій та підприємств активних дій для покращення умов праці та забезпечення безпеки працівників.

На сучасних підприємствах, незалежно від їх розміру або сфери діяльності, інциденти трапляються регулярно. Розуміння того, які інциденти є найбільш поширеними, дозволяє компаніям підготовлюватися до потенційних загроз та відповідати на них більш ефективно.

Технічні несправності часто стають джерелом проблем для багатьох підприємств. Це може бути нефункціонування обладнання, відмови у

системах або збої в програмному забезпеченні. Такі проблеми можуть призвести до зупинки виробництва, втрати даних або інших втрат.

В сучасному цифровому контексті інформаційні інциденти, такі як кібератаки та витоки даних, стають все частішими, акцентуючи на необхідності надійного захисту даних

Отже, інциденти на підприємствах, пов'язані з наявністю недозволених предметів, становлять серйозний ризик для безпеки працівників і можуть мати далекосяжні наслідки для підприємства та суспільства в цілому. Ці інциденти можуть призвести до фізичних травм, смертельних наслідків та юридичних проблем.

1.2 Роль технологій виявленні недозволених предметів

Системи відеоспостереження виникли у 1940-х роках як прості чорно-білі камери для спостереження за державними об'єктами. З 1970-х, завдяки новим технологіям, з'явилися кольорові камери і можливість зберігання на магнітних стрічках. Цифрова ера 1980-1990-х принесла цифрові камери та жорсткі диски для зберігання, а також можливість передачі відео через Інтернет.

Останніми роками інтелектуальні алгоритми інтегрувались у системи відеоспостереження, дозволяючи автоматично виявляти рух, розпізнавати особи та аналізувати поведінку. З розвитком хмарних технологій відеодані стали зберігатися у хмарі, забезпечуючи доступ із будь-якої точки світу. Сучасні системи відеоспостереження інтегруються з іншими безпековими системами, використовуючи передові технології для забезпечення захисту.

Використання автоматизованих систем розпізнавання в сучасних технологічних процесах набуває все більшої актуальності, зміщуючи традиційні методи контролю. Існує ряд переваг цих систем, які демонструють їх ефективність і доцільність застосування:

- Перш за все, автоматизовані системи розпізнавання володіють великою швидкістю обробки даних. Там, де людина може витратити значну кількість часу на аналіз інформації та прийняття рішення, комп'ютерні системи роблять це миттєво. Така швидкість дозволяє оптимізувати робочі процеси, знижує ризик помилок, спричинених людським фактором, і підвищує продуктивність роботи.

- Автоматизовані системи здатні обробляти величезний обсяг інформації. Сучасний світ переповнений даними, і обробка таких масивів інформації вручну є майже неможливою. Автоматизовані системи розпізнавання можуть аналізувати тисячі і навіть мільйони записів, виділяючи ключову інформацію та забезпечуючи високий рівень точності.

- Точність і консистентність також є важливими перевагами. Людська робота завжди схильна до впливу емоцій, втоми або розсіяності. У випадку з автоматизованими системами, вони працюють на основі заданих алгоритмів, забезпечуючи однаковий рівень точності незалежно від обсягу даних або часу роботи.

- Крім того, автоматизовані системи розпізнавання часто здатні "навчатися". Застосування технологій штучного інтелекту та машинного навчання дозволяє системам адаптуватися до нових умов, вдосконалювати свої алгоритми та підвищувати ефективність роботи.

Однак, необхідно враховувати, що з впровадженням автоматизованих систем розпізнавання можуть виникнути і деякі виклики, такі як необхідність в удосконаленні та актуалізації алгоритмів, захисту даних та ін. Проте, з урахуванням всіх переваг, такі системи стають ключовими інструментами в багатьох галузях сучасної економіки.

Інтеграція систем автоматичного розпізнавання з існуючими системами безпеки підприємства є важливим аспектом в сучасних умовах. Ця інтеграція забезпечує підвищення ефективності заходів безпеки, розширення

можливостей аналітики та збільшення швидкості реакції на різноманітні загрози.

Одним із ключових аспектів такої інтеграції є взаємна сумісність компонентів систем. На практиці це означає, що системи автоматичного розпізнавання, такі як розпізнавання облич, голосу чи інших біометричних параметрів, повинні бути сумісними з архітектурою та протоколами комунікації існуючих систем безпеки. Для цього часто використовуються стандартні інтерфейси та API, які дозволяють легко інтегрувати різні системи і компоненти між собою.

Також важливо враховувати питання швидкодії та надійності. Системи автоматичного розпізнавання, особливо в режимі реального часу, вимагають великих обчислювальних потужностей. Тому при інтеграції з системами безпеки необхідно враховувати можливість масштабування, а також забезпечувати резервування та відмовостійкість систем.

Крім того, питання конфіденційності та захисту даних є особливо актуальними при роботі з біометричними даними. Ці дані відносяться до особистої інформації, і їх витік може призвести до серйозних наслідків. Тому системи автоматичного розпізнавання повинні мати високий рівень захисту, включаючи зашифрування даних, механізми аутентифікації та контролю доступу. Додатково, щоб забезпечити ефективну роботу системи, важливо забезпечити належну підготовку персоналу. Співробітники повинні знати принципи роботи систем автоматичного розпізнавання, а також основи їх інтеграції з існуючими системами безпеки.

З урахуванням зазначеного можна робити висновок, що інтеграція систем автоматичного розпізнавання з існуючими системами безпеки є не лише технічно складним завданням, але й важливим аспектом забезпечення конфіденційності, надійності та ефективності діяльності підприємства в сфері безпеки.

Виявлення недозволених предметів є важливим завданням в багатьох сферах, включаючи безпеку аеропортів, місць суспільного зібрання, кордонний контроль та інші. Сучасні технології надають значущий внесок у вирішення цієї проблеми на світовому рівні.

Один з найбільш поширених методів виявлення недозволених предметів – це рентгенівська інтроскопія, яка дозволяє контролювати вміст багажу та інших вантажів. Сучасні рентгенівські системи забезпечують високу роздільну здатність, що дозволяє виявляти найменші деталі та предмети.

Застосування нейронних мереж у цій сфері стає все більш поширеним. Ці мережі навчаються розпізнавати різноманітні об'єкти на базі великої кількості даних і можуть автоматично ідентифікувати недозволені предмети на сканованих зображеннях. Додаткова перевага нейронних мереж полягає в їх спроможності адаптуватися до нових типів загроз, які можуть з'явитися з часом.

Терагерцева імагінгова технологія є іншим важливим інструментом виявлення зброї, вибухових речовин та інших недозволених предметів під одягом людини. Терагерцеві хвилі можуть проникати через більшість неметалевих матеріалів, не завдаючи шкоди організму.

На додачу, системи відеоспостереження оснащені алгоритмами штучного інтелекту вже використовуються для аналізу поведінки осіб у публічних місцях, виявлення осіб, які можуть мати при собі недозволені предмети або здійснювати підозрілі дії.

Резюмуючи, важливо підкреслити вирішальну роль штучного інтелекту у виявленні недозволених предметів. Штучний інтелект, особливо в комбінації з нейронними мережами, представляє собою найперспективніший і найбільш ефективний підхід до розробки систем, які можуть адаптуватися до змінюваних умов та нових видів загроз. Він дозволяє автоматично аналізувати великі обсяги даних, вивчати патерни і тенденції та робити точні прогнози. Ці способи мають величезний потенціал для підвищення рівня безпеки в різних

секторах. Інтеграція систем на основі штучного інтелекту в комплексні системи безпеки не лише посилює їх ефективність, але й відкриває нові можливості для протидії потенційним загрозам. Тому можна стверджувати, що використання штучного інтелекту в сфері виявлення недозволених предметів є найбільш вдалим та обґрунтованим рішенням у сучасних умовах.

1.3 Категорія недозволених предметів

На сучасних підприємствах протоколи безпеки є невід'ємною частиною функціонування, маючи на меті забезпечити добробут та безпеку як працівників, так і матеріальних активів. Особлива увага приділяється обмеженню або забороні на введення певних предметів на територію підприємства, серед яких виділяються різноманітні види озброєння та приховані засоби самооборони.

Вогнепальна зброя, така як револьвери, пістолети, автомати та гвинтівки, займає важливе місце серед потенційних загроз. Хоча вони можуть бути законними в деяких регіонах або країнах, багато підприємств активно обмежують або повністю забороняють їх наявність у своїх приміщеннях. Також до списку потенційно небезпечних предметів можна віднести холодну зброю, яка включає в себе предмети, як ножі, шаблі, мечі та бойові кастети.

Крім традиційного озброєння, існують також електрошокові пристрої, які використовують електричний струм для паралізації чи поранення, а також газові балончики, які містять хімічний агент, що може тимчасово ослабити або паралізувати людину.

Окрему увагу заслуговують приховані або масковані засоби самооборони. Це можуть бути предмети, які на перший погляд виглядають абсолютно невинними, але при потребі можуть використовуватися як засоби самооборони або атаки. Наприклад, масковані ножі, які можуть бути вбудовані в ручки або гребінці; кастети, що маскуються під браслети;

електрошокові пристрої у вигляді кілець або брелоків; а також спреї самооборони, які маскуються під звичайні предмети типу парфумів.

Вибухові речовини існують у великому різноманітті, від традиційних форм до новітніх інновацій. Традиційно, вибухові речовини такі як тротил, динаміт, анфо та нітрогліцерин служили основою для виробництва бомб та інших вибухових пристроїв.

Щодо наркотичних речовин, світ знає велику кількість видів. Опіюди (наприклад, морфін, героїн), стимулятори (кокаїн, амфетамін), депресанти (барбітурати, бензодіазепіни) та психоделіки (LSD, псилоцибін) — це лише кілька прикладів основних класів наркотиків. Ці речовини взаємодіють з нервовою системою, викликаючи зміни в сприйнятті, настрої, свідомості та поведінці.

У зв'язку з вищевказаним, існує необхідність у постійному моніторингу та розумінні цих речовин та інструментів, які використовуються для їхнього вживання або активізації, з метою забезпечення громадської безпеки, особливо, особливо на підприємствах, на місцях громадського скупчення.

На основі аналізу вищезазначених даних стає очевидною актуальність проблеми контролю за рухом потенційно небезпечних та недозволених речовин та предметів. Задля вирішення проблеми незаконного введення холодної зброї на території підприємств, що може становити загрозу для безпеки, було вирішено створити застосунок для її виявлення на основі штучного інтелекту.

Необхідність регулювання володіння та використання холодною зброєю, як і будь-якою іншою зброєю, впливає з потреби забезпечення громадської безпеки, запобігання преступним діям та забезпечення порядку. Однак це також питання балансу прав громадян на самозахист та запобігання незаконному або небезпечному використанню зброї.

Враховуючи все вищезазначене, для підприємств критично важливо розробляти та дотримуватися чітких правил щодо недозволених предметів на

своїй території. Такі заходи сприяють забезпеченню безпеки, мінімізації ризиків та зниженню ймовірності небажаних інцидентів.

Резюмуючи, використання нейронних мереж для виявлення таких предметів надає можливість автоматизувати та значно підвищити ефективність процесу контролю. Сучасні технології, зокрема глибоке навчання, дозволяють розпізнавати складні об'єкти на зображеннях або в даних з великою точністю. Таким чином, прийняття рішення про написання роботи на цю тему є відгуком на актуальні потреби суспільства у застосуванні передових технологій для підвищення рівня безпеки на підприємствах та в інших сферах.

1.4 Технології аналізу та обробки зображень і методи розпізнавання об'єктів на зображенні

З початком вивчення візуального сприйняття людини відкрилися нові горизонти для розуміння того, як ми сприймаємо та обробляємо інформацію. Технології обробки зображень стали науковим відгалуженням, спрямованим на вивчення, вдосконалення та імітацію цього процесу за допомогою комп'ютерних систем.

Історично технології обробки зображень мають корені ще у період аналогової ери, коли спеціалісти маніпулювали фотографіями у фотолабораторіях. Однак з появою цифрової ери і розвитком комп'ютерних технологій обробка зображень отримала значущий поштовх до прогресу. Сучасні техніки, такі як ймовірнісні методи, машинне навчання та нейронні мережі, дозволили досягти небаченого раніше рівня обробки та аналізу зображень.

У сучасному світі значення цифрової обробки зображень неможливо переоцінити. Вона знаходить застосування в різних галузях – від медицини, де використовуються методи діагностики на основі зображень, до космічної промисловості, де потрібно аналізувати величезні об'єми космічних знімків.

Цифрова обробка зображень стає необхідною складовою розвитку технологій розширеної реальності, автономних транспортних засобів та безлічі інших інноваційних технологій.

Процес цифрової обробки зображень можна розбити на кілька ключових етапів (рис. 1.1).

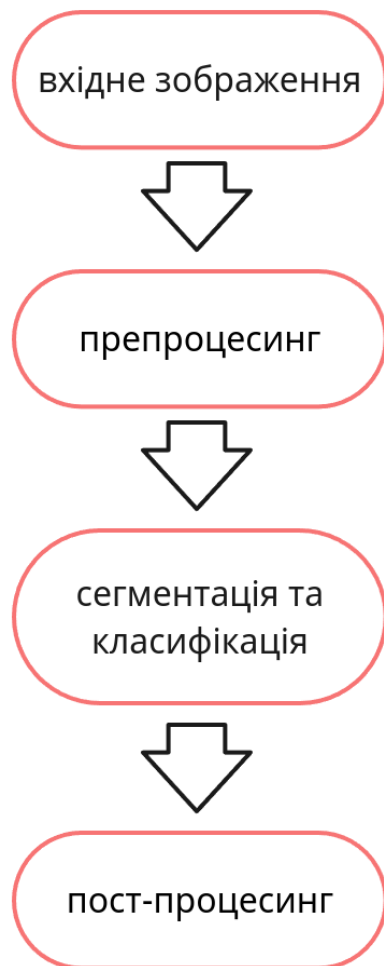


Рисунок 1.1 Процес цифрової обробки зображень

Перший етап - це завантаження вхідних даних. Другий етап – це препроцесинг, під час якого вихідне зображення піддається первинній обробці для покращення його якості. Це може включати в себе корекцію контрасту, шумоподавлення, згладжування або інтерполяцію.

Далі настає етап сегментації. Під час цього етапу зображення розбивається на сегменти або області, які мають певний смисл чи

відповідають певним об'єктам [4]. Ця процедура є основою для подальшого аналізу зображення та виділення необхідної інформації.

Заключний етап – пост-процесинг. На цьому етапі отримані результати можуть бути додатково оброблені для досягнення кінцевої мети. Це може включати в себе реконструкцію, морфологічну обробку або інтерпретацію отриманих даних.

Про кожен з цих етапів детальніше у другому розділі.

В усіх цих процесах важливу роль відіграють алгоритми та математичні моделі, які використовуються для аналізу та трансформації зображень. Ось деякі з них.

Алгоритми препроцесингу:

- Фільтр Гаусса є одним із популярних методів шумоподавлення на зображеннях. Він використовує гауссівську функцію для розмиття високочастотних деталей на зображенні, таких як шум чи гранули, підсилюючи зображення та зменшуючи його шум [6].

- Згладжування зображення, за допомогою фільтра Гаусса або інших методів, сприяє розмиттю границь об'єктів, що допомагає у покращенні однорідності зображення та зниженні впливу шуму.

- До інших методів препроцесингу можна віднести нормалізацію зображення, яка вирівнює яскравість та контраст, а також корекцію кольорової балансу для покращення якості зображення.

Алгоритми сегментації:

- Порогова обрізка (Thresholding) - це метод сегментації, де встановлюється поріг яскравості, і кожен піксель класифікується як об'єкт або фон, в залежності від того, чи перевищує він цей поріг.

- Сегментація за кольором дозволяє виділити області на зображенні, які мають схожий колір або відтінок, і розділити їх на

окремі сегменти, що допомагає в розпізнаванні об'єктів на основі їх кольору.

- Керована сегментація (Watershed Segmentation) використовує геометричні особливості зображення, такі як вершини гірського ландшафту, для виділення об'єктів та границь між ними, що особливо корисно для сегментації об'єктів зі складною геометрією.

Алгоритми пост-процесинга:

- Морфологічна обробка включає операції розширення (додавання пікселів) та стискання (видалення пікселів), що допомагають вдосконалити форми та структури об'єктів на зображенні, особливо в процесі видалення шуму та незначних дефектів.

- Реконструкція дозволяє відновити втрачені деталі на зображенні, використовуючи інформацію з попередніх етапів обробки, що може бути корисним для відновлення об'єктів зі складною структурою.

- Інтерпретація отриманих даних включає аналіз результатів сегментації та визначення атрибутів об'єктів, таких як їх розмір, форма та положення, що може бути корисним для подальшого використання, наприклад, в медичних дослідженнях чи робототехніці.

Також техніки аналізу зображень є важливою частиною обробки цифрових зображень і включають в себе ряд методів та інструментів для аналізу та трансформації візуальної інформації [5]:

- Гістограми зображень
- Фільтрація зображень
- Трансформації зображень

Гістограми зображень представляють собою графічне відображення розподілу інтенсивності пікселів на зображенні. Вони можуть бути

одновимірними (якщо розглядається тільки один канал кольору або яскравість) або багатовимірними, коли розглядаються кілька каналів кольору одночасно. Гістограми можуть надавати цінну інформацію про контраст, яскравість та розподіл кольорів на зображенні. На основі гістограми можна визначити потребу в корекції яскравості і контрасту зображення.

Фільтрація зображень - це процес модифікації зображення за допомогою спеціальних операцій, що змінюють значення пікселів на зображенні. Фільтри можуть бути лінійними та нелінійними. Лінійні фільтри використовують комбінацію пікселів в околі досліджуваного пікселя для отримання нового значення пікселя, тоді як нелінійні фільтри можуть базуватися на статистичних методах, таких як медіанна фільтрація. Фільтрація зображень може використовуватися для зменшення шуму, видалення артефактів, підсилення країв або інших деталей зображення.

Трансформації зображень - це операції, що змінюють положення та/або значення пікселів на зображенні [5]. Афінні трансформації включають масштабування, поворот та зсув і зберігають паралельність ліній. Проективні трансформації, відомі також як гомографії, використовуються для перетворення зображення з однієї площини на іншу, при цьому паралельні лінії можуть перетинатися. Морфологічні трансформації використовують структурний елемент для модифікації зображення, виконуючи операції розширення, стиснення, відкриття та закриття, що дозволяє виділяти, аналізувати та модифікувати структурні особливості зображення.

Узагальнюючи, техніка аналізу зображень надає набір інструментів для аналізу, модифікації та трансформації цифрових зображень, що дозволяє вдосконалити якість візуальної інформації, виявити ключові особливості зображення та адаптувати зображення для конкретних потреб.

Крім цього дуже важливим аспектом виявлення предметів на зображенні є методи розпізнавання об'єктів.

Методи розпізнавання об'єктів на зображенні поділяються на класичні та сучасні [3], засновані на глибокому навчанні (рис. 1.2). Давайте детальніше розглянемо ці методи.

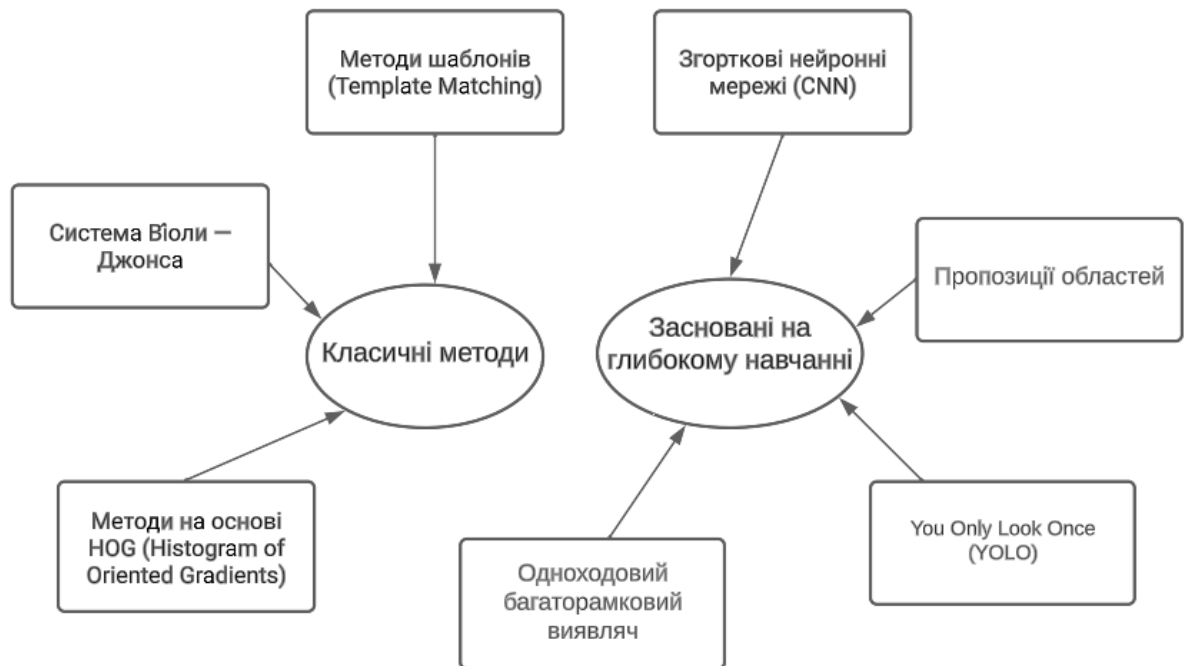


Рисунок 1.2 Класичні та сучасні методи розпізнавання об'єктів

Класичні методи розпізнавання об'єктів зазвичай використовують різноманітні алгоритми та техніки для обробки зображень, щоб визначити та класифікувати об'єкти. Основні принципи цих методів полягають у використанні характеристик зображення, таких як текстури, кольори, форми та контури. Ці методи спрямовані на пошук локальних особливостей на зображенні, які є стійкими до масштабних змін, обертання та інших трансформацій.

Метод гістограми направлених градієнтів (HOG) є важливим інструментом у сфері комп'ютерного зору. Він базується на аналізі градієнтів яскравості зображення, де кожна частина зображення розділяється на невеликі області, звані "клітинами", в яких обчислюються градієнти та орієнтації країв. Ці градієнти та орієнтації потім зводяться в гістограми, які утворюють основу

опису форми та структури об'єктів у межах зображення. Ефективність НОГ полягає у його здатності захоплювати інформацію про форму та текстуру об'єктів, роблячи його корисним у різноманітних застосуваннях, від розпізнавання облич до моніторингу пішоходів.

Система Віоли-Джонса, розроблена Полом Віолою та Майклом Джонсом, є однією з перших ефективних методик для розпізнавання об'єктів у реальному часі, зокрема облич у зображеннях. Ця система використовує так звані "прикметні ознаки Хаара", які швидко обчислюються завдяки інтегральним зображенням, що дозволяє ефективно виявляти структурні властивості об'єктів. Вона застосовує каскадний класифікатор, де кожен наступний етап рішення є більш складним, що допомагає швидко відкидати нерелевантні області зображення. Незважаючи на те, що цей метод був розроблений у 2001 році, він продовжує залишатися значущим у дослідженнях комп'ютерного зору та покладає основу для більш сучасних методів розпізнавання об'єктів.

Метод шаблонів (Template Matching) є одним із класичних підходів у комп'ютерному зорі, який застосовується для визначення місцезнаходження об'єктів на зображенні. Цей метод полягає у порівнянні фрагменту зображення (шаблону) з цільовим зображенням для ідентифікації відповідностей за допомогою кореляції. Ефективність методу шаблонів залежить від різниці між шаблоном та цільовим зображенням, а також від умов освітлення та змін у перспективі. Хоча цей метод є простим у реалізації та інтуїтивно зрозумілим, він часто обмежений у випадках, коли об'єкт зазнає значних змін у масштабі, орієнтації чи освітленні.

Методи, засновані на глибокому навчанні, в основному використовують нейронні мережі з глибокими архітектурами для автоматичної здатності витягувати особливості з зображення та виконувати класифікацію.

Метод пропозицій областей, відомий як Region Proposal Network (RPN), є ключовою складовою багатьох архітектур нейронних мереж, у галузі

розпізнавання об'єктів на зображеннях. Він автоматично визначає потенційні області на зображенні, які можуть містити об'єкти, значно покращуючи швидкість та точність в порівнянні з попередніми методами. RPN одночасно вирішує дві задачі: класифікацію областей як містять або не містять об'єкти та регресію координат обмежувальних прямокутників. Це дозволяє ефективно інтегрувати виявлення об'єктів та класифікацію в єдиний, оптимізований процес.

Одноходовий багаторамковий виявляч (Single Shot MultiBox Detector, SSD) є передовим методом у сфері розпізнавання об'єктів, заснованим на глибокому навчанні. Цей алгоритм вирізняється своєю здатністю виявляти об'єкти на різних масштабах, використовуючи одночасну екстракцію ознак на кількох рівнях. SSD об'єднує переваги точності та швидкості, здатний працювати в реальному часі з високою ефективністю. Його архітектура дозволяє ефективно обробляти великі масиви даних, що робить його ідеальним для складних задач розпізнавання об'єктів у різноманітних додатках.

Згорткові нейронні мережі (CNN) є важливим класом глибоких нейронних мереж, які широко використовуються в області комп'ютерного зору. Ці мережі відрізняються своєю здатністю автоматично та ефективно виявляти ключові характеристики візуальних даних завдяки використанню згорткових шарів. CNN структуровані таким чином, що вони імітують спосіб, яким людське зорове сприйняття збирає та обробляє інформацію, прогресуючи від низькорівневих ознак до вищих абстракцій. Ці мережі складаються з різних шарів, включаючи згорткові шари, пулінгові шари та повнозв'язні шари, кожен з яких виконує унікальні функції під час процесу навчання. CNN широко використовуються в багатьох застосуваннях, від автоматичного розпізнавання об'єктів до аналізу медичних зображень, завдяки своїй здатності до глибокого та ефективного аналізу візуальної інформації.

Метод "You Only Look Once" (YOLO) є революційним підходом в галузі розпізнавання об'єктів, який заснований на принципі швидкої та інтегрованої обробки зображень. Відмінною рисою YOLO є його спроможність аналізувати весь образ за один прохід, що різко підвищує швидкість виявлення об'єктів порівняно з традиційними методами, які аналізують окремі частини зображення. Цей метод ефективно використовує конволюційні нейронні мережі для одночасного прогнозування кількох класів і розташування об'єктів, мінімізуючи потребу в додаткових кроках обробки. Однак, незважаючи на високу швидкість, YOLO може мати певні обмеження у точності виявлення дрібних або перекритих об'єктів.

Перевагами глибокого навчання є висока точність розпізнавання, автоматичний вибір особливостей та здатність обробляти великі набори даних. Однак вони вимагають великої кількості даних для навчання та значних обчислювальних ресурсів.

Відмінності між класичними методами та глибоким навчанням полягають в тому, що перші зазвичай вимагають ручного вибору особливостей та знань про домен, у той час як методи на основі глибокого навчання можуть автоматично вивчати особливості без значного втручання людини.

Однак, варто зазначити, що обидва підходи мають свої спільні риси. Обидва прагнуть розпізнати та класифікувати об'єкти на зображенні, використовуючи характеристики зображення. Також обидва підходи можуть бути застосовані в комбінованій манері, де класичні методи використовуються для попередньої обробки або підсилення результатів глибокого навчання.

Підсумовуючи, можна сказати, що розпізнавання об'єктів на зображенні є динамічною областю досліджень, яка постійно розвивається завдяки комбінації класичних методів та сучасних технік глибокого навчання. Детальніше про ці методи та інші методи у другій частині кваліфікаційної роботи.

1.5 Висновки до першого розділу

В рамках першого розділу була розглянута проблема виявлення недозволених предметів на підприємствах і актуальність її вирішення за допомогою сучасних технологій. Проблема виявлення недозволених предметів, особливо на великих підприємствах та в інфраструктурних об'єктах, відіграє критичну роль в забезпеченні безпеки. Крім прямої небезпеки, яку можуть представляти такі предмети, є також питання про репутаційний ризик для організацій, що не можуть гарантувати безпеку своїх працівників та відвідувачів.

Сучасні технології обробки та аналізу зображень відкривають нові можливості для вирішення цієї проблеми. Техніки, такі як гістограми, морфологічні операції та методи фільтрації, дозволяють ефективно обробляти зображення, виділяючи на них ключові особливості, необхідні для подальшого аналізу.

Зокрема, методи розпізнавання об'єктів, базовані на глибокому навчанні, демонструють вражаючі результати в задачах розпізнавання об'єктів на зображенні, що робить їх вельми перспективними для виявлення недозволених предметів.

Розглядаючи висвітлену інформацію, можна констатувати, що інтеграція сучасних технологій в системи безпеки підприємств не тільки підвищить ефективність виявлення недозволених предметів, але й може слугувати превентивною мірою, відсікаючи спроби порушення безпеки.

Розділ 2. Аналіз методів розпізнавання об'єктів на зображенні

2.1. Аналіз методів аналізу зображень

Огляд методів аналізу зображень охоплює широкий спектр технік та

підходів, які є фундаментальними для розуміння та обробки візуальної інформації в цифровому вигляді.

Основні принципи та концепції обробки зображень стартують з цифрового представлення зображень. Ключовим елементом є піксель – найменша одиниця зображення, яка має конкретний колір. Кольори у цифрових зображеннях представлені через різні формати, як-от RGB чи CMYK, де кожен колір визначається своїм унікальним набором значень. Основні операції обробки зображень включають регулювання контрасту, яскравості та застосування різноманітних фільтрів. Методи покращення зображення, такі як шумоподавлення та збільшення різкості, дозволяють підвищити якість зображень, що особливо важливо в медичній візуалізації та супутниковій зйомці.

Класичні методи обробки зображень засновані на алгоритмах, які застосовують елементарні трансформації, такі як масштабування та обертання. Детектування країв – ключова процедура в аналізі зображень – включає в себе алгоритми, такі як Canny, Sobel та Prewitt. Морфологічні операції, як-от ерозія та дилатація, змінюють форму та структуру об'єктів у зображенні, що застосовується в області біомедичного зображення та промислового машинного зору.

Методи машинного навчання в аналізі зображень привносять новий рівень можливостей у цю область. Навчання з підкріпленням використовується для розв'язання завдань, де потрібно визначити найкращу стратегію дій на основі зображень. Згорткові нейронні мережі (CNN) є фундаментом сучасного комп'ютерного зору, особливо ефективні у класифікації та розпізнаванні об'єктів у зображеннях. Автоенкодери, з іншого боку, дозволяють відновлювати та класифікувати зображення, виходячи з неповної або пошкодженої інформації, застосовуючи алгоритми глибинного навчання для визначення ключових характеристик у вхідних даних.

Додатковим розвитком у сфері аналізу зображень є використання генеративно-змагальних мереж (GANs). Ці мережі здатні генерувати нові

зображення, які важко відрізнити від справжніх, знаходячи застосування в областях, таких як створення фотореалістичних зображень та глибоке фальсифікування. Крім того, техніки переносу стилю, які дозволяють модифікувати візуальний стиль зображення, використовуючи шаблони з інших зображень, відкривають нові можливості для творчості та дизайну.

У галузі обробки зображень також набувають популярності методи розширеного та доповненого реальності. Ці технології дозволяють інтегрувати віртуальні об'єкти у реальний світ, забезпечуючи взаємодію між віртуальними та фізичними об'єктами. Такий підхід має значний потенціал в освіті, розвагах та промисловості.

Заключно, у сфері аналізу зображень також відбуваються значні досягнення в області інтерпретації та аналізу емоцій. Системи, що здатні аналізувати вирази обличчя та мову тіла, знаходять застосування в психології, маркетингу та системах безпеки.

Цей огляд методів аналізу зображень демонструє, що ця галузь постійно розвивається, інтегруючи новітні технології та алгоритми, що розширюють можливості візуального аналізу та обробки. Завдяки цьому процесу, комп'ютерний зір та обробка зображень стають все більш потужними та гнучкими інструментами у різноманітних додатках, від автоматичного виявлення медичних патологій до створення високоякісних візуальних ефектів у кіноіндустрії.

Розвиток цифрової обробки зображень тісно пов'язаний з прогресом у галузі обчислювальної техніки та штучного інтелекту. Поліпшення обчислювальної потужності та алгоритмічної ефективності відкриває нові горизонти для обробки великих наборів даних, дозволяючи швидше та точніше аналізувати зображення. Ключову роль у цьому процесі відіграють алгоритми машинного навчання та глибокого навчання, які здатні автоматично виявляти складні закономірності та характеристики в даних, значно вдосконалюючи якість та точність обробки зображень.

Інноваційні алгоритми, такі як ті, що застосовуються в генеративно-змагальних мережах (GANs) та згорткових нейронних мережах (CNN), демонструють здатність моделювати складні візуальні структури та контексти, що використовуються в різноманітних областях, від автоматичного створення змісту до розробки нових лікарських препаратів. Особливо значущими є досягнення у сфері автоматичного розпізнавання та аналізу обличчя, які знаходять застосування в безпекових системах, рекламі та інших областях.

Також важливим аспектом є застосування алгоритмів обробки зображень в медичній візуалізації. Використання розширених методів обробки зображень дозволяє значно підвищити точність діагностики та лікування. Зокрема, методи комп'ютерного зору активно використовуються для аналізу медичних знімків, таких як МРТ, КТ та рентгенівські зображення, що дозволяє виявляти патологічні зміни на ранніх стадіях. Впровадження глибинного навчання та інших технік машинного навчання в медичну візуалізацію забезпечує можливість більш точного аналізу структур і функцій організму, що важливо для планування лікування та відстеження динаміки захворювань.

На додаток до медичної візуалізації, методи обробки зображень знаходять застосування в галузі біометричної ідентифікації. Системи розпізнавання обличчя, відбитків пальців та сітківки ока дозволяють забезпечувати безпеку та ідентифікацію особистості у багатьох сферах, включаючи фінансові послуги, безпеку аеропортів та мобільні пристрої. Розвиток цих технологій відбувається завдяки поєднанню класичних методів обробки зображень з сучасними алгоритмами машинного навчання, які здатні виявляти та аналізувати складні візуальні шаблони.

Ще декілька важливих застосувань технологій обробки зображень:

- Автоматизовані системи спостереження та контролю. Використання алгоритмів обробки зображень для аналізу відеопотоків з камер спостереження дозволяє виявляти підозрілі дії, управляти трафіком та забезпечувати загальну безпеку. Ці системи використовують складні алгоритми для обробки великих

обсягів даних у реальному часі, виявляючи важливі події та забезпечуючи швидке реагування.

- У сфері розваг та медіа, техніки обробки зображень використовуються для створення спецефектів у кіно та телебаченні. Завдяки прогресу в області комп'ютерної графіки та візуальних ефектів, реалізуються складні візуальні концепції, що значно підвищує якість візуального контенту. Особливо помітними є досягнення у створенні фотореалістичних 3D-моделей, анімації та віртуальної реальності, що відкриває нові перспективи для кіноіндустрії та відеоігор.

- У сфері освіти та наукових досліджень, методи обробки зображень використовуються для візуалізації складних наукових концептів та даних. Від візуалізації астрономічних даних до мікроскопічних зображень у біології, ці техніки дозволяють науковцям та студентам глибше зрозуміти складні процеси та структури.

- Важливою тенденцією в сучасній обробці зображень є інтеграція з Інтернетом речей (IoT). Це забезпечує можливість аналізувати візуальні дані з множини пристроїв, що взаємодіють між собою. Така інтеграція відкриває шлях для розумних домів, промислових автоматизованих систем та управління міським середовищем.

Зрештою, важливо зазначити, що обробка зображень невід'ємно пов'язана з питаннями етики та приватності. Використання технологій розпізнавання облич та аналізу візуальних даних порушує питання конфіденційності та захисту персональних даних. Тому розвиток цієї галузі вимагає не тільки технічного прогресу, але й розвитку нормативно-правових рамок, що регулюють використання та захист інформації.

В цілому, методи аналізу зображень представляють собою динамічну та багатогранну область, що впливає на багато аспектів сучасного життя. Від медичної діагностики до кіноіндустрії, від наукових досліджень до розваг, ці технології продовжують відкривати нові можливості для інновацій та розвитку.

Більш детальний опис цих методів в наступних розділах.

2.2. Дослідження попередньої обробки зображень

В області обробки зображень, попередня обробка відіграє ключову роль у підготовці зображень до подальшого аналізу та інтерпретації. Однією з фундаментальних технік є фільтрація зображень, яка включає в себе різні підходи, такі як лінійна та нелінійна фільтрація.

Дослідження попередньої обробки зображень включає в себе кілька ключових етапів. Ось основні кроки та їх детальний опис:

1. Фільтрація зображень:

1.1. Лінійна фільтрація: Використовує лінійні фільтри для обробки пікселів зображення, засновані на лінійній комбінації їх значень у визначеному околі. Наприклад, фільтр середнього значення використовує середнє значення пікселів, а медіанний фільтр вибирає медіанне значення для зменшення шуму.

1.2. Нелінійна фільтрація: Включає складніші методи, такі як білатеральний фільтр, який зменшує шум, зберігаючи при цьому гостроту країв, та Гаусівський фільтр, який використовує нормальний розподіл для згладжування зображень.

2. Нормалізація зображень:

Процес коригування інтенсивностей пікселів для покращення контрасту зображень. Лінійна нормалізація змінює масштаб інтенсивності пікселів, тоді як гістограмна еквалізація розподіляє частоту використання інтенсивностей більш рівномірно.

3. Сегментація зображень:

Розділення зображення на декілька частин або сегментів зі спільними характеристиками. Порогова сегментація використовується для розділення пікселів за значеннями інтенсивності, а метод вододілу визначає границі між різними областями зображення.

4. Морфологічна обробка:

Включає операції, такі як ерозія та дилатація, які допомагають покращити структуру та форму об'єктів на зображенні.

5. Корекція спотворень та виправлення перспективи:

Техніки, що коригують геометричні спотворення, викликані лінзами камери або кутом зйомки, для створення більш реалістичного зображення об'єктів.

6. Видалення артефактів:

Включає методи для видалення шуму, наприклад, на основі хвильового перетворення, щоб покращити якість зображення, мінімізуючи втрату важливої інформації.

Лінійна фільтрація, як правило, включає в себе фільтри, які обробляють пікселі зображення на основі лінійної комбінації їхніх значень у визначеному околі. Наприклад, фільтр середнього значення використовує середнє арифметичне значення пікселів у визначеному околі. Для пікселя $P(x,y)$, середнє значення обчислюється як

$$M(x, y) = \frac{1}{N} \sum_{i=-k}^k \sum_{j=-k}^k I(x + i, y + j)$$

де $I(x,y)$ - інтенсивність пікселя на позиції (x,y) , N - кількість пікселів у околі (наприклад, для 3×3 околу $N=9$), k - радіус околу (для 3×3 $k=1$).

Тоді як медіанний фільтр вибирає медіанне значення пікселів для зменшення шуму. Процедура полягає в упорядкуванні значень інтенсивності пікселів у околі пікселя $P(x,y)$ і виборі середнього значення (медіани) з цього списку.

Нелінійна фільтрація, у свою чергу, включає більш складні методи, такі як білатеральний фільтр та фільтр Гаусса [6].

$$BF[I]_p = \frac{1}{W_p} \sum_{q \in S} I_q \cdot f_r(\|I_p - I_q\|) \cdot g_s(\|p - q\|)$$

де I - вхідне зображення, p та q - пікселі зображення, I_p та I_q - інтенсивності відповідних пікселів, S - підмножина зображення, f_r - фільтр діапазону (зазвичай гауссівський), що залежить від різниці інтенсивностей, g_s

- просторовий гауссівський фільтр, що залежить від відстані між пікселями, та W_p нормалізаційний коефіцієнт.

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

де G - гауссівський фільтр, x та y - координати пікселя відносно центру фільтра, а σ - стандартне відхилення гауссівського розподілу, яке контролює рівень згладжування.

Білатеральний фільтр ефективно зменшує шум, зберігаючи при цьому гостроту країв, тоді як Гауссівський фільтр використовує нормальний розподіл для згладжування зображень.

Нормалізація зображень є ще одним важливим аспектом попередньої обробки. Це процес коригування інтенсивностей пікселів для покращення контрасту зображень. Лінійна нормалізація зображення змінює масштаб інтенсивності пікселів, тоді як гістограмна еквалізація розподіляє частоту використання інтенсивностей пікселів більш рівномірно по всьому зображенню.

Формула лінійної нормалізації:

$$I_{\text{norm}} = \frac{I - I_{\min}}{I_{\max} - I_{\min}} \times (L_{\max} - L_{\min}) + L_{\min}$$

де:

- I - оригінальна інтенсивність пікселя.

- $\min I_{\min}$, $\max I_{\max}$ - мінімальне та максимальне значення інтенсивностей в зображенні.

- $\min L_{\min}$, $\max L_{\max}$ - новий мінімальний та максимальний рівень інтенсивності (наприклад, 0 та 255 для 8-бітного зображення).

- $\text{norm}I_{\text{norm}}$ - нормалізована інтенсивність пікселя.

Формула гістограмної еквалізації:

$$I_{\text{eq}} = \left\lfloor \frac{C(I)}{N} \times (L_{\max} - 1) \right\rfloor$$

де:

- $C(I)$ - кумулятивна сума гістограми інтенсивностей до рівня I .
- N - загальна кількість пікселів у зображенні.
- $\max L_{\max}$ - максимальна інтенсивність (наприклад, 256 для 8-бітного зображення).

- I_{eq} - інтенсивність пікселя після еквалізації.

Сегментація зображень, полягає у розділенні зображення на декілька частин або сегментів, що мають спільні характеристики. Порогова сегментація використовується для розділення пікселів на основі їхніх значень інтенсивності. Метод вододілу, з іншого боку, імітує процес розливу води для визначення границь між різними областями зображення.

Алгоритм порогової сегментації:

Відповідно до порогової сегментації, для заданого пікселя на координатах (x,y) із інтенсивністю $I(x,y)$, вихідний сигнал сегментації $S(x,y)$ визначається за наступною умовою:

- $S(x,y)=1$, якщо $I(x,y)$ більше ніж порогове значення T .
- $S(x,y)=0$, у всіх інших випадках.

Це означає, що піксель вважається частиною сегменту (і відповідно має значення 1), якщо його інтенсивність перевищує заданий поріг T . В іншому випадку, піксель не включається до сегменту (і має значення 0).

Кожен з цих методів попередньої обробки зображень має свої унікальні застосування та виклики, і правильний вибір техніки залежить від специфіки задачі обробки зображень. Важливо розуміти, що вибір методу обробки має адаптуватися до характеристик вхідного зображення та очікуваного результату.

На додаток до зазначених методів, розглянемо морфологічну обробку, що включає операції, такі як ерозія (рис. 2.1) та дилатація, що допомагають покращити структуру та форму об'єктів на зображенні. Ерозія зменшує розміри об'єктів та видаляє дрібні шуми. Це досягається шляхом "еродування" країв об'єктів. Дилатація збільшує розміри об'єктів та заповнює дірки та розриви у контурах об'єктів.

Початкове зображення



Зображення після використання ерозії



Рисунок

2.1

Приклад

роботи

ерозії

Початкове зображення



Зображення після процесу дилатації



Рисунок

2.2

Приклад

роботи

дилатації

Алгоритм ерозії:

$$(A \ominus B)(x,y) = \min_{(bx,by) \in B} [A(x-bx, y-by)],$$

де:

- A - оригінальне зображення.
- B - структурний елемент (наприклад, диск або квадрат), який визначає форму ерозії.

- $A \ominus B$ - результат ерозії зображення A зі структурним елементом B .

Алгоритм дилатації:

$$(A \oplus B)(x,y) = \max_{(bx,by) \in B} [A(x+bx, y+by)],$$

де:

- A - оригінальне зображення.
- B - структурний елемент.
- $A \oplus B$ - результат дилатації зображення A зі структурним елементом B .

Ще одним важливим аспектом є корекція спотворень та виправлення перспективи. Це включає техніки, що коригують геометричні спотворення, викликані лінзами камери або кутом зйомки, для створення більш реалістичного зображення об'єктів. Одним з підходів для корекції цих спотворень є використання перетворення гомографії.

Гомографія - це тип перетворення, яке використовується для виправлення перспективних спотворень. Воно дозволяє перетворювати зображення так, ніби воно було знято з іншого кута.

Формула перетворення гомографії:

$$p' = H \cdot p,$$

де:

- p - вектор пікселя у вихідному зображенні у формі $[x, y, 1]^T$.
- p' - вектор пікселя у виправленому зображенні.
- H - матриця перетворення гомографії розміром 3×3 .

Матриця H зазвичай обчислюється на основі набору відповідностей між пікселями вихідного та виправленого зображень. Це може включати використання характерних точок та їх відповідностей у двох зображеннях.

Крім того, важливим є видалення артефактів, таких як шум від сенсора камери або стиснення зображень. Методи, такі як видалення шуму на основі хвильового перетворення, використовуються для покращення якості зображення, мінімізуючи втрату важливої інформації.

Хвильове перетворення - це метод, який розкладає зображення на компоненти, які представляють різні рівні деталізації. Потім шум може бути видалений шляхом модифікації або видалення цих компонентів.

Формула хвильового перетворення:

$$W(f)(a,b) = |a|^{-1} \int_{-\infty}^{\infty} f(t) \cdot \psi(at-b) dt,$$

де:

- $f(t)$ - функція зображення в часовому домені.
- $\psi(t)$ - хвильова функція (wavelet function), яка діє як базис для перетворення.
- a, b - параметри масштабування і зсуву в хвильовому перетворенні.
- $W(f)(a,b)$ - хвильовий коефіцієнт на певному масштабі a та зсуві b .

Процес видалення шуму включає застосування хвильового перетворення до зображення, модифікацію або видалення коефіцієнтів, які відповідають шуму, та зворотне перетворення для отримання виправленого зображення.

У сфері машинного зору та штучного інтелекту, попередня обробка зображень є критично важливою для забезпечення точності та надійності подальшого аналізу. Наприклад, у галузі медичної діагностики, точна сегментація та нормалізація можуть допомогти у визначенні патологій на медичних знімках. Точно так само, в автомобільній індустрії, ефективна обробка зображень відіграє ключову роль у розвитку систем автономного водіння.

Узагалі, успіх обробки зображень залежить від глибокого розуміння взаємодії між різними техніками та їх адаптації до конкретних вимог проекту.

Враховуючи швидкий розвиток технологій цифрового зображення, ця область продовжує розвиватися, пропонуючи все більш передові методи для обробки та аналізу зображень.

2.3. Аналіз нейронних мереж в розпізнаванні об'єктів

Нейронні мережі відіграють ключову роль у сучасних системах розпізнавання об'єктів, що є одним із найбільш динамічно розвиваючихся напрямків у галузі штучного інтелекту. Основою цього прогресу є згорткові та рекурентні нейронні мережі, які відтворюють деякі аспекти людського сприйняття та обробки інформації.

Згорткові нейронні мережі (CNN) становлять основу багатьох систем розпізнавання об'єктів. Їх архітектура імітує спосіб, яким людський зір виявляє форми та об'єкти, шляхом послідовного застосування фільтрів, що виявляють краї, кути, та інші особливості об'єктів. Ці мережі здатні виявляти взаємозв'язки у великих масивах даних, що робить їх ідеальними для обробки зображень.

Вибір оптимальної архітектури мережі є ключовим, оскільки він визначає базову здатність моделі до сприйняття та класифікації образів. Це може включати в себе використання глибоких нейронних мереж, таких як згорткові нейронні мережі (CNN), для ефективного аналізу зображень. Тюнінг гіперпараметрів моделі, таких як швидкість навчання, кількість шарів, розмір фільтрів тощо, є критично важливим для підвищення точності та ефективності моделі. На рис.2.3 показано схему базової CNN, зв'язки якої залежать від конкретної реалізації.

Архітектура стандартної CNN

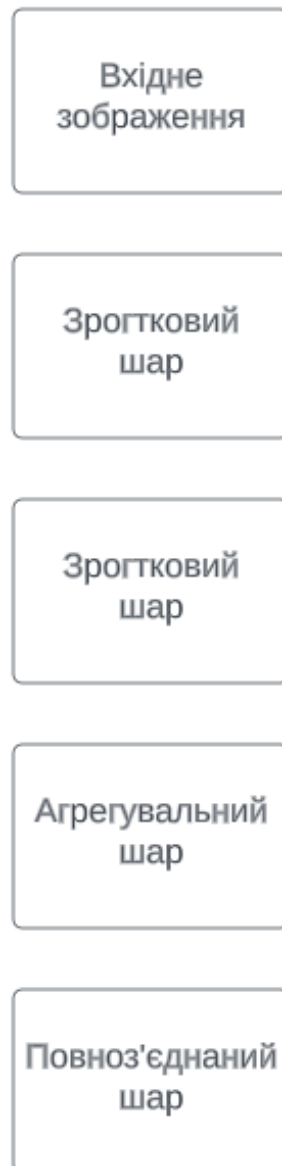


Рисунок 2.3 Схеми архітектури базової згорткової нейронної мережі.

Згортковий шар представляє собою фундаментальний елемент у архітектурах згорткових нейронних мереж (CNN). Цей шар обробляє вхідні дані за допомогою набору навчальних фільтрів, ефективно вловлюючи специфічні характеристики, такі як краї, текстури та візерунки.

Опис дії та композиція згорткового шару:

1. Фільтри (ядра згортки): У згортковому шарі використовуються декілька фільтрів. Кожен фільтр є невеликою матрицею ваг.

2. Згортка: Кожен фільтр застосовується до вхідних даних, зображення, шляхом проведення операції згортки. Це означає, що фільтр "переміщується" по всьому зображенню, і для кожної позиції обчислюється скалярний добуток між елементами фільтра та відповідними елементами вхідних даних.

3. Активаційна функція: Після обчислення згортки до результату часто застосовується активаційна функція, така як ReLU (Rectified Linear Unit). Це додає нелінійність до обробки даних, що є важливим для глибокого навчання.

4. Згорткова карта (feature map): Вихід згорткового шару є згортковою картою, яка представляє виявлені характеристики (наприклад, краї, текстури). Кожен фільтр генерує свою власну згорткову карту.

Формула для згортки можна представити у такому вигляді [7]:

$$O_{ij} = \sum_m \sum_n I(i+m)(j+n) \times K_{mn}$$

де:

- O_{ij} — це елемент у згортковій карті.
- I — вхідні дані (наприклад, зображення).
- K — фільтр (ядро згортки).
- m, n — індекси, що пробігають по фільтру.

ReLU (Rectified Linear Unit) представляє собою ключовий компонент в даній архітектурі глибокого навчання. Ця функція активації, визначена як

$$f(x) = \max(0, x)$$

де x є вхідним сигналом нейрону [10]. Ця функція пропускає всі позитивні значення без змін, а всі негативні значення замінює на нуль. Така проста характеристика дозволяє уникнути проблеми зникнення градієнтів, яка є значною проблемою для традиційних сигмоїдних або гіперболічних тангенс активаційних функцій, особливо в глибоких нейронних мережах. Крім цього у ReLU є проблема "Вмираючих ReLU": Якщо вхідний сигнал є негативним, вихід

стає постійно нульовим, що може призвести до того, що деякі нейрони більше не активуються (вмирають) під час навчання.

Наступний шар - агрегувальний. Агрегувальний шар відіграє ключову роль у зменшенні просторових розмірів функціональних карт. Цей шар вибірково сканує вхідні функціональні карти, залишаючи лише максимальні значення у визначеному вікні (рис.2.4). Ця операція не тільки зменшує обсяг обчислень та кількість параметрів моделі, але й забезпечує певний рівень інваріантності до місцезнаходження об'єктів у зображенні.



Рисунок 2.4 Приклад принципу дії агрегувального шару

Опис дії та композиція агрегувального шару:

1. Розмір вікна: Вибирається розмір вікна (наприклад, 2x2 або 3x3 пікселів), який буде використовуватися для сканування вхідної функціональної карти.

Крок (stride): Визначається крок, з яким вікно буде переміщуватися по карті.

2. Крок часто дорівнює розміру вікна, що призводить до неперекривання вікон.

3. Операція максимального згортання: Для кожного вікна вхідної карти вибирається максимальне значення. Це максимальне значення представляє весь регіон вікна в вихідній згортковій карті.

Формула агрегувального шару [12]:

$$O_{ij} = \max(I_{kl})$$

де:

- O_{ij} — елемент у вихідній агрегованій карті.
- I_{kl} — елементи вхідної карти в межах поточного вікна.
- k, l — індекси, що пробігають по вікну.

Цей процес повторюється для кожного вікна, що сканує вхідну карту. Результатом є зменшена за розміром карта характеристик, де кожен елемент відображає максимальне значення відповідного регіону вхідної карти. Ця операція допомагає зменшити обсяг обчислень, кількість параметрів моделі та забезпечує певний рівень інваріантності до місцезнаходження об'єктів у зображенні.

Повнозв'язні шари (fully connected layers), часто згадувані як FC шари, є стандартним компонентом у багатьох типах нейронних мереж, включаючи згорткові нейронні мережі (CNN). Ці шари зазвичай розміщуються після згорткових і агрегувальних шарів у CNN і використовуються для класифікації або інших завдань, що вимагають виведення високорівневих висновків з низькорівневих характеристик, отриманих з попередніх шарів [13].

Опис дії та композиція повнозв'язного шару:

1. Векторизація вхідних даних: Вхідні дані (зазвичай вихід з попереднього шару) перетворюються у одновимірний вектор. У випадку з CNN, це означає перетворення вихідних активаційних карт у лінійний вектор.

2. Зв'язок з нейронами: Кожен елемент вектора вхідних даних повністю зв'язується з кожним нейроном у FC шарі. Це означає, що кожен вхідний елемент має ваговий коефіцієнт для кожного нейрону.

3. Обчислення виходів: Вихід кожного нейрону обчислюється як взважена сума вхідних даних, до якої додається значення зсуву (bias).

Формула для повнозв'язного шару може бути представлена так [10]:

$$O_i = \text{активація} \left(\sum_j W_{ij} X_j + b_i \right)$$

де:

- O_i — вихід i -го нейрону.
- W_{ij} — вага, що з'єднує j -й вхідний елемент з i -м нейроном.
- X_j — j -й елемент вхідного вектора.
- b_i — зсув (bias) для i -го нейрону.
- $\text{активація}()$ — активаційна функція, така як ReLU, сигмоїд, або тангенс гіперболічний.

Повнозв'язні шари дозволяють нейронній мережі вивчати високорівневі абстракції та зв'язки у вхідних даних, що є критично важливим для завдань класифікації та регресії.

Рекурентні нейронні мережі (RNN), з іншого боку, особливо ефективні у випадках, де дані мають послідовну структуру, як-от в мовному розпізнаванні або прогнозуванні часових рядів. Вони здатні зберігати інформацію про попередні входи, що дає їм можливість враховувати контекст при обробці нових даних.

У контексті розпізнавання об'єктів, RNN можуть бути використані для аналізу послідовностей зображень або відео. Наприклад, вони можуть допомагати у розпізнаванні об'єктів у відео, де контекст попередніх кадрів може дати додаткову інформацію для точнішого розпізнавання об'єктів у поточному кадрі.

Тим не менш, для безпосереднього розпізнавання об'єктів на статичних зображеннях частіше використовуються згорткові нейронні мережі (CNN), оскільки вони спеціалізуються на обробці візуальної інформації та здатні ефективно виділяти характеристики з зображень. RNN, в свою чергу, можуть

доповнювати CNN, надаючи додатковий контекст або допомагаючи у зв'язку між об'єктами у відео або серії зображень.

Стандартна архітектура рекурентної нейронної мережі (RNN) включає в себе ключові елементи, які дозволяють їй ефективно обробляти послідовні дані. Основні компоненти цієї архітектури:

1. Вхідний шар (Input Layer): Цей шар приймає вхідні дані. У випадку послідовностей, вхідні дані подаються в мережу по одному елементу за раз (наприклад, одне слово у реченні або один кадр у відео).

2. Прихований шар (Hidden Layer): Прихований шар є серцем RNN. Він здійснює обробку вхідних даних з урахуванням контексту, отриманого з попередніх кроків. Цей шар використовує певну вагову матрицю для трансформації вхідних даних та поточного стану прихованого шару в новий стан. Таким чином, він зберігає інформацію про попередні входи (тобто "пам'ять" мережі).

3. Вихідний шар (Output Layer): На цьому шарі генерується кінцевий вихід на основі поточного стану прихованого шару. Вихід може бути у формі класифікованого об'єкта, прогнозованого значення тощо.

4. Повторювані зв'язки (Recurrent Connections): Ці зв'язки є особливістю RNN. Вони дозволяють передавати інформацію від одного кроку до наступного. У кожному кроці обробки послідовності, стан прихованого шару використовується разом з новим входом для генерації наступного стану та виходу.

5. Функції активації: Такі функції, як тангенс гіперболічний (tanh) або сигмоїд, використовуються для впровадження нелінійності в обчислення прихованого шару.

Обчислення стану прихованого шару [11]:

$$h_t = f(W_{hh}h_{t-1} + W_{xh}x_t + b_h)$$

Де:

- h_t - стан прихованого шару в часовий момент t
- h_{t-1} - стан прихованого шару в попередній момент часу $t-1$
- x_t - вхідний вектор в часовий момент t
- W_{hh} - матриця ваг між прихованим шаром в різні моменти часу
- W_{xh} - матриця ваг між вхідним і прихованим шарами
- b_h - вектор зміщення прихованого шару
- f - функція активації, зазвичай \tanh або ReLU

Обчислення виходу мережі:

$$y_t = g(W_{hy}h_t + b_y)$$

Де:

- y_t - вихідний вектор в часовий момент t
- W_{hy} - матриця ваг між прихованим шаром і виходом
- b_y - вектор зміщення виходу
- g - функція активації для виходу, яка може варіюватися залежно від задачі (наприклад, softmax для класифікації)

Важливим аспектом є те, що RNN використовує свій попередній стан h_{t-1} разом з поточним входом x_t для генерації поточного стану h_t . Це дозволяє мережі зберігати інформацію про попередні дані в послідовності, яка потім використовується для прийняття рішень або прогнозів.

Softmax є функцією активації, яка перетворює вектор виходів від повнозв'язного шару у розподіл ймовірностей. Ключова характеристика softmax полягає в тому, що вона присвоює ймовірності кожному класу таким чином, що їх загальна сума дорівнює одиниці. Це дозволяє інтерпретувати виходи мережі як ймовірності приналежності вхідного зразка до кожного з класів.

Функція softmax визначає ймовірність приналежності вхідного зразка до класу

j за формулою[15]:

$$P(y = j | \mathbf{x}) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

де x - це вхідний вектор, z_j - вихід повнозв'язного шару для класу j , і K - загальна кількість класів.

У контексті глибокого навчання, ключовими підходами у розпізнаванні об'єктів є трансферне навчання та використання великих наборів даних. Трансферне навчання дозволяє застосовувати знання, отримані однією моделлю на одному наборі даних, до інших моделей та даних. Це значно знижує необхідність у великих обсягах тренувальних даних та обчислювальних ресурсів.

Щодо великих наборів даних, їх використання дозволяє системам глибокого навчання "вчитися" з більш різноманітних та складних прикладів, що збільшує точність та надійність систем розпізнавання об'єктів.

Випадки використання нейронних мереж у розпізнаванні об'єктів включають розпізнавання облич та біометричні системи, де CNN ефективно виявляють та ідентифікують людські обличчя, а також виявлення та класифікацію об'єктів у реальному часі, що є ключовим для безпілотних автомобілів, систем відеоспостереження та інших технологій.

2.4. Дослідження специфіки застосування нейронних мереж для виявлення недозволених предметів

Специфіка застосування нейронних мереж для виявлення недозволених предметів є складною та багатогранною задачею, що вимагає глибокого аналізу та ретельного підходу на кожному етапі її реалізації. Початковим кроком у цій сфері є розуміння контексту задачі, що передбачає визначення ключових характеристик недозволених предметів та аналіз контексту та середовища їх використання. Це включає в себе вивчення фізичних властивостей предметів, їх зовнішнього вигляду, а також умов, в яких вони зазвичай виявляються. Важливо також врахувати потенційні зміни у зовнішності предметів залежно від різних факторів, таких як освітлення, кути огляду тощо.

Далі, процес включає розробку архітектури та тренування нейронних мереж, специфічно адаптованих для виявлення цих предметів.

Згорткові нейронні мережі (CNN) є фундаментальним компонентом у задачах комп'ютерного зору, особливо у виявленні недозволених предметів. Основна перевага CNN полягає у їх здатності автоматично та ефективно виявляти важливі особливості на зображеннях без необхідності ручного втручання або попереднього визначення характеристик. Це досягається завдяки використанню згорткових шарів, які сканують вхідне зображення фільтрами, здатними розпізнавати різні аспекти образу, такі як краї, кути, текстури тощо.

Глибина та складність CNN можуть значно варіюватися залежно від конкретної задачі. Наприклад, простіші мережі можуть бути достатніми для виявлення стандартних форм, в той час як більш складні мережі з більшою кількістю шарів та фільтрів потрібні для розпізнавання складних образів у складних умовах. Успіх CNN у виявленні недозволених предметів значною мірою залежить від якості та різноманітності тренувальних даних. Забезпечення, щоб набір даних включав широкий спектр прикладів з різними варіаціями освітлення, кутів, масштабу та контексту, є критично важливим для розвитку мережі, здатної ефективно працювати в реальних умовах. Крім того, важливим аспектом є неперервне тестування та налаштування мережі на основі отриманих результатів, що дозволяє ітеративно покращувати її точність та надійність.

Останнім етапом є оцінка ефективності та точності розроблених моделей. Це включає в себе застосування методів валідації та тестування моделей на реальних або симульованих даних для визначення їх здатності правильно ідентифікувати недозволені предмети у різноманітних умовах. Аналіз помилок, які виникають під час тестування, та розробка стратегій їх усунення є критично важливими для подальшого удосконалення моделі. Це може включати в себе регулювання гіперпараметрів, додавання нових даних для тренування або зміну самої структури мережі.

Таким чином, використання нейронних мереж для виявлення недозволених предметів є складним процесом, що вимагає інтегрованого підходу, починаючи від розуміння задачі та закінчуючи детальним аналізом ефективності використаних моделей.

2.5. Аналіз і порівняння існуючих аналогів

У сучасному світі технологій обробки зображень різні платформи пропонують унікальні функції та спеціалізації, щоб задовольнити різноманітні потреби користувачів. Наприклад, Google Cloud Vision API використовує передові техніки глибокого навчання для досягнення високої точності у виявленні об'єктів, тексту, емоцій та інших характеристик в зображеннях, одночасно інтегруючись з широким спектром сервісів Google Cloud таблиця 2.1. Це забезпечує гнучкість і масштабованість, хоча й може бути відносно високою за ціною.

Amazon Rekognition, у свою чергу, глибоко інтегрується з Amazon Web Services, забезпечуючи сильну платформу для проектів великого масштабу, особливо в галузях безпеки та нагляду. Однак, він може бути менш гнучким для деяких специфічних задач у порівнянні з іншими інструментами.

Clarifai відзначається високою точністю в детальному аналізі зображень і пропонує налаштування та модульні рішення, що можна адаптувати під специфічні потреби. Однак, ця платформа може вимагати більше технічних знань для ефективного використання.

Anyline спеціалізується на мобільних застосунках, зосереджуючись на розпізнаванні тексту та об'єктів, і легко інтегрується за допомогою SDK. Ця платформа ефективна у своїй ніші, але може бути менш універсальною у порівнянні з ширшими рішеннями,

У контексті розвитку ринку та тенденцій, спостерігається інноваційний рух до розвитку більш точних, швидких та гнучких технологій. Зокрема,

важливість спеціалізованих рішень для конкретних галузей, таких як охорона здоров'я чи автономне водіння, зростає.

Для подальшого розвитку, важливо зосередитись на гнучкості та адаптивності систем, здатних виконувати специфічні задачі в різноманітних умовах, а також на покращенні інтеграції з іншими технологічними платформами. Інвестиції в інновації є ключовими для розробки нових технологій, які можуть підвищити точність та швидкість обробки зображень.

Таблиця 2.1

Порівняльна таблиця з існуючих рішень

Система	Час відгуку системи	Масштабованість	Підтримка різних типів даних	Ціноутворення	Приблизна Точність Розпізнавання (%)
Google Cloud Vision API	Швидкий	Висока	Зображення, текст, обличчя, логотипи	Платно	95%
Amazon Rekognition	Швидкий	Висока	Обличчя, текст, небезпечний контент, логотипи, емоції	Платно	90%
Clarifai	Швидкий	Висока	Зображення, відео, текст, аудіо	Платно	92%

Anyline	Швидкий	Висока	OCR, штрих-коди, метричні дані, ID-документи, VIN номери	Платно	88%
---------	---------	--------	--	--------	-----

2.6 Висновки до другого розділу

У ході нашого дослідження ми провели всебічний аналіз сучасних методів розпізнавання об'єктів на зображеннях, зокрема, з акцентом на використанні нейронних мереж. Початково було розглянуто загальні підходи до обробки зображень, від основних принципів цифрового представлення зображень до складніших методів, таких як машинне навчання та глибоке навчання. Особлива увага приділялась попередній обробці зображень, що включає фільтрацію, нормалізацію, сегментацію, морфологічну обробку та корекцію спотворень. Згорткові нейронні мережі (CNN) були визначені як особливо ефективні у класифікації та розпізнаванні об'єктів, в той час як рекурентні нейронні мережі (RNN) виявилися корисними в обробці послідовних даних. Важливість трансферного навчання та використання великих наборів даних для підвищення точності систем розпізнавання також була підкреслена.

Далі ми зосередилися на специфіці застосування нейронних мереж для виявлення недозволених предметів, зазначивши необхідність ретельного вибору та налаштування моделей для досягнення максимальної точності та ефективності. Аналіз і порівняння існуючих технологій та платформ, таких як Google Cloud Vision API, Amazon Rekognition, Clarifai та Anyline, дозволив нам виявити їх унікальні переваги та обмеження, а також зрозуміти загальні тенденції розвитку галузі, які вказують на зростаючу потребу в гнучких, швидких і точних системах обробки зображень. Загальний висновок дослідження підтверджує, що галузь обробки зображень і розпізнавання об'єктів продовжує динамічно

розвиватися, з використанням прогресу в обчислювальній техніці та методах штучного інтелекту, підкреслюючи необхідність адаптації та розуміння цих змін для успішного застосування в різноманітних сферах.

Розділ 3. Розробка архітектури програмного рішення

3.1 Аналіз програмного рішення

У контексті актуальних напрямів у галузі комп'ютерного зору та глибокого навчання, розроблюваний програмний додаток був створений на основі конволюційної нейронної мережі (CNN) для точної класифікації об'єктів на зображеннях. Ця мета вбирає в себе використання многовимірних даних для автоматизованого виявлення та класифікації візуальних патернів, що є фундаментальним аспектом у дослідженнях машинного навчання.

З огляду на поставлені завдання, застосунок покликаний реалізувати ефективну модель, здатну впізнавати та класифікувати об'єкти в категоріях: ніж, пістолет, смартфон, коробки, а також ідентифікувати об'єкти та образи, що не відповідають цим класам. Це передбачає створення нейронної мережі, здатної оптимально обробляти значний обсяг неструктурованих візуальних даних.

В ході імплементації програмного застосунку було прийняте рішення про використання таких фреймворків як TensorFlow та Keras, що сприяло створенню мережі з згортковими (конволюційними), агрегувальними (MaxPooling), повнозв'язними (Dense) та з шарами виключення (Dropout) для ефективного виявлення ознак та запобігання перенавчання. Нормалізація даних і техніки збільшення даних дозволяють моделі адаптуватися до різноманітності вхідних зображень.

Особливу увагу було приділено вибору архітектури моделі, параметрів навчання та стратегій збільшення даних. Ефективність моделі постійно спостерігалася через візуалізацію динаміки втрат та точності, що відіграє ключову роль у забезпеченні адекватності навчального процесу. В подальшому, передбачається дослідження впливу різних гіперпараметрів та оптимізація структури мережі для підвищення її продуктивності та точності.

Як було зазначено вище, у кваліфікаційній роботі було здійснено застосування сучасних технологій програмування та штучного інтелекту, зокрема Python, TensorFlow і Keras, що забезпечило розробку ефективної

системи для класифікації зображень.

За статистикою, дев'яносто відсотків додатків для машинного навчання пишеться на мові Python, оскільки у Python широкий спектр застосування, а також підтримка численних бібліотек і фреймворків. Саме тому, Python був використаний як основна мова для написання додатку.

Для реалізації проекту були використані бібліотеки TensorFlow та Keras. TensorFlow - це потужна відкрита платформа для машинного навчання, розроблена Google, яка забезпечує гнучкість та ефективність у роботі з великими наборами даних. Keras, інтегрований у TensorFlow, є високорівневим інтерфейсом для побудови та тренування моделей глибокого навчання, забезпечуючи інтуїтивне програмування.

TensorFlow відіграв ключову роль у проекті, надаючи можливість обробки великих об'ємів даних та реалізації складних алгоритмів глибокого навчання. Основні можливості TensorFlow включають підтримку різноманітних операцій на тензорах, автоматичне диференціювання для обчислення градієнтів, а також масштабування обчислень на CPU та GPU. Використання TensorFlow у реалізації моделі забезпечило високу продуктивність та точність при класифікації зображень.

Keras, як частина TensorFlow, використовувався для спрощення процесу створення та навчання нейронних мереж. Використання Keras у розробці нейронних мереж сприяло ефективному тренуванню та валідації моделі, що дозволило досягти високої точності класифікації.

3.2 Дослідження підготовки та обробки даних

У контексті реалізації нейронної мережі для класифікації зображень, критичним аспектом є детальний вибір датасету. Задля підвищення рівня безпеки на підприємствах, було прийняте рішення класифікувати такі недозволені предмети такі як, ножі та пістолети. Також, задля розширення варіативності класифікації було додано клас коробок та смартфонів. Отже, досліджуваний

датасет охоплює зображення категоризовані у чотири основні класи: ножі, пістолети, смартфони та коробки, також, в ході реалізації НМ, було прийняте рішення додати категорію "інше", що представляє негативні зразки даних класів, що допоможе розпізнавати інші предмети, та збільшити датасет.

Класи для класифікації:

- Ножі
- Пістолети
- Смартфони
- Коробки
- Інше

Ця множинність категорій забезпечує багатofакторний спектр для навчання моделі, що ускладнює навчання. Джерела даних було сегментовано за класами, що сприяє систематичному збору та обробці даних. Було використано формат JPG для зображень, що відповідає загальноприйнятим стандартам у галузі обробки зображень.

Важливим аспектом є характеристика зображень, зокрема їх розмір та кольорова гамма. Також було впроваджено нормалізацію до розміру (64, 64) та (128, 128) в останній з імплементацій моделі, із трьома каналами (RGB), що є оптимізаційним компромісом між детальністю та обчислювальними вимогами, це було впроваджено зміною гіперпараметра розміру зображення. В результаті отримані вхідні тензори (багатовимірний масив), $64*64*3$, та $128*128*3$ в останньому експерименті .

Присвоєння унікальних міток кожному класу об'єктів було автоматизовано в програмному коді, що забезпечує дискримінативну здатність моделі. Включення категорії "інше" з унікальною міткою для негативних зразків сприяє поліпшенню загальної точності класифікації.

Кожна категорія відіграє специфічну роль у навчанні, дозволяючи моделі вчинити розрізнення між різними об'єктами. Це є основою для ефективного навчання класифікації. Структурована система міток, де кожному класу

присвоєна унікальна мітка, забезпечує точність у визначенні відповідності зображень до певних класів, а також ефективно виявляє випадки, що не відповідають жодному з визначених класів.

Аналіз розміру зображень відіграє ключову роль у балансі між вимогами до обчислювальних ресурсів та необхідною деталізацією для ефективного навчання моделі. Перед вибором розміру були протестовані такі розміри зображень:

- 1) 32*32 пікселі
- 2) 64*64 пікселі
- 3) 128*128 пікселів
- 4) 256*256 пікселів

З аналізу стало зрозуміло що розмір 32*32 пікселі є недостатнім для ефективного процесу класифікації, а розмір 256*256 не підходить об'ємам обчислювальних ресурсів, а саме, процес навчання займає занадто багато оперативної пам'яті.

Кількість та різноманітність даних впливають на здатність моделі адаптуватися до різних умов. В процесі реалізації програмного додатку було зроблено висновок, що кількість даних взятих з публічних датасетів виявилась недостатньою, тож було прийняте рішення про додавання аугментації. Використання методів аугментації даних є ефективним засобом збільшення обсягу датасету, забезпечуючи моделі більшу точність [9].

Аугментація даних виконується з метою штучного розширення набору даних, що включає модифікації як у геометричному, так і у візуальному вимірах. Це дозволяє створити додаткові варіації даних, які можуть бути використані для підвищення точності та здатності моделі до узагальнення. Аугментація даних має безпосередній вплив на зниження ризику перенавчання моделі, забезпечуючи її здатність до адаптації до нових, невідомих зразків.

В програмному коді було використано методи аугментації, такі як ротація, зсув, нахил, масштабування, та горизонтальне віддзеркалення. Ці трансформації

реалізовані через модуль фреймворку Keras `ImageDataGenerator` , що автоматично застосовує ці зміни до набору даних під час навчання, сприяючи підвищенню стійкості та гнучкості моделі. Цілісність трансформацій полягає у створенні даних, які максимально відповідають реальним умовам застосування, де об'єкти можуть бути представлені під різними кутами та в різних масштабах. Ефективність таких трансформацій визначається їх здатністю зменшити ризик перенавчання та збільшити загальну точність моделі.

Оперативне розділення даних на навчальні та тестувальні набори є фундаментальною необхідністю у запобіганні явища перенавчання. Цей процес має забезпечити, що модель не просто запам'ятовує вхідні дані, а здатна ефективно адаптуватися до нових, невідомих даних.

Основні критерії, які були враховані у дослідженні, включають:

1) Співвідношення навчальних та тестувальних даних: вибір пропорції поділу даних є ключовим; стандартними є співвідношення 80/20 або 70/30 для навчальних та тестувальних наборів відповідно. З урахуванням потреби в більшій кількості даних, було прийняте рішення про пропорцію 80/20.

2) Також була використана рандомізація з урахуванням стратифікації. Рандомізація гарантує репрезентативність вибірки, тоді як стратифікація забезпечує приблизно рівну кількість представників кожного класу у наборах даних.

3) Важливим фактором є наявність валідаційного набору даних у процесі навчання та тестуванні. Використання валідаційного набору дозволяє точно калібрувати гіперпараметри перед оцінкою моделі на тестовому наборі. Використання валідаційного набору при навчанні поліпшує процес раннього виявлення перенавчання та оптимізації моделі.

3.3 Опис розробки та архітектура моделі

В рамках дослідження, здійсненого для даної роботи, були розглянуті різноманітні архітектури нейронних мереж. Важливо визначити специфіку

кожної архітектури, щоб зрозуміти їх придатність для різних доменів задач. Архітектури, такі як AlexNet, VGG, Inception та ResNet, представляють собою важливі віхи в розвитку глибокого навчання. Наприклад, архітектура ResNet, що характеризується значною глибиною, є здатною ідентифікувати складніші особливості, в той час як більш поверхневі архітектури, такі як AlexNet, можуть бути більш ефективними для задач із меншою складністю [8]. Кожна архітектура нейронної мережі має свої специфічні переваги та недоліки. Глибокі мережі, такі як ResNet, можуть демонструвати високу точність, але потребують значних обчислювальних ресурсів та тривалого часу для навчання. З іншого боку, менш глибокі архітектури можуть бути швидшими у навчанні та менш вимогливими до обсягів навчальних даних, хоча їх ефективність може знижуватися при розв'язанні складних задач. Враховуючи поставлену задачу було прийняте рішення створити власну архітектуру.

Задача, що стоїть перед дослідженням, полягає у класифікації недозволених зображень предметів на підприємстві (ножі, пістолети, смартфони, коробки). Враховуючи відносну простоту цієї задачі, не потребується застосування особливо глибоких нейронних мереж. Створена архітектура моделі, відповідає вимогам цієї задачі. Програмний код включає завантаження та обробку зображень, створення, навчання та оцінку моделі (рис. 3.1). Також реалізована візуалізація процесу навчання, демонструючи динаміку змін точності та втрат протягом навчальних епох. Це є важливим для оцінки ефективності та стабільності процесу навчання.

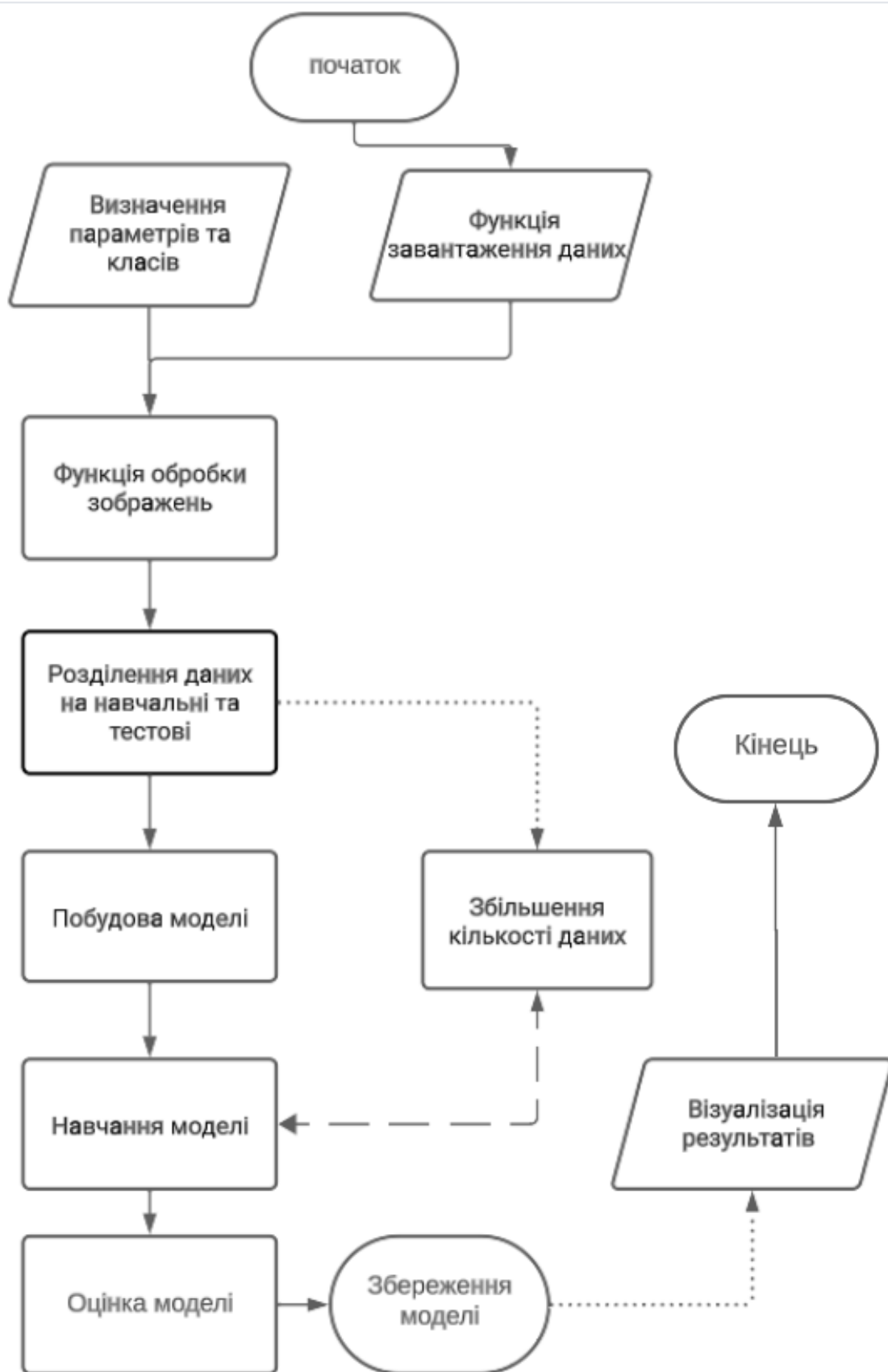


Рисунок 3.1 Діаграма програмного коду створення моделі
Загальний алгоритм створення, навчання і оцінки НМ:

1. Ініціалізація параметрів та класів:
 - 1.1. Встановлення класів об'єктів, розмірів вхідного зображення, кількості каналів та інших параметрів.
2. Функція обробки зображень:
 - 2.1. Завантаження та обробка зображень з директорії для кожного класу.
 - 2.2. Нормалізація зображень.
3. Завантаження та підготовка даних:
 - 3.1. Збір зображень та їх міток у масиви.
 - 3.2. Розділення даних на навчальний та тестувальний набори.
4. Створення та компіляція моделі:
 - 4.1. Определение архитектуры сверточной нейронной сети.
 - 4.2. Компіляція моделі з визначенням функції втрат та оптимізатора.
5. Аугментація даних:
 - 5.1. Збільшення різноманітності навчальних даних шляхом застосування трансформацій зображення.
6. Навчання моделі:
 - 6.1. Навчання моделі на аугментованих даних.
7. Оцінка моделі:
 - 7.1. Перевірка ефективності моделі на тестувальному наборі даних.
8. Збереження моделі:
 - 8.1. Зберігання навченої моделі для подальшого використання.
9. Візуалізація результатів навчання:
 - 9.1. Побудова графіків для візуального представлення процесу навчання (втрати та точності).

Алгоритми для кожної з основних функцій.

Функція обробки зображень (`process_image``)

Вхід: Шлях до зображення, цільовий розмір.

Вихід: Оброблене зображення.

Кроки:

- 1) Завантаження зображення з вказаного шляху.
- 2) Зміна розміру зображення до заданого розміру.
- 3) Конвертація зображення у масив.
- 4) Нормалізація пікселів зображення (ділення на 255).
- 5) Повернення обробленого зображення.

Підготовка даних (`'load_data'`)

Вхід: Розмір вхідного зображення, кількість каналів.

Вихід: Масиви зображень та їх міток.

Кроки:

- 1) Ітерація по кожному класу.
- 2) Завантаження та обробка зображень для кожного класу.
- 3) Додавання зображень та їх міток у відповідні масиви.
- 4) Завантаження та обробка "негативних" зразків (які не належать до визначених класів).
- 5) Повернення масивів зображень та міток.

Створення та компіляція моделі

Кроки:

- 1) Визначення архітектури згорткової нейронної мережі
- 2) Компіляція моделі з використанням `'categorical_crossentropy'` як функції втрати та `'adam'` як оптимізатора.

Навчання моделі

Вхід: Навчальні дані (зображення та мітки), параметри навчання.

Кроки:

- 1) Навчання моделі на аугментованих навчальних даних.
- 2) Використання тестувального набору даних для валідації.

Оцінка та збереження моделі

Кроки:

- 1) Оцінка моделі на тестувальному наборі даних.
- 2) Виведення результатів (точність та втрати).
- 3) Збереження моделі.

Була досліджена наукова праця [14] 2020 року, в якій ретельно проаналізовано використання менших згорткових ядер. Автори зазначають, що класичні CNN часто використовують великі згорткові ядра (наприклад, 11×11), що призводить до збільшення кількості тренувальних параметрів. Натомість вони пропонують використання менших ядер (наприклад, 3×3 або 7×7), що може зменшити кількість параметрів та поліпшити ефективність навчання. Автори проводять експерименти, використовуючи свою оптимізовану структуру CNN на різноманітних наборах даних, порівнюючи результати з іншими відомими алгоритмами, такими як AlexNet та VGG. Цей підхід демонструє поліпшення у точності розпізнавання зображень. Після аналізу даної праці було прийняте рішення про використання менших згорткових ядер.

Після детального аналізу архітектур згорткових нейронних мереж класифікації, та декількох тестових архітектур, було створено власну архітектуру моделі (рис. 3.2), що підходить до вимог та ефективно вирішує поставлену задачу класифікації зображень. На рисунку показані згорткові шари тільки з 1 фільтром, інші фільтри будуть помітні при тривимірному моделюванні.

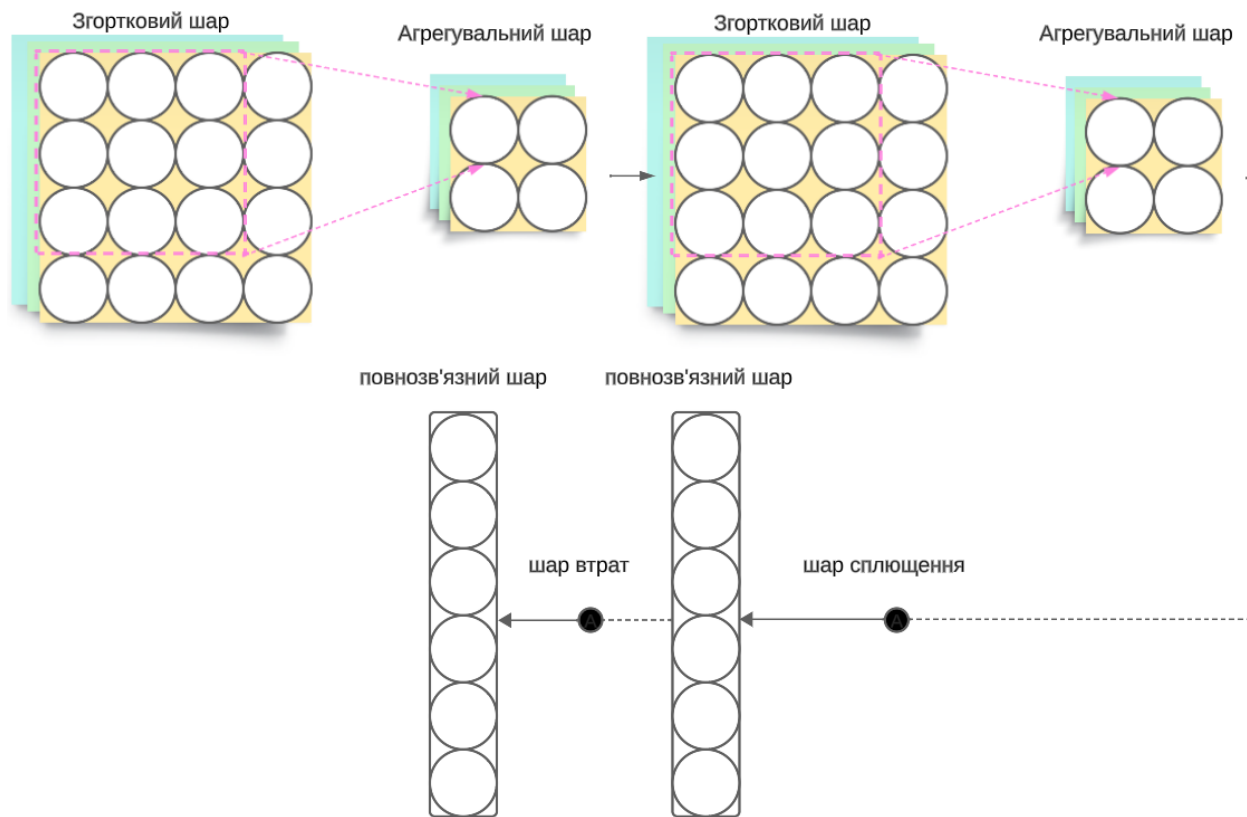


Рисунок 3.2 Архітектура моделі

Список шарів розробленої згорткової нейронної мережі:

- 1) Згортковий шар (32 фільтри, розмір фільтра 3x3, функція активації ReLU)
- 2) Агрегувальний шар (розмірність агрегації шару 2x2)
- 3) Згортковий шар (64 фільтри, розмір фільтра 3x3, функція активації ReLU)
- 4) Агрегувальний шар (розмірність агрегації шару 2x2)
- 5) Випрямляючий шар (Flatten)
- 6) Повнозв'язний шар (кількість нейронів 128, функція активації ReLU)
- 7) Спеціалізований шар втрат (процент відключених нейронів 50%)
- 8) Повнозв'язний шар (кількість нейронів 5, функція активації softmax)

В архітектурі розроблюваної моделі використовувалися згорткові шари Conv2D. У цьому дослідженні, шар Conv2D імплементовано двічі: перший з 32

фільтрами 3×3 та другий з 64 фільтрами, також обидва шари використовують функцію активації ReLU (Rectified Linear Unit).

У рамках даної задачі класифікації, Conv2D сприяє ідентифікації властивостей різних об'єктів, таких як ножі, пістолети, смартфони та коробки. Шари Conv2D послідовно виявляють просторові характеристики від простих до складніших. Конволюційні шари складаються з набору фільтрів (або ядер). Кожен фільтр є невеликою матрицею ваг, яка "пересувається" (конволює) по вхідному зображенню або функції, створюючи карту особливостей. Процес множення елементів фільтру на відповідні елементи вхідного зображення називається конволюцією. В результаті конволюції формується карта активації, яка відображає, де та як сильно виявлені особливості в зображенні. Фільтри автоматично навчаються виявляти корисні особливості, такі як краї, кути, текстури та інші важливі візуальні елементи зображення. Кількість фільтрів вказує, скільки різних фільтрів буде використовуватися в шарі. Більша кількість фільтрів дозволяє мережі виявляти більшу кількість особливостей, але також збільшує обчислювальні витрати, в ході експериментів було обрано 32 та 64 фільтри у першому та другому згортковому шарі відповідно. Розмір (зазвичай 3×3 , 5×5) фільтра впливає на область, яку він охоплює на вхідному зображенні. Менші фільтри вловлюють більш детальні особливості, в той час як більші фільтри забезпечують більш широке охоплення, тому було прийнято рішення про використання розміру 3×3 , тому що розміри вхідного зображення невеликі 64×64 та 128×128 у останніх двох експериментах.

Також в даній моделі використовувався агрегувальний шар MaxPooling2D. У розробленій моделі, MaxPooling2D з розмірністю 2×2 , застосовується після кожного згорткового шару, забезпечуючи прогресивне зниження розмірності просторових даних. Це сприяє зосередженню моделі на більш важливих атрибутах зображення та запобігає перенавчанню.

Отже через послідовне застосування згорткових Conv2D та агрегувальних MaxPooling2D шарів, створюється мережа, яка ефективно виявляє

характеристики об'єктів на зображеннях, а також зменшує розмірність даних перед передачею їх в повнозв'язні шари для класифікації.

Наступним шаром в моделі є випрямляючий шар (Flatten). Flatten є фундаментальним шаром у глибинних нейронних мережах, відповідальним за конвертацію багатовимірних тензорів даних у одновимірні вектори. Ця операція є критичною для забезпечення сумісності між великорозмірними вихідними даними з конволюційних та пулінгових(агрегувальних) шарів та наступними повнозв'язними шарами. У програмному кодї, Flatten виконує важливу функцію інтерфейсу, що дозволяє передавати дані з просторово-орієнтованих шарів до шарів, які виконують високорівневу класифікацію. Перетворення вихідних даних конволюційних шарів у формат, придатний для повнозв'язних шарів через Flatten, є ключовим для ефективної інтеграції просторової інформації з глобальним аналізом.

Далі у архітектурі даної моделі йде два Dense шари , також відомі як повнозв'язні шари, він є структурним компонентом глибинних нейронних мереж. Кожен нейрон у Dense шарі формує повні зв'язки з усіма активаціями попереднього шару, тим самим здійснюючи функцію інтеграції та синтезу інформації. З цього випливає, що інформація з усіх нейронів враховується при розрахунку виходу. Кожне з'єднання між нейронами має вагу, а кожен нейрон має зсув (bias), які під час навчання адаптуються, щоб мережа могла краще виконувати задану задачу. У даному дослідженні, застосування Dense шару спрямоване на здійснення кінцевої класифікації зображень, де кожний клас представлений окремим виходом нейрона. Кількість нейронів у Dense шарі має безпосередній вплив на здатність мережі до узагальнення. Підвищена кількість нейронів може підвищити ризик перенавчання, особливо при обмеженому обсязі тренувальних даних. Натомість, недостатня кількість нейронів може обмежити потенціал мережі для виявлення складних залежностей у даних, тому враховуючи це була обрана оптимальна кількість нейронів 128.

Між двома повнозв'язними шарами знаходиться спеціалізований шар втрат

(Dropout). Застосування Dropout значно покращує здатність моделі до узагальнення, коли використовується в складних мережах з великою кількістю. Даний шар допомагає зменшити перенавчання, зберігаючи при цьому ефективність моделі на невидимих раніше даних. Цей шар випадково "відключає" деякі нейрони під час тренування, що допомагає уникнути перенавчання (overfitting). В даній роботі було обрано значення відключення 0.5, 50% нейронів буде випадково відключено, що допоможе зменшити залежність нейронної мережі від конкретних властивостей вхідних даних, сприяючи тим самим розподілу "навантаження навчання" по всій мережі. Це важливо для забезпечення того, що модель не стане "занадто специфічною" до тренувального набору даних, що може призвести до поганого узагальнення на нових, невиданих раніше даних. Крім того, такий підхід сприяє формуванню більш міцних і стійких ознак у моделі, оскільки вона навчається не покладатися на будь-яку окрему групу нейронів, що може бути особливо корисним у випадках, коли вхідні дані містять шум або відхилення.

Також було прийняте рішення про використання у вихідному повнозв'язному шарі, функції активації softmax. Загалом, підхід та імплементація коректно відповідають вимогам сучасного глибинного навчання та ефективно розв'язують поставлену задачу класифікації зображень.

3.4 Навчання, оцінка та візуалізація результатів моделі

Розмір партії (batch size) та кількість епох є ключовими гіперпараметрами у процесі навчання нейронної мережі. У даному дослідженні, після декількох тестів, розмір партії було встановлено як 64, що є балансом між обчислювальною ефективністю та здатністю мережі до загальної адаптації. Вибір кількості епох, у цьому випадку 10, ґрунтується на потребі забезпечити достатню конвергенцію моделі без перенавчання.

Для оптимізації використовується алгоритм Adam. Алгоритм Adam (Adaptive Moment Estimation) представляє собою метод оптимізації, який

використовується для оновлення ваг у нейронних мережах на основі обчислення першого (моменту) та другого (невиправлених моментів) моментів градієнтів. Особливість Adam полягає в його здатності адаптувати швидкість навчання для кожного параметра.

Структурні компоненти алгоритму Adam:

1) Оцінка Першого Моменту (середнього): Adam обчислює експоненційно зважене рухоме середнє градієнтів. Це допомагає згладити шум у градієнтах та забезпечує більш стабільне оновлення ваг.

2) Оцінка Другого Моменту (невиправлений момент): Алгоритм також використовує квадрат градієнтів, який допомагає адаптувати швидкість навчання відповідно до величини градієнтів, забезпечуючи більш точне налаштування швидкості навчання для кожного параметра.

3) Коригування Моментів: Оскільки обидва моменти ініціалізуються як нулі, вони мають зміщення до нуля на початкових етапах навчання. Алгоритм Adam коригує ці моменти, щоб компенсувати їхнє зміщення.

Цей алгоритм має такі переваги

1) Адаптивність: автоматично регулює швидкість навчання для кожного параметра, що робить його ефективним для задач з різними шкалами та різною чутливістю до параметрів.

2) Стабільність: Через використання згладжування градієнтів, Adam демонструє більшу стабільність під час навчання в порівнянні з іншими методами оптимізації.

Процес навчання у даному випадку включає ітеративні кроки, де кожен крок (або епоха) включає пряме поширення (forward propagation) та зворотне поширення помилок (backpropagation). Пряме поширення передбачає проходження вхідних даних через мережу та отримання виходу. Потім, під час зворотного поширення, обчислювалися градієнти втрати (loss) щодо кожного параметра мережі та здійснювалося їх оновлення, щоб мінімізувати ці втрати.

Під час навчання моделі, що розробляється відбувався періодичний моніторинг. Моніторинг процесу навчання має ключове значення для визначення прогресу та ефективності навчання моделі. Це включає в себе спостереження за змінами у втратах та точності на тренувальних та валідаційних даних. Для полегшення процесу моніторинга та коригування моделі, було впроваджено візуалізацію точності та втрат на графіку.

3.5 Висновки до третього розділу

У третьому розділі була представлена розробка та реалізація програмного застосунку, заснованого на конволюційній нейронній мережі (CNN), для точної класифікації об'єктів на зображеннях. Основна увага була зосереджена на створенні та оптимізації моделі, яка здатна ефективно впізнавати та класифікувати різні об'єкти, зокрема ніж, пістолет, смартфон, коробки та ідентифікувати зображення, що не належать до цих категорій.

У процесі розробки було використано сучасні технології та інструменти, зокрема мову програмування Python та фреймворки TensorFlow та Keras. Це дозволило створити мережу з різними типами шарів, включаючи згорткові (конволюційні), агрегувальні (MaxPooling), повнозв'язні (Dense) та шари виключення (Dropout), які сприяли ефективному виявленню характеристик та запобіганню перенавчанню.

Було проведено глибокий аналіз підготовки та обробки даних, який включав вибір датасету, нормалізацію та аугментацію зображень для забезпечення високої точності моделі. Також було ретельно вивчено архітектуру моделі, оскільки вибір кожного компонента має важливе значення для ефективності класифікації.

З цього можемо зробити висновок, що ретельно продумана архітектура мережі, а також уважний підхід до підготовки та обробки даних, є ключовими для успішної реалізації програмного застосунку. Використання сучасних технологій та методологій у сфері машинного навчання та комп'ютерного зору

дозволяє досягти високої точності та ефективності у класифікації зображень, що робить розроблений програмний застосунок важливим інструментом у вирішенні задач сучасного комп'ютерного зору.

Розділ 4. Опис, тестування та аналіз роботи застосунку

4.1 Вступ до четвертого розділу

У четвертому розділі кваліфікаційної роботи представлено детальний аналіз та опис роботи розробленого застосунку, який базується на моделі глибокого навчання для класифікації об'єктів.

Основною метою цього розділу є демонстрація функціональності та ефективності застосунку, а також висвітлення ключових аспектів його розробки та тестування.

Робота застосунку базується на використанні глибокої нейронної мережі, розробленої з використанням TensorFlow та Keras. Архітектура мережі включає послідовність згорткових шарів (Conv2D), шарів пулінгу (MaxPooling2D), випрямляючого шару (Flatten), повнозв'язних шарів (Dense) та шару Dropout для запобігання перенавчанню.

Модель навчалася на даних, які включали зображення чотирьох класів об'єктів: ножі, пістолети, смартфони та коробки. Для кожного класу було зібрано відповідний набір даних зображень, що дозволило ефективно тренувати мережу.

Також розглядається інтерактивна частина застосунку, що включає відображення результатів класифікації. Забезпечується візуальне представлення класифікованих об'єктів та відсоткове вираження ймовірності приналежності до кожного з класів. Це дозволяє користувачеві швидко оцінити результати роботи моделі. Важливою складовою застосунку є його графічний інтерфейс, реалізований за допомогою бібліотеки Tkinter. Це забезпечує простий та зручний спосіб взаємодії з користувачем. Інтерфейс дозволяє виконувати завантаження зображень з файлової системи або захоплювати їх безпосередньо з веб-камери, що робить застосунок гнучким та багатофункціональним.

4.2 Аналіз тестування та проведених експериментів

Параметр `steps_per_epoch`, визначає кількість ітерацій навчання в кожній епосі. Це дозволяє контролювати, скільки разів мережа буде оновлюватися

протягом однієї епохи, що є критично важливим для забезпечення адекватного навчання на великих наборах даних. Збільшення кількості кроків може покращити здатність мережі до загальної адаптації, але також збільшує обчислювальні вимоги. В ході першого експерименту було встановлено 1000 ітерацій в кожній з десяти епох, графіки точності й валідації на (рис. 4.1) та (рис. 4.2). Після чого, під час другого експерименту було збільшено цей параметр до 2000 що дало приріст точності до 85 відсотків (рис. 4.4), що є непоганим результатом для такої кількості вхідних даних у навчанні, втрати знизилися до 45%. В наступному експерименті, було збільшено кількість ітерацій до 5000. Що дало значний приріст у точності моделі, до 87% (рис. 4.6), що безперечно є гарним результатом. Крім цього зменшилася кількість втрат до 37% (рис. 4.5). Після низки експериментів з гіперпараметрами було прийняте рішення збільшити розмірність пікселів зображення до 128*128 з кількістю ітерацій 5000 графіки (рис. 4.7) та (рис.4.8). Аналізуючи графіки, на кривих валідації, точність та втрати покращилися. Криві стали більш плавними, що є ознакою якісного навчання моделі. Також, була спроба збільшити кількість ітерацій за епоху до 10000, проте модель почала “запам’ятовувати” відповіді, що було спричинено невеликою кількістю даних для навчання, а саме 24600 зображень, без урахування аугментації даних.

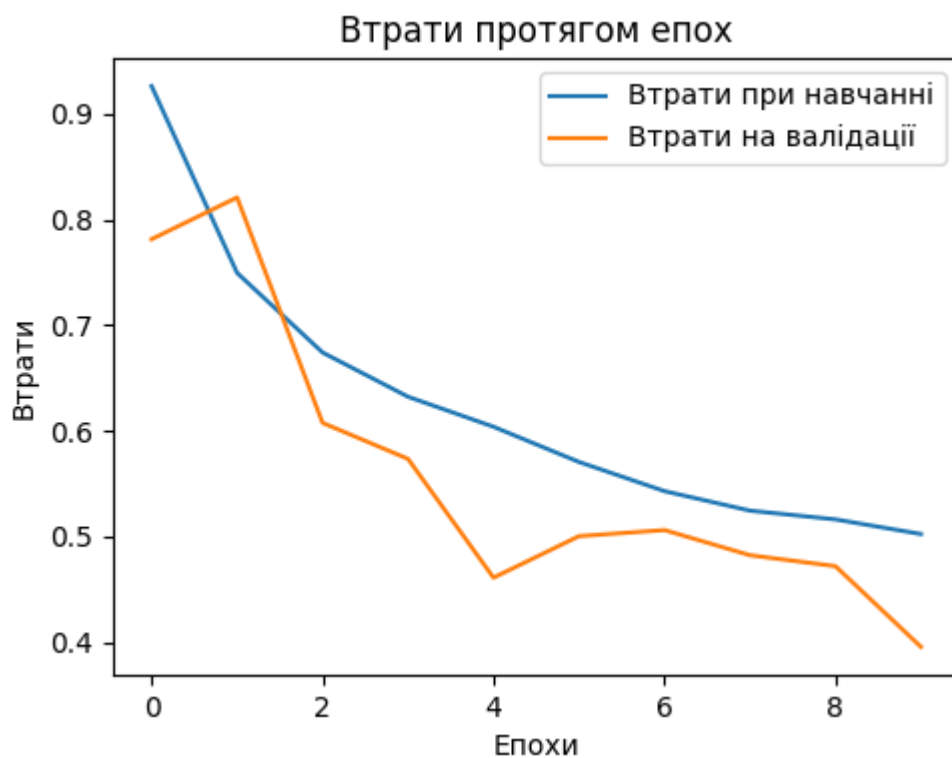


Рисунок 4.1 Графік втрат протягом епох 1й експеримент

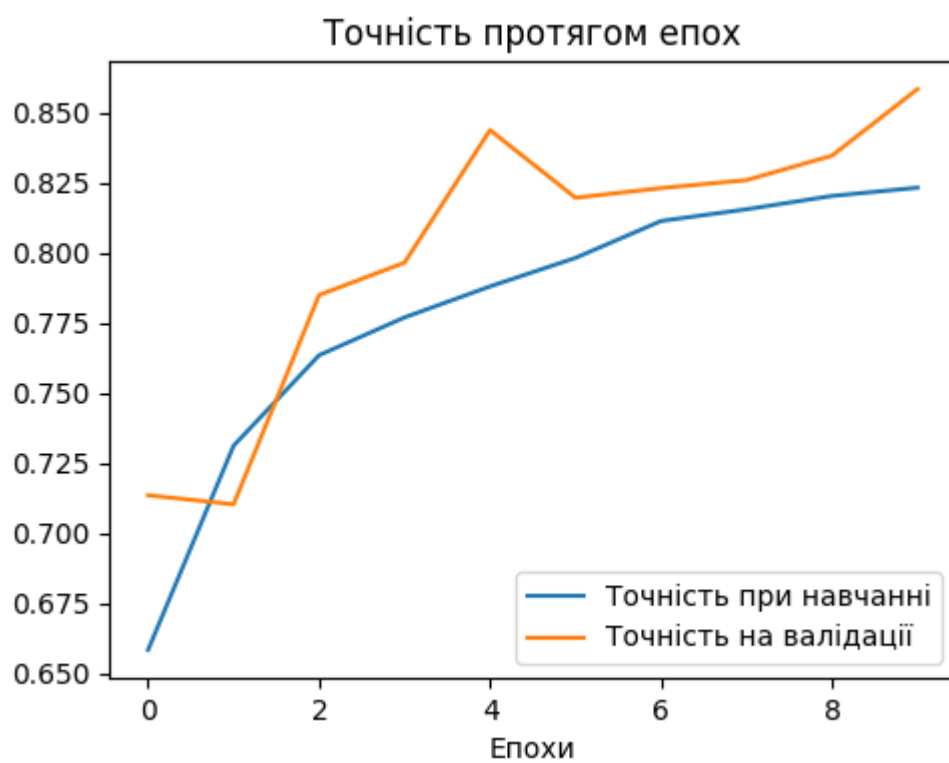


Рисунок 4.2 Графік точності протягом епох 1й експеримент

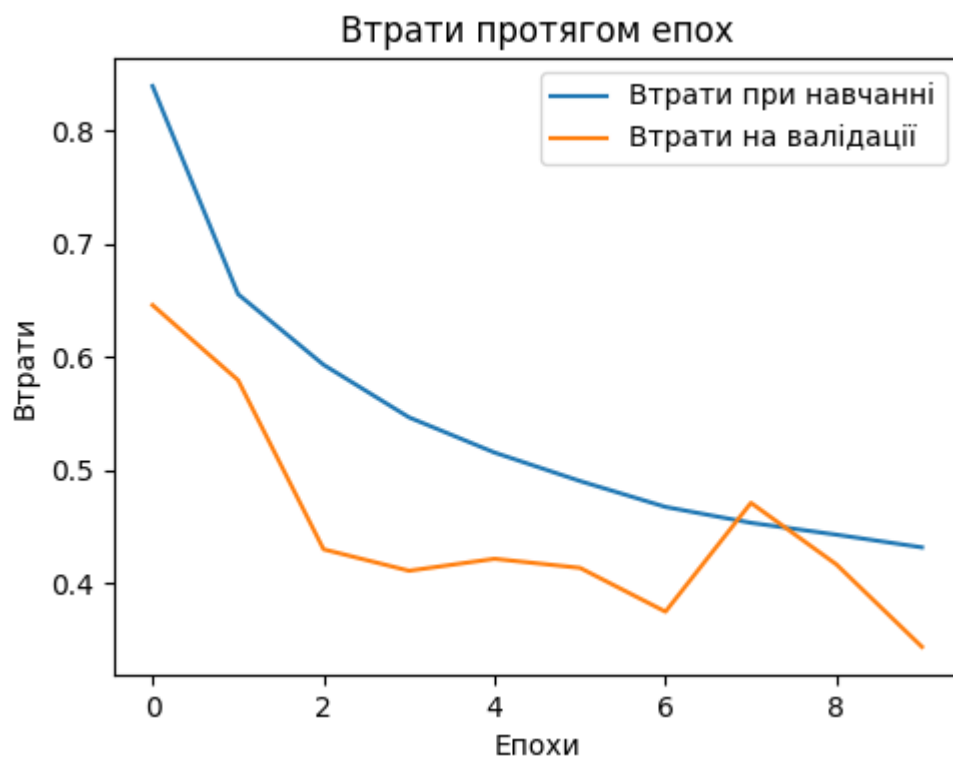


Рисунок 4.3 Графік втрат протягом епох 2й експеримент

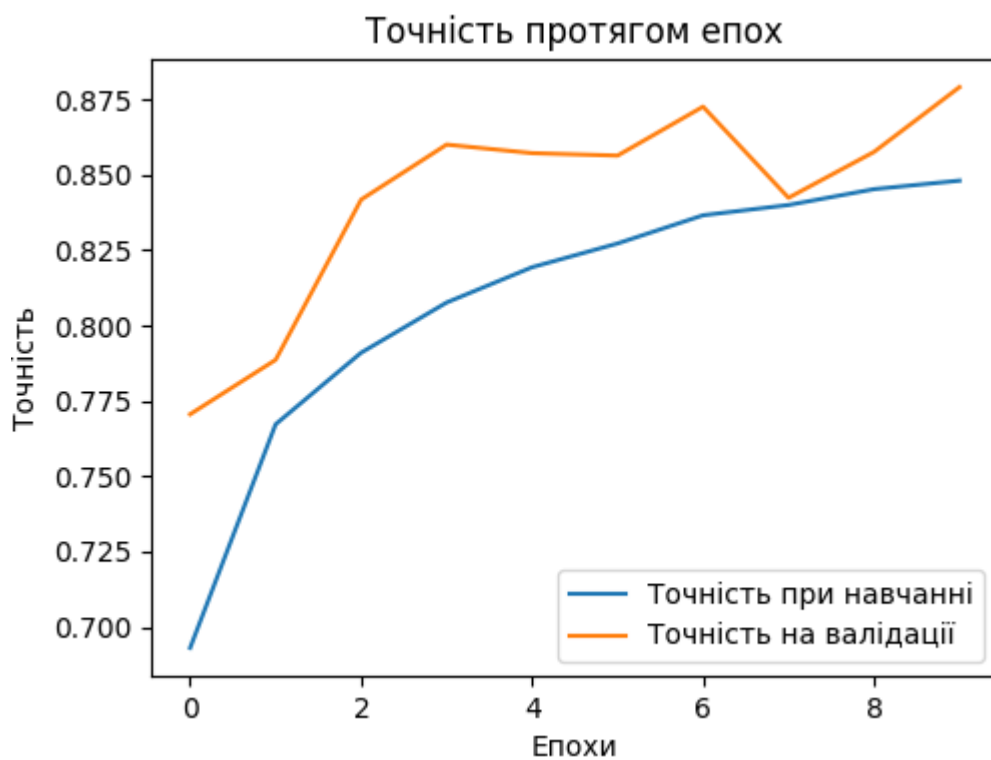


Рисунок 4.4 Графік точності протягом епох 2й експеримент

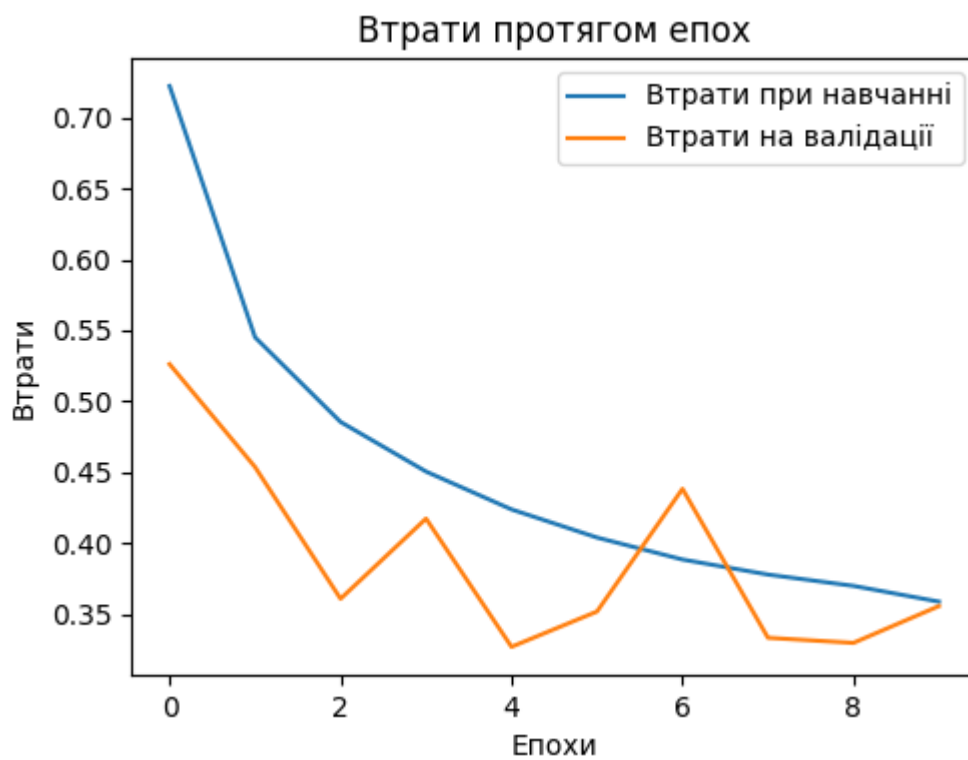


Рисунок 4.5 Графік втрат протягом епох 3й експеримент

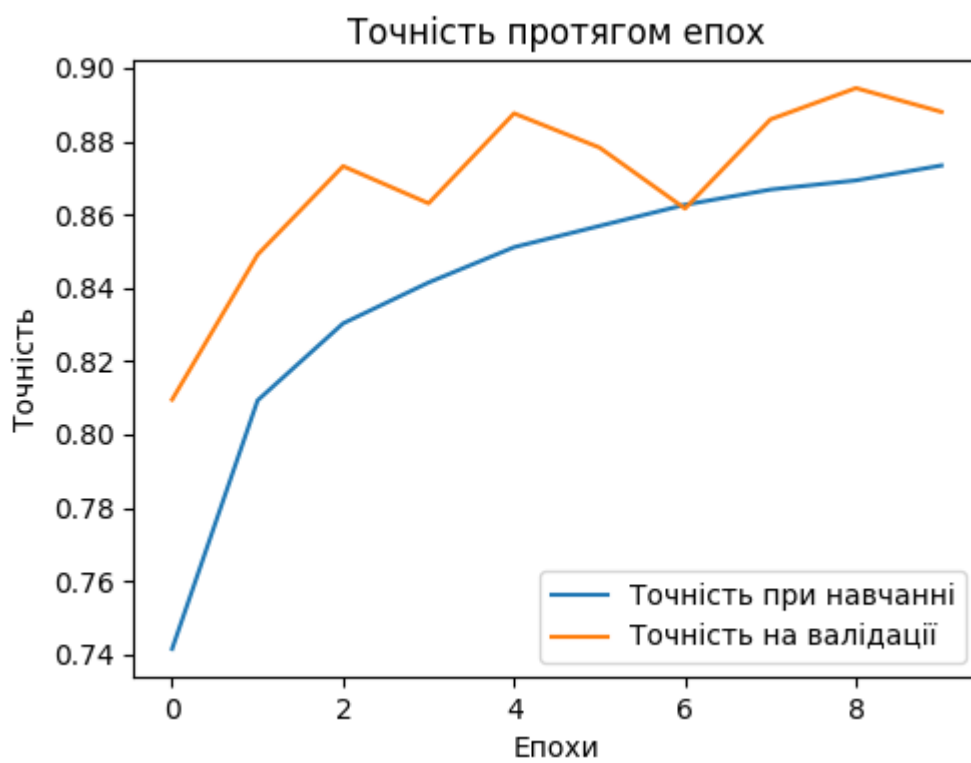


Рисунок 4.6 Графік точності протягом епох 3й експеримент

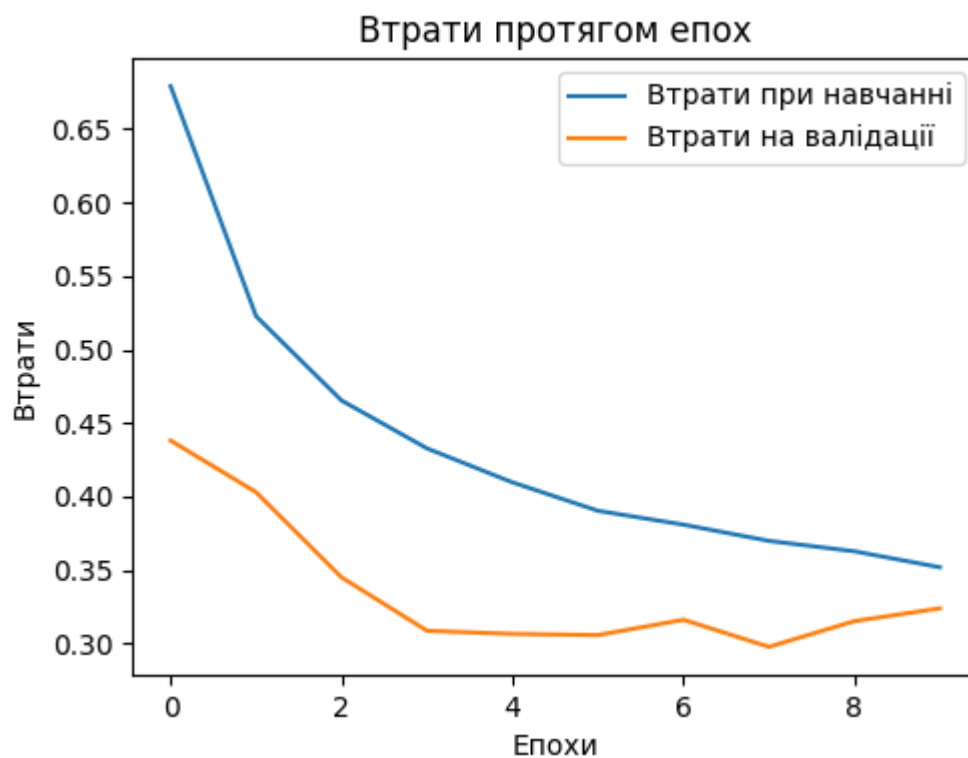


Рисунок 4.7 Графік втрат протягом епох 4й експеримент

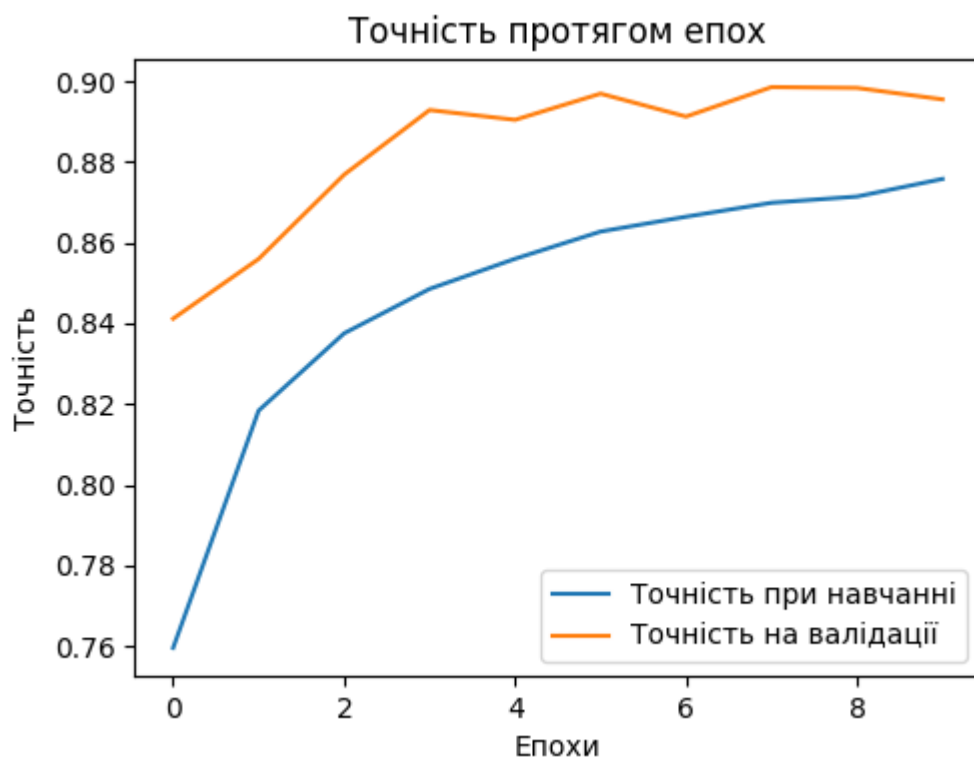


Рисунок 4.8 Графік точності протягом епох 4й експеримент

`Batch_size` є важливим гіперпараметром у процесі навчання нейронних мереж. Він визначає кількість зразків, які будуть оброблятися за один раз під час навчання. `Batch_size` визначає кількість тренувальних зразків (зображень), які потрібно обробити перед оновленням внутрішніх параметрів моделі (вагів). Іншими словами, це кількість зразків, які нейронна мережа "бачить", перш ніж зробити крок в процесі навчання (крок означає оновлення вагів). У даній роботі після експериментів і спостереженнями за оперативною пам'яттю, цей параметр був встановлений як 64, що є збалансованим вибором для більшості обчислювальних середовищ. Вищі значення можуть прискорити навчання, але також вимагають більше пам'яті та можуть призвести до нестабільності процесу навчання.

Розмір вхідного зображення впливає на обчислювальні вимоги та здатність мережі вчитися з високою точністю. У розробленому коді розміри вхідного зображення були встановлені як (64, 64), що забезпечує баланс між обчислювальною ефективністю та здатністю мережі розпізнавати деталі на зображенні. Проте після трьох позитивних тестів з різною кількістю ітерацій, розмір було змінено на (128,128), що гарно вплинуло на втрати у валідації, це можна спостерігати на (рис. 4.7) та (рис. 4.8) в порівнянні з іншими графіками. Також маю зазначити, що значно збільшився час навчання. Функція `process_image` ефективно адаптує розміри зображень до потрібного формату, зберігаючи при цьому адекватну якість візуальної інформації.

Модель визначає класи по наступним датасетам:

- Ножі: 4,427 зображень
- Пістолети: 3,975 зображень
- Коробки: 2,050 зображень
- Смартфони: 2,444 зображень
- Інше: 11,756 зображень

Загальний об'єм даних становить 24,652 зображення.

Резюмуючи, у рамках дослідження було проведено п'ять основних експериментів для вивчення впливу гіперпараметрів на навчання нейронної мережі. Ключовими параметрами які були перевірені, це кількість ітерацій на епоху та розмір вхідних зображень.

1) Перший експеримент: Зосереджено на впливі кількості ітерацій на епоху. Було встановлено 1000 ітерацій в кожній з десяти епох, із використанням зображень розміром 64x64 пікселів. Результати (рис. 4.1, 4.2) показали задовільну точність та втрати.

2) Другий експеримент: Кількість ітерацій було збільшено до 2000 для тих самих 10 епох, зберігаючи розмір зображень 64x64. Результати (рис. 4.3, 4.4) вказали на підвищення точності до 85% та зменшення втрат до 45%.

3) Третій експеримент: Кількість ітерацій зросла до 5000, що ще більше покращило точність (до 87%) і знизило втрати (до 37%), що видно з графіків на рис. 4.5 та 4.6.

4) Четвертий експеримент: Внесено зміни до розміру зображень, збільшивши їх до 128x128 пікселів, залишивши кількість ітерацій 5000. Результати (рис. 4.7, 4.8) показали подальше зниження втрат та незначне підвищення точності.

5) П'ятий експеримент: Збільшено кількість ітерацій до 10000, що, на жаль, призвело до перенавчання моделі через недостатню кількість даних.

Для підсумкової оцінки результатів експериментів створена зведена таблиця 4.1.

Таблиця 4.1

Результуюча таблиця експериментів

Експеримент	Ітерації на епоху	Розмір зображення	Точність (%)	Втрати (%)
1-й	1000	64x64	82	52
2-й	2000	64x64	85	45
3-й	5000	64x64	87	37
4-й	5000	128x128	87.5	35.5
5-й	10000	128x128	Перенавчання	Піднялися до 100%

За результатами експериментів, можна зробити висновок, що збільшення кількості ітерацій на епоху сприяє підвищенню точності моделі, але існує межа, після якої модель може перенавчатися. Також зміна розміру вхідних зображень на більший розмір допомагає знизити втрати і злегка підвищити точність, проте це вимагає більше обчислювальних ресурсів та часу на навчання.

4.3 Тестування роботи застосунку

Застосунок має графічний інтерфейс користувача, розроблений за допомогою бібліотеки Tkinter, що включає кнопки для завантаження зображень, приклад роботи завантаження зображень на (рис. 4.9) та управління автоматичним захопленням зображень з камери (рис. 4.10). Інтерфейс відображає зображення та результати класифікації, включаючи назву класу та ймовірність (рис. 4.11).

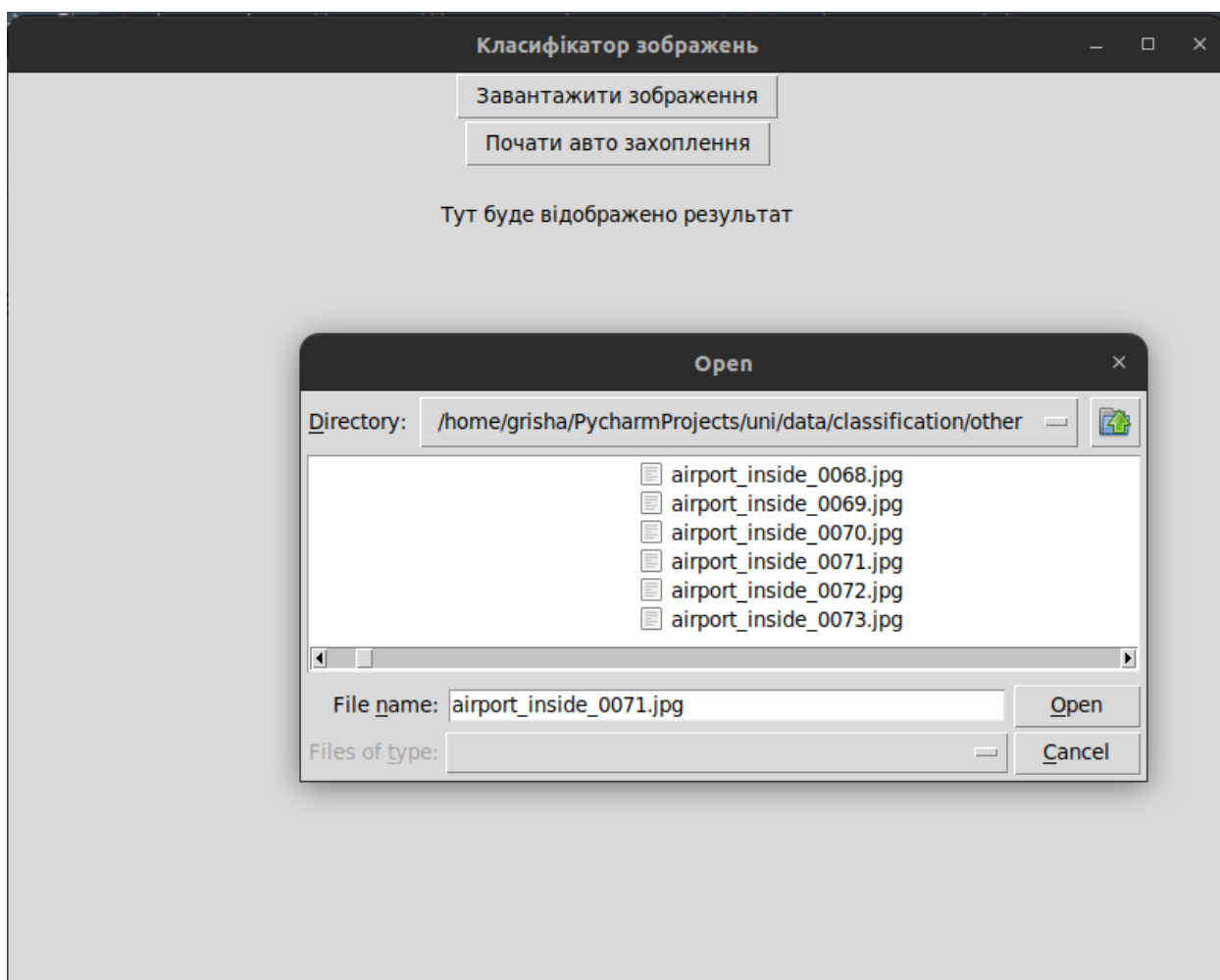


Рисунок 4.9 Інтерфейс завантаження зображення

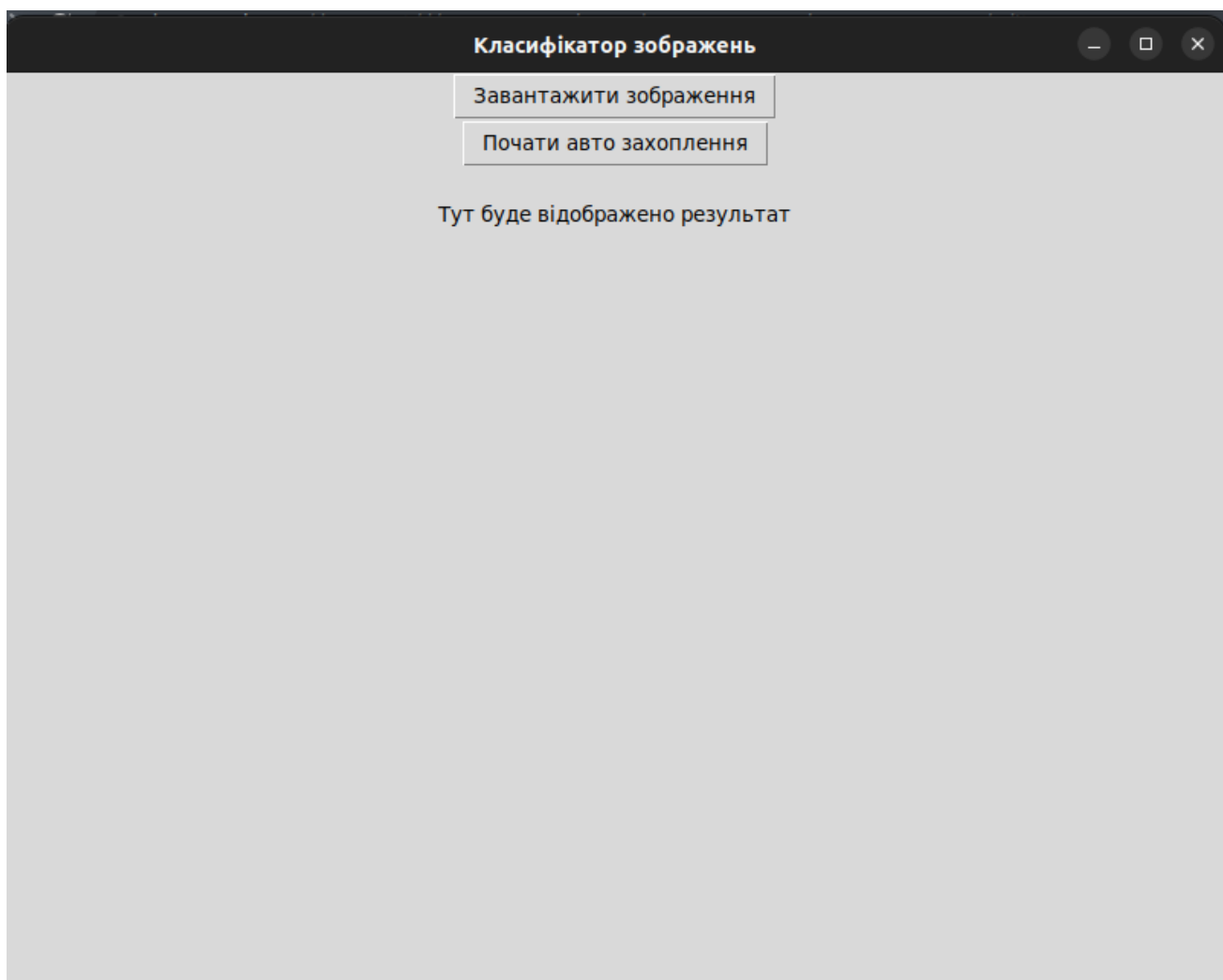


Рисунок 4.10 Інтерфейс додатку

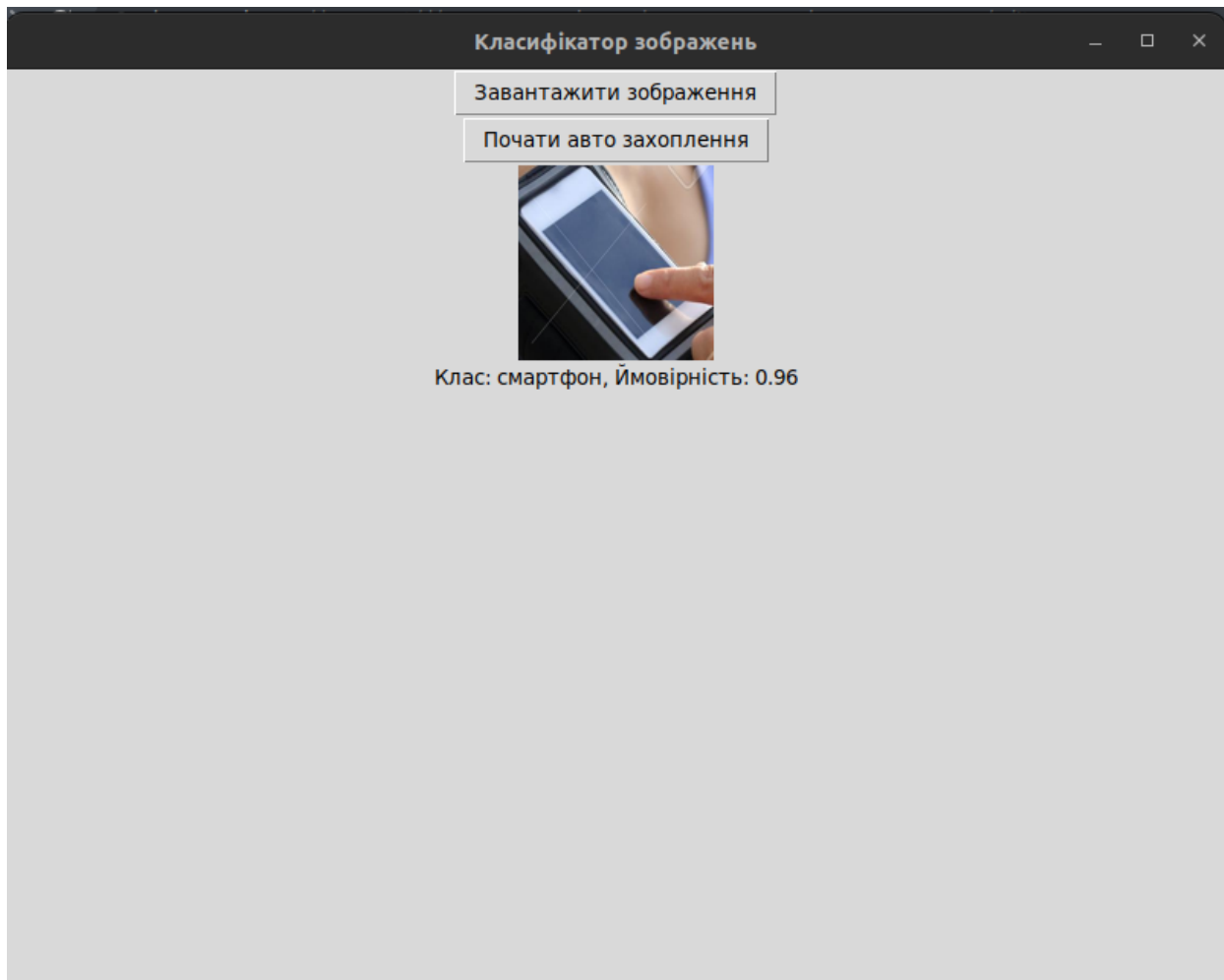


Рисунок 4.11 Приклад роботи класифікації у застосунку

Застосунок використовує навчену модель нейронної мережі для класифікації зображень. Користувач може завантажити зображення або використовувати камеру для захоплення зображень 4.12, після натискання на дану кнопку кожні 3 секунди робиться фотографія з підключеної до системи камери. Після завантаження, зображення обробляється, нормалізується, і подається на вхід моделі. Модель повертає ймовірність приналежності зображення до одного з класів.

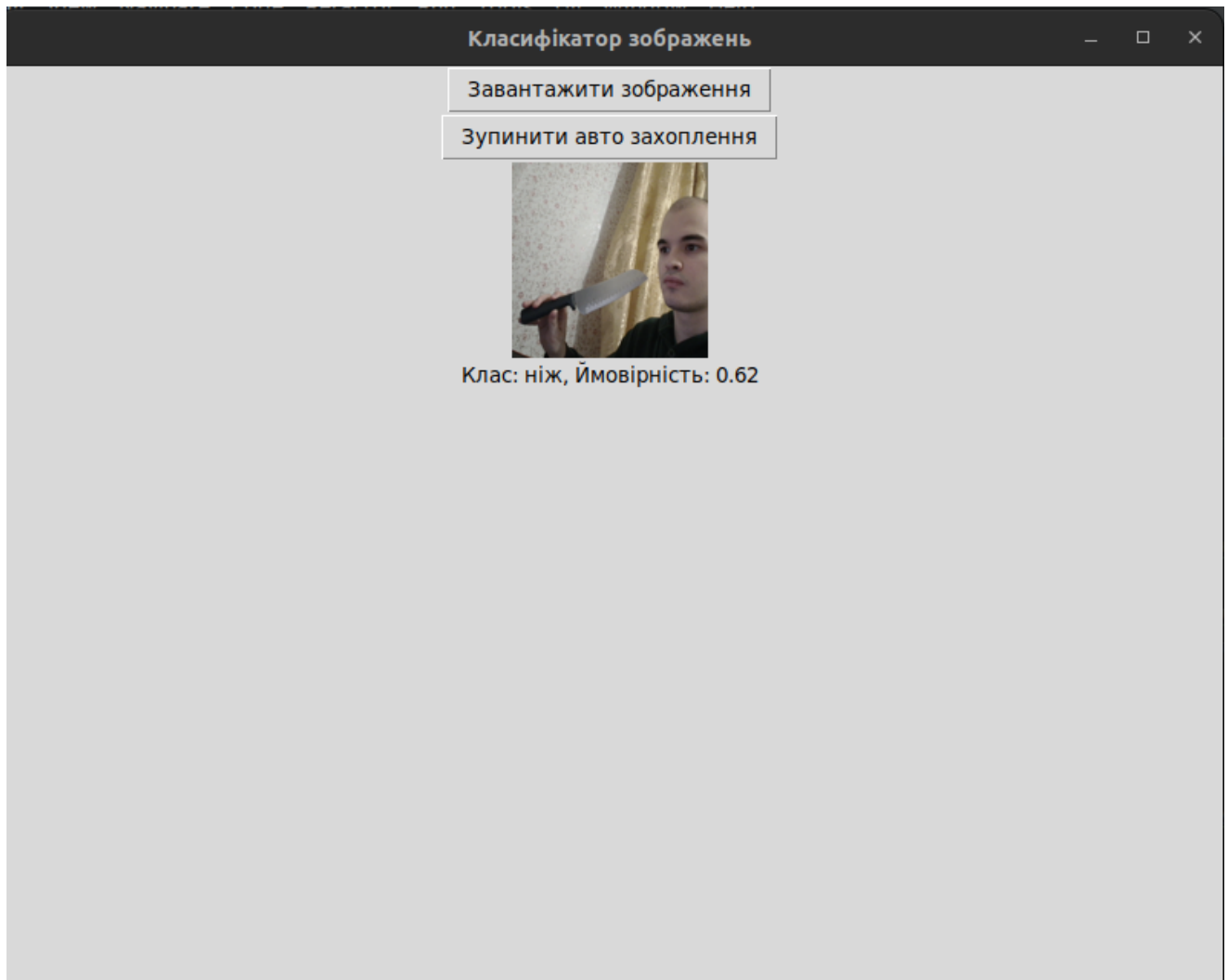


Рисунок 4.12 Робота автозахоплення зображень з камери та класифікації у застосунку

4.5 Аналіз порівняння з аналогами

В цьому підрозділі було проведено порівняння розроблюваного додатка на основі навченної моделі з аналогами. Також його було додано у таблицю 4.2 з аналогами з другого розділу.

Таблиця 4.2

Порівняльна таблиця з існуючих рішень з розробленим додатком

Система	Час відгуку системи	Масштабованість	Підтримка різних типів даних	Ціноутворення	Приблизна Точність Розпізнавання (%)
Google Cloud Vision API	Швидкий	Швидкий	Зображення, текст, обличчя, логотипи	Платно	95%
Amazon Rekognition	Швидкий	Висока	Обличчя, текст, небезпечний контент, логотипи, емоції	Платно	90%
Clarifai	Швидкий	Висока	Зображення, відео, текст, аудіо	Платно	92%
Anyline	Швидкий	Висока	OCR, штрих-коди, метричні дані, ID-документи, VIN номери	Платно	88%
Додаток з розробленою моделлю	Швидкий	Середня	Зображення	Безкоштовна	87%

Google Cloud Vision API вирізняється високою точністю в розпізнаванні об'єктів, тексту та емоцій, а також зручною інтеграцією з іншими сервісами

Google Cloud. Однак, його використання може бути відносно дорогим, а також вимагає інтеграції в екосистему Google Cloud.

Amazon Rekognition ефективно інтегрується з AWS і ідеально підходить для проектів великого масштабу, особливо в галузях безпеки та нагляду. Проте, ця система може бути менш гнучкою для деяких специфічних задач.

Clarifai надає високу точність в детальному аналізі зображень і пропонує налаштовані та модульні рішення. Незважаючи на це, вона вимагає більше технічних знань для ефективного використання.

Anyline спеціалізується на мобільних застосунках і вирізняється легкістю інтеграції. Однак, вона має обмежений функціонал і є менш універсальною порівняно з більш широкими рішеннями розпізнавання зображень.

Розроблена в даній роботі модель демонструє точність до 87%, що є дуже конкурентоспроможним показником, з урахуванням, що втрати складають до 37%. Однією з ключових переваг є здатність моделі ефективно навчатися на малих обсягах вхідних даних, що є важливим аспектом для проектів з обмеженими датасетами. Крім того, процес навчання моделі не є ресурсозатратним, що робить її доступною для використання на менш потужному обладнанні.

На основі проведеного порівняльного аналізу можна зробити висновок, що моя розроблена модель є конкурентоспроможною на ринку рішень розпізнавання зображень. Вона вирізняється здатністю до навчання на малих наборах даних і ефективністю в умовах обмежених ресурсів, що робить її цінною для проектів з обмеженими бюджетами та обладнанням. Однак, слід врахувати, що точність моделі трохи нижча порівняно з провідними рішеннями на ринку, що може бути предметом подальшого дослідження та оптимізації.

4.6 Висновки до четвертого розділу

У четвертому розділі було представлено детальний аналіз та опис роботи розробленого програмного застосунку, заснованого на моделі глибокого

навчання для класифікації об'єктів. Основна увага у цьому розділі зосереджена на демонстрації функціональності та ефективності застосунку, а також на висвітленні ключових аспектів його розробки та тестування. Використання глибокої нейронної мережі, розробленої з використанням TensorFlow та Keras, є фундаментальним для функціонування застосунку, забезпечуючи високу точність та надійність у класифікації об'єктів.

Було розглянуто процес коригування під час навчання, включаючи зміну гіперпараметрів, таких як `'steps_per_epoch'` і `'batch_size'` та збільшення розміру вхідних зображень. Ці зміни дозволили значно підвищити точність моделі та зменшити втрати. Додатково, було надано інформацію про графічний інтерфейс користувача, розроблений з використанням бібліотеки Tkinter, який забезпечує простоту та зручність взаємодії з користувачем.

Також у цьому розділі проведено порівняльний аналіз розробленого додатка з іншими аналогічними рішеннями, що дозволило виявити його переваги та недоліки в контексті ринкових альтернатив. Встановлено, що розроблений застосунок, хоч і має трохи нижчу точність у порівнянні з провідними рішеннями, проте вирізняється своєю гнучкістю, невибагливістю до обчислювальних ресурсів та здатністю ефективно навчатися на малих наборах даних.

З цього можна зробити висновок, що розроблений застосунок є конкурентоспроможним рішенням у галузі розпізнавання зображень, особливо для проектів з обмеженими бюджетами та технічними ресурсами. Його гнучкість, висока точність та зручний графічний інтерфейс роблять його цінним інструментом для різноманітних застосувань у цій області.

Розділ 5. Охорона навколишнього середовища

Екологічні рухи, які зародилися у 19-му столітті та досягли зрілості у 20-му та 21-му століттях, відіграють ключову роль у формуванні глобальної екологічної свідомості. У цій історичній ретроспективі ми досліджуємо їхній розвиток, філософію та вплив на суспільство. Ранні екологічні рухи виникли в основному у відповідь на промислову революцію, яка кардинально змінила взаємодію людини з природою. Це був період, коли почала формуватися екологічна свідомість, але більша частина діяльності була фокусована на збереженні окремих природних територій та видів.

Протягом 1960-1970-х років відбувся значний розвиток у сфері екологічної свідомості, зумовлений глобальними екологічними кризами та розумінням глибоких взаємозв'язків між людством та природою. Це був час, коли екологія перетворилася з наукової дисципліни в широку суспільну та політичну течію. Сучасні тенденції екологічних рухів характеризуються глобалізацією, технологічними інноваціями та зростаючою міжнародною співпрацею.

Філософія та ідеологія екологічних рухів різноманітні та еволюціонували з часом. Екологічний гуманізм, який акцентує на гармонії між людиною та природою, підкреслює важливість екологічної обізнаності для досягнення сталого розвитку. Глибинна екологія йде ще далі, пропонуючи фундаментальний перегляд ролі людини в екосистемі, де людина розглядається як частина більш широкої живої системи.

Екологічний реалізм та радикалізм відображають різні підходи до вирішення екологічних проблем. Екологічний реалізм прагне до розробки практичних, науково обґрунтованих рішень, тоді як радикальні напрямки можуть закликати до докорінних змін у суспільних та економічних системах.

Основними напрямками та цілями екологічних рухів є захист природних ресурсів, боротьба з кліматичними змінами та просування сталого розвитку. Ці

цілі досягаються через різноманітні методи, від освітніх кампаній до прямих дій та політичного лобіювання.

Впливові екологічні організації, такі як Greenpeace, WWF, та інші, мають значний вплив на глобальному рівні. Їхні кампанії часто зосереджені на конкретних питаннях, таких як збереження видів, захист тропічних лісів, зменшення викидів парникових газів, та розвиток відновлюваних джерел енергії. Ці організації також грають ключову роль у формуванні міжнародної екологічної політики та стандартів.

В цілому, екологічні рухи продовжують бути важливою силою у формуванні глобальної екологічної політики та свідомості. Їхня робота адаптується до сучасних викликів, таких як кліматична криза та загроза біорізноманіттю, що вимагає постійного переосмислення та інновацій у підходах до захисту навколишнього середовища.

Впливові екологічні організації, такі як Greenpeace, World Wildlife Fund (WWF) та Extinction Rebellion, стали значними гравцями у глобальних зусиллях з охорони навколишнього середовища. Їхні кампанії відіграють важливу роль у формуванні громадської думки, впливі на політику та залученні уваги до критичних екологічних питань. Greenpeace, заснований у 1971 році, знаний своєю прямою дією та креативними кампаніями. WWF, заснований у 1961 році, концентрується на збереженні природи та вплив на політику в області охорони дикої природи. Extinction Rebellion, новіша організація, яка з'явилася в 2018 році, використовує тактику непокори та масових протестів для привернення уваги до кліматичної кризи.

Сучасні екологічні рухи стикаються з низкою викликів та проблем. Політичний та економічний тиск часто впливає на їхню діяльність, оскільки уряди та корпорації можуть виступати проти екологічних ініціатив, які загрожують існуючим інтересам. Проблеми фінансування та ресурсів також є значними, особливо для менших та менш відомих організацій. Зв'язок із

соціальною справедливістю стає все важливішим, оскільки багато екологічних проблем тісно переплетені з питаннями нерівності та прав людини.

Вплив екологічних рухів можна спостерігати в ряді областей. Законодавчі зміни та політичні ініціативи, ініційовані екологічними рухами, сприяли встановленню нових стандартів у сфері охорони навколишнього середовища. Громадська свідомість та зміна поведінки є одним із найбільших досягнень цих рухів, оскільки вони спонукали мільйони людей переглянути свої звички та вибір. Міжнародна співпраця та угоди, такі як Паризька кліматична угода, також були посилені активною роботою екологічних організацій.

Майбутнє екологічних рухів здається багатообіцяючим, але водночас непередбачуваним. Нові тенденції та стратегії з'являються у відповідь на змінювані екологічні, політичні та соціальні ландшафти. Залучення молоді та інноваційні підходи стають все більш важливими, оскільки молоде покоління займає активну позицію у питаннях захисту навколишнього середовища. Глобальні виклики, такі як кліматична криза, втрата біорізноманіття та забруднення, продовжують вимагати широкомасштабних зусиль та інноваційних рішень.

Екологічні рухи продовжують розвиватися, адаптуючись до нових викликів та можливостей. Вони стали важливою частиною глобального діалогу про майбутнє нашої планети та її жителів. Їхня робота не тільки сприяє захисту природи та зменшенню впливу на навколишнє середовище, але й стимулює політичні, економічні та соціальні зміни, які є життєво необхідними для сталого майбутнього.

ВИСНОВКИ

У цій роботі було зосереджено увагу на проблематиці виявлення недозволених предметів на підприємствах і важливість її вирішення за допомогою передових технологій. Було встановлено, що дана проблема відіграє ключову роль у забезпеченні безпеки працівників підприємств та може суттєво впливати на репутацію організацій.

В результаті виконання кваліфікаційної роботи отримані наступні результати:

- Проаналізовано сучасні методи цифрової обробки зображень та алгоритми розпізнавання об'єктів, особливу увагу було приділено методам з використанням глибокого навчання та на основі результату проведеного аналізу було обрано методи машинного навчання для вирішення задачі виявлення недозволених предметів.
- Розроблено програмний застосунок виявлення недозволених предметів на підприємстві, в якому інтегровано згорткову нейронну мережу (CNN) та технології обробки зображень, що дало змогу ефективно виявляти та класифікувати об'єкти в різних умовах.
- Протестовано розроблений програмний застосунок, що дало змогу дослідити доцільність використання розробленого застосунку для вирішення поставленої задачі, з акцентом на точність виявлення та швидкість обробки даних.

В цілому, дана робота підкреслює значення інтеграції сучасних технологій обробки зображень та глибокого навчання у системи забезпечення безпеки. Результати дослідження та розробки демонструють, що такий підхід може істотно підвищити ефективність виявлення недозволених предметів, відкриваючи можливості для покращення безпеки у важливих секторах.

Список використаних джерел:

1. Accidents at work statistics [Електронний ресурс] — Режим доступу до ресурсу: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Accidents_at_work_statistics#:~:text=There%20were%203%20347%20fatal,compared%20with%20the%20year%20before.&text=In%202021%2C%20more%20than%20a,place%20with%20the%20construction%20sector.

2. Analysis of National Major Work Safety Accidents in China [Електронний ресурс] — Режим доступу до ресурсу: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4822395/#:~:text=A%20total%20of%20877%20major,have%20declined%20in%20recent%20years.>

3. Castleman KR (1996) Digital image processing. Prentice-Hall, Upper Saddle River, NJ Sawyer LC, Grubb DT (1996) Polymer microscopy, 2nd edn. Chapman and Hall, London Kirkland EJ (1998) Advanced computing in electron microscopy. / NJ Sawyer LC, Grubb DT / [Електронний ресурс] — Режим доступу до ресурсу: https://link.springer.com/chapter/10.1007/978-3-540-36352-1_8

4. Image Analysis and Processing — Python OpenCV Example. [Електронний ресурс] — Режим доступу до ресурсу: <https://medium.com/elucidate-ai/image-analysis-and-processing-python-opencv-example-c09fcf929d8>

5. Image Transformations. [Електронний ресурс] — Режим доступу до ресурсу: https://www.tutorialspoint.com/dip/image_transformations.htm

6. Analysis of Various Image Feature Extraction Methods against Noisy Image: SIFT, SURF and HOG. [Електронний ресурс] — Режим доступу до ресурсу: <https://jespublication.com/upload/2019-V10I22019-2.pdf>

7. Introduction to Convolutional Neural Networks. [Електронний ресурс] — Режим доступу до ресурсу:

<https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/>

8. LeCun, Yann. LeNet-5, convolutional neural networks. /LeCun, Yann/ [Электронный ресурс] — Режим доступа до ресурсу:

<https://web.archive.org/web/20210224225707/http://yann.lecun.com/exdb/lenet/>

9. Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda (2003). Subject independent facial expression recognition with robust face detection using a convolutional neural network. / Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda / [Электронный ресурс] — Режим доступа до ресурсу:

https://web.archive.org/web/20131213022740/http://www.iro.umontreal.ca/~pift6080/H09/documents/papers/sparse/matsugo_et_al_face_expression_conv_nnet.pdf

10. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie & Laith Farhan / Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie & Laith Farhan / [Электронный ресурс] — Режим доступа до ресурсу:

<https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00444-8>

11. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. / Mohammad Mustafa Taye / [Электронный ресурс] — Режим доступа до ресурсу:

<https://www.mdpi.com/2079-3197/11/3/52>

12. Understanding of Convolutional Neural Network (CNN): A Review. / Purwono, Alfian Ma'arif, Wahyu Rahmani, Haris Imam Karim Fathurrahman, Aufaclav Zatu Kusuma Frisky, Qazi Mazhar ul Haq / [Электронный ресурс] — Режим доступа до ресурсу:

https://www.researchgate.net/publication/367157330_Understanding_of_Convolutional_Neural_Network_CNN_A_Review

13. Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. / Sakshi Indolia, Anil Kumar Goswami, S.P. Mishra, Pooja Asora / [Электронный ресурс] — Режим доступа до ресурсу:

<https://www.sciencedirect.com/science/article/pii/S1877050918308019>

14. Recent Advances in Convolutional Neural Networks / Jiuxiang Gu, Zhenhua Wangb, Jason Kuen, Lianyang Ma, Amir Shahroudy , Bing Shuai , Ting Liu , Xingxing Wang , Li Wang , Gang Wang , Jianfei Cai , Tsuhan Chen / [Электронный ресурс] — Режим доступа до ресурсу:

<https://arxiv.org/pdf/1512.07108.pdf%C3%A3%E2%82%AC%E2%80%9A>

15. Deep learning: basics and convolutional neural networks (CNN) / Maria Vakalopoulou, Stergios Christodoulidis, Ninon Burgos, Olivier Colliot, Vincent Lepetit / [Электронный ресурс] — Режим доступа до ресурсу:

<https://hal.science/hal-03957224/document>

Код програми

```
import os
import numpy as np
import glob
import tensorflow as tf
from tensorflow.python.keras.models import Sequential
from tensorflow.python.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense,
Dropout
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt

# Визначення класів
classes = ['knife', 'pistol', 'smartphone', 'boxes']
num_classes = len(classes) # Кількість класів предметів

# Розміри вхідного зображення
input_shape = (64, 64)
# Кількість каналів у зображенні (RGB)
num_channels = 3
# Кількість кроків на епоху під час навчання
steps_per_epoch = 5000

# Функція для обробки зображення для декількох класів
def process_image(image_path, target_size):
    try:
        # Завантаження та трансформація зображення
```

```

image = tf.keras.utils.load_img(image_path, target_size=target_size)
image = tf.keras.utils.img_to_array(image)
# Нормалізація пікселів
image = image / 255.0
return image
except Exception as e:
    print(f"Помилка при обробці зображення {image_path}: {e}")
# Повернення порожнього зображення у випадку помилки
return tf.zeros([target_size[0], target_size[1], num_channels])

# Функція для завантаження набору даних з директорії
def load_data(input_shape, num_channels):
    images = []
    labels = []

    # Завантаження зображень для кожного класу
    for idx, class_name in enumerate(classes):
        directory = f"data/classification/{class_name}"
        for file in glob.glob(os.path.join(directory, '*.*jpg')) +
glob.glob(os.path.join(directory, '*.*JPG')):
            image = process_image(file, target_size=input_shape)
            if tf.reduce_sum(image) > 0:
                images.append(image)
                labels.append(idx)

    # Завантаження негативних зразків
    directory = "data/classification/other"

```

```

for file in glob.glob(os.path.join(directory, '*.jpg')) +
glob.glob(os.path.join(directory, '*.JPG')):
    image = process_image(file, target_size=input_shape)
    if tf.reduce_sum(image) > 0:
        images.append(image)
        labels.append(num_classes) # Унікальний мітка для негативних зразків

return np.array(images), tf.keras.utils.to_categorical(labels, num_classes + 1)

```

```
# Завантаження даних
```

```
X, y = load_data(input_shape, num_channels)
```

```
# Розділення даних на навчальний та тестувальний набори
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
```

```
# Архітектура моделі для декількох класів
```

```

model = Sequential([
    Conv2D(32, (3, 3), activation='relu', input_shape=(input_shape[0], input_shape[1],
num_channels)),
    MaxPooling2D(pool_size=(2, 2)),
    Conv2D(64, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Flatten(),
    Dense(128, activation='relu'),
    Dropout(0.5),
    Dense(num_classes + 1, activation='softmax')
])

```

```
# Компіляція моделі
model.compile(loss='categorical_crossentropy', optimizer='adam',
metrics=['accuracy'])

# Параметри збільшення даних
datagen = tf.keras.preprocessing.image.ImageDataGenerator(
    rotation_range=20,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True,
    fill_mode='nearest')

# Навчання моделі
history = model.fit(datagen.flow(X_train, y_train, batch_size=64),
    epochs=10,
    steps_per_epoch=steps_per_epoch,
    validation_data=(X_test, y_test))

# Оцінка моделі
test_loss, test_acc = model.evaluate(X_test, y_test)
print(f"Точність на тестових даних: {test_acc:.3f}, Втрати на тестових даних:
{test_loss:.3f}")

# Збереження моделі
model.save("model/multiclass_classification_model_3.h1")
```

```
# Візуалізація результатів навчання
# Візуалізація втрат
plt.figure(figsize=(12, 4))
plt.subplot(1, 2, 1)
plt.plot(history.history['loss'], label='Втрати при навчанні')
plt.plot(history.history['val_loss'], label='Втрати на валідації')
plt.title('Втрати протягом епох')
plt.xlabel('Епохи')
plt.ylabel('Втрати')
plt.legend()

# Візуалізація точності
plt.subplot(1, 2, 2)
plt.plot(history.history['accuracy'], label='Точність при навчанні')
plt.plot(history.history['val_accuracy'], label='Точність на валідації')
plt.title('Точність протягом епох')
plt.xlabel('Епохи')
plt.ylabel('Точність')
plt.legend()

plt.show()
import os
import numpy as np
import tensorflow as tf
from tensorflow.python.keras.models import load_model
import glob

# Функція для обробки зображень
def process_image(image_path, target_size):
```

```

try:
    # Завантаження та перетворення зображення
    image = tf.keras.utils.load_img(image_path, target_size=target_size)
    image = tf.keras.utils.img_to_array(image)
    # Нормалізація пікселів
    image = image / 255.0
    return image
except Exception as e:
    # Повідомлення про помилку обробки зображення
    print(f"Помилка обробки зображення {image_path}: {e}")
    # Повернення порожнього зображення у разі помилки
    return tf.zeros([target_size[0], target_size[1], num_channels])

# Функція для завантаження даних
def load_data(input_shape, num_channels):
    images = []
    labels = []
    # Визначення класів
    classes = ['knife', 'pistol', 'smartphone', 'boxes', 'other'] # Включно з 'other' для
негативних зразків

    # Завантаження зображень для кожного класу
    for idx, class_name in enumerate(classes):
        directory = f"data/classification/{class_name}"
        for file in glob.glob(os.path.join(directory, '*.jpg')) +
glob.glob(os.path.join(directory, '*.JPG')):
            image = process_image(file, target_size=input_shape)
            if tf.reduce_sum(image) > 0:
                images.append(image)

```

```
        labels.append(idx)

    return np.array(images), tf.keras.utils.to_categorical(labels, len(classes))

# Завантаження моделі
model = load_model("model/multiclass_classification_model_1.h1")

# Встановлення параметрів вхідного зображення
input_shape = (64, 64)
num_channels = 3

# Завантаження тестових даних
X_test, y_test = load_data(input_shape, num_channels)

# Оцінка моделі
test_loss, test_acc = model.evaluate(X_test, y_test)

# Виведення результатів тестування
print(f"Точність на тестових даних: {test_acc:.3f}, Втрати на тестових даних:
{test_loss:.3f}")

import tkinter as tk
from tkinter import filedialog
from PIL import Image, ImageTk
import numpy as np
import tensorflow as tf
import cv2

# Завантаження моделі
model = tf.keras.models.load_model('model/multiclass_classification_model_4.h1')

# Назви класів
```

```
classes = ['ніж', 'пістолет', 'смартфон', 'коробки', 'інше']

# Флаг для відстеження стану автоматичного захоплення зображень
capturing = False

def upload_image():
    # Функція для завантаження зображення з файлу
    file_path = filedialog.askopenfilename()
    img = Image.open(file_path)
    classify_image(img)

def capture_image():
    # Функція для захоплення зображення з камери
    global capturing
    if capturing:
        cap = cv2.VideoCapture(0) # Використовуйте 0 або індекс камери
        ret, frame = cap.read()
        if ret:
            # Конвертуємо зображення в формат PIL
            cv_img = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)
            img = Image.fromarray(cv_img)
            classify_image(img)
        cap.release()
    window.after(3000, capture_image) # Запускаємо функцію знову через 3
секунди
```



```
def classify_image(img):  
    # Функція класифікації зображення  
    if img.mode != 'RGB':  
        img = img.convert('RGB')  
  
    img = img.resize((128, 128))  
    img_array = np.array(img) / 255.0  
    img_array = img_array.reshape((1, 128, 128, 3))  
  
    # Відображення зображення  
    img = ImageTk.PhotoImage(img)  
    image_panel.config(image=img)  
    image_panel.image = img  
  
    # Класифікація зображення  
    predictions = model.predict(img_array)  
    predicted_class = np.argmax(predictions, axis=1)  
    predicted_probability = np.max(predictions)  
  
    # Відображення назви класу та ймовірності  
    class_name = classes[predicted_class[0]]  
    probability_text = f'Клас: {class_name}, Ймовірність: {predicted_probability:.2f}'  
    label.config(text=probability_text)  
  
def toggle_auto_capture():  
    # Функція для управління автоматичним захопленням зображень  
    global capturing  
    capturing = not capturing
```

```
if capturing:
    capture_button.config(text='Зупинити авто захоплення')
    capture_image()
else:
    capture_button.config(text='Почати авто захоплення')
```

```
# Налаштування GUI
```

```
window = tk.Tk()
window.title('Класифікатор зображень')
window.geometry('800x600')
```

```
# Кнопка для завантаження зображення
```

```
open_button = tk.Button(window, text='Завантажити зображення',
command=upload_image)
open_button.pack()
```

```
# Кнопка для управління автоматичним захопленням
```

```
capture_button = tk.Button(window, text='Почати авто захоплення',
command=toggle_auto_capture)
capture_button.pack()
```

```
# Створення Label для зображення
```

```
image_panel = tk.Label(window)
image_panel.pack()
```

```
label = tk.Label(window, text='Тут буде відображено результат')
label.pack()
```

```
# Запуск програми  
window.mainloop()
```