



КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

на тему Метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання

Галузь знань 12 – Інформаційні технології
Шифр і назва галузі знань
Спеціальність 122 – Комп'ютерні науки
Шифр і назва спеціальності
Освітня програма Комп'ютерні науки
Назва освітньої програми

Виконала: студентка групи КН-21-1  Ольга ШЕБЕТКО
Група виконавця Підпис Ім'я, ПРІЗВИЩЕ

Керівник: д.т.н., зав. кафедри КН  Олександр БАРМАК
Науковий ступінь, посада Підпис Ім'я, ПРІЗВИЩЕ

Нормоконтроль: к.т.н., доц. каф. КН  Руслан БАГРІЙ
Науковий ступінь, посада Підпис Ім'я, ПРІЗВИЩЕ

До захисту допускаю:
зав. кафедри КН, д.т.н., професор  Олександр БАРМАК
Підпис Ім'я, ПРІЗВИЩЕ

09 06 2025 р.

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет інформаційних технологій

Кафедра комп'ютерних наук

Освітній ступінь бакалавр

Галузь знань 12 – Інформаційні технології

Спеціальність 122 – Комп'ютерні науки

ЗАТВЕРДЖУЮ

Завідувач кафедри комп'ютерних наук


(підпис)

д.т.н., професор Олександр БАРМАК

« 10 » 02 2025 року

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ БАКАЛАВРА**

1. Тема кваліфікаційної роботи бакалавра: «Метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання»

2. Завдання видано студентці Ользі ШЕБЕТКО

(Ім'я, прізвище)

3. Керівник роботи зав. каф. КН, Олександр БАРМАК

(посада, ім'я, прізвище)

4. Затверджено наказом університету від « 07 » 02 2025 р. № 23

5. Дата видачі завдання студенту: « 10 » 02 2025 р.

6. Зміст пояснювальної записки (перелік задач) та вихідні дані:

Мета роботи – допомога у підборі друкованої інформації через аналіз зображень їх обкладинок за допомогою відповідного методу штучного інтелекту імплементованого у мобільне застосування. Метод має здійснювати розпізнавання типу, жанру та інших характеристик друкованих видань, а також формувати рекомендації щодо вибору літератури на основі рівня володіння англійською мовою користувача. Відповідний мобільний застосунок, у який імплементовано зазначений метод, повинен забезпечити функції отримання, обробки, аналізу зображень обкладинок, а також надати зручний інтерфейс для взаємодії з користувачем.

7. Календарний план виконання кваліфікаційної роботи бакалавра:

№	Назва етапів (розділів) кваліфікаційної роботи бакалавра	Термін виконання	Примітка
1	Вибір напрямку дослідження та узгодження тематики кваліфікаційної роботи бакалавра з керівником, складання календарного графіка виконання роботи	січень 2025	
2	Ознайомлення з предметною областю, формулювання мети та задач дослідження, визначення об'єкта та предмета дослідження	лютий 2025	
3	Робота над розділом 1 «Характеристика предметної області: аналіз моделей, методів та реалізацій»	березень 2025	
4	Робота над розділом 2 «Метод автоматичної ідентифікації друкованих видань та особливості його реалізації»	квітень 2025	
5	Робота над розділом 3 «Експериментальне дослідження програмної реалізації запропонованого методу»	травень 2025	
6	Розробка презентаційних матеріалів та попередній захист кваліфікаційної роботи	травень 2025	
7	Отримання відгуку керівника, рецензії, перевірка на плагіат, нормоконтроль	червень 2025	
8	Підготовка до захисту та захист кваліфікаційної роботи бакалавра	червень 2025	

Виконавець: студентка групи КН-21-1
Група виконавця


Підпис

Ольга ШЕБЕТКО
Ім'я, ПРІЗВИЩЕ

Керівник: д.т.н., зав. кафедри КН
Науковий ступінь, посада


Підпис

Олександр БАРМАК
Ім'я, ПРІЗВИЩЕ

Анотація

Тема кваліфікаційної роботи бакалавра: «Метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання»

Виконавець кваліфікаційної роботи бакалавра: студентка групи КН-21-1 Ольга ШЕБЕТКО

Керівник кваліфікаційної роботи бакалавра: д.т.н., проф. Олександр БАРМАК

Кваліфікаційна робота бакалавра містить:

Пояснювальна записка				Кількість додатків
Сторінок	Рисунків	Таблиць	Джерел інформації	
54	25	3	53	2

Метою кваліфікаційної роботи бакалавра є допомога у підборі друкованої інформації через аналіз зображень їх обкладинок за допомогою відповідного методу штучного інтелекту імplementованого у мобільне застосування.

У межах роботи розроблено та реалізовано метод ідентифікації друкованої продукції, який дозволяє визначати жанр видання, назву, автора та інші релевантні характеристики на основі візуального вигляду обкладинки. Застосунок поєднує інструменти комп'ютерного зору, OCR і нейронних мереж для точного розпізнавання та класифікації видань. Практичне значення полягає у зручності автоматичного підбору літератури та наданні персоналізованих рекомендацій, що може бути корисним у бібліотеках, освітніх сервісах або особистому користуванні. Застосунок має інтуїтивний інтерфейс і функціонал, достатній для повноцінної взаємодії користувача з системою.

Ключові слова: комп'ютерний зір, глибоке навчання, обкладинка, книга, жанр, рекомендації.

Виконавець: студентка групи КН-21-1
Група виконавця


Підпис

Ольга ШЕБЕТКО
Ім'я, ПРІЗВИЩЕ

Зміст

Перелік скорочень	3
Вступ.....	4
Розділ 1 Характеристика предметної області: аналіз моделей, методів та реалізацій.....	6
1.1 Аналіз інформаційних моделей.....	6
1.2 Огляд підходів до розв’язку подібних задач.....	8
1.3 Аналіз існуючих програмних засобів та наукових рішень.....	10
1.4 Мета, задачі та вимоги до розробки та реалізації.....	13
Розділ 2 Метод автоматичної ідентифікації друкованих видань та особливості його реалізації.....	15
2.1 Метод ідентифікації друкованих видань за зображенням їх обкладинки ...	15
2.1.1. Ідея методу.....	15
2.1.2 Модель нейромережі	15
2.1.3. Кроки методу.....	17
2.2. Особливості реалізації запропонованого методу	20
2.3. Проектування інформаційної системи для ідентифікації зображень обкладинки друкованих видань.....	23
2.4 Висновки до розділу 2	29
Розділ 3 Експериментальне дослідження програмної реалізації запропонованого методу.....	31
3.1 Опис інтерфейсу користувача.....	31
3.2 Результати експериментальних досліджень.....	37
3.3 Висновки до розділу 3	44
Загальні висновки.....	47
Перелік посилань.....	50
Додатки	

Перелік скорочень

Скорочення, термін, позначення	Пояснення
ІС	Інформаційна система
ІТ	Інформаційні технології
КРБ	Кваліфікаційна робота бакалавра
КН	Комп'ютерні науки
НМ	Нейронна мережа
ПП	Програмний продукт
ХНУ	Хмельницький національний університет.
RGB	Red, Green, Blue
СМЯК	Сяан, Magenta, Yellow, Key color
CNN	Convolutional Neural Networks
SVM	Support Vector Machine
MPT	Магнітно-резонансна томографія
DNN	Deep Neural Network
API	Application Programming Interface
CSV	Comma-Separated Values
CPU	Central Processing Unit
GPU	Graphics Processing Unit
UML	Unified Modeling Language
IOS	iPhone, iPod, iPad Operating System
CEFR	Common European Framework of Reference
OCR	Optical Character Recognition

Вступ

Актуальність. Розвиток інформаційних технологій є надзвичайно стрімким. Використання мобільних пристроїв зростає, вони здатні допомагати користувачам у вирішенні щоденних задач, зокрема в сфері підбору друкованих матеріалів. Сучасний ринок книг, журналів та інших друкованих видань постійно збільшується, що створює потребу в ефективних інструментах для допомоги у виборі відповідних матеріалів серед великої кількості доступної інформації. Однак, у багатьох випадках користувачам складно знайти літературу, яка відповідає їхнім інтересам та вимогам через величезну різноманітність видань.

Технології комп'ютерного зору та глибокого навчання дозволяють автоматизувати процес ідентифікації та класифікації друкованих матеріалів за зображенням обкладинок, що є дуже актуальним для користувачів, які бажають швидко та ефективно знайти потрібну книгу чи підручник. Завдяки розпізнаванню обкладинок можна отримати не лише інформацію про жанр, автора, але й зрозуміти, чи відповідає видання потребам користувача.

Тому виникає необхідність у створенні методу ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання, який здатний аналізувати зображення обкладинок та пропонувати користувачам персоналізовані рекомендації. З розвитком мобільних технологій такі рішення стають доступними для широкого кола користувачів, адже вони не потребують додаткового спеціалізованого обладнання.

Крім того, завдяки використанню методів глибокого навчання, можна досягти високої точності в розпізнаванні та класифікації різних типів друкованих видань, що дозволяє створювати зручні та ефективні інтерфейси для кінцевих користувачів.

Такий метод штучного інтелекту допоможе розробити універсальний інструментом для всіх, хто шукає конкретні книги чи підручники, а також для тих, хто бажає дізнатися більше про нові видання, не витрачаючи час на пошук серед численних варіантів.

Об'єкт дослідження – процес аналізу зображень для виявлення та декодування текстової інформації.

Предмет дослідження – методи та технології глибокого навчання для виявлення та декодування тексту на зображеннях за допомогою глибокого навчання.

Мета кваліфікаційної роботи бакалавра – допомога у підборі друкованої інформації через аналіз зображень їх обкладинок за допомогою відповідного методу штучного інтелекту імplementованого у мобільне застосування.

Завдання кваліфікаційної роботи бакалавра – провести аналіз предметної області автоматизованого розпізнавання друкованих видань за зображеннями обкладинок, а саме розпізнавання назв, жанрів, авторів та типів матеріалів. Проаналізувати наявні рішення подібних задач. Розробити метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання. Інтегрувати метод у мобільний застосунок, який може аналізувати зображення обкладинок та надавати користувачеві персоналізовані рекомендації на основі розпізнаних характеристик видання. Провести тестування розробленого застосунку та оцінити точність розпізнавання та ефективність рекомендацій для різних груп користувачів, враховуючи рівень володіння англійською мовою.

Розділ 1 Характеристика предметної області: аналіз моделей, методів та реалізацій

1.1 Аналіз інформаційних моделей

У сучасному світі процес навчання є одним з найважливіших у житті кожної людини. Технології стрімко розвиваються, а інформація змінюється кожного дня, вміння вчитися впродовж усього життя стає необхідністю [1]. Сучасний ринок праці вимагає від працівників не тільки базової освіти, а й постійного підвищення кваліфікації. Постійний розвиток допомагає знайти кращу роботу, покращити навички та отримати можливості для кар'єрного зростання [2].

Сьогодні складно уявити життя без знання англійської мови. З усіх іноземних мов англійська є першою і вже давно досягла міжнародного рівня. Тому дуже важливо, щоб кожен мав доступ до матеріалів для її вивчення [3].

Вибір правильних підручників і книг – важливий фактор у процесі вивчення та вдосконалення англійської мови. Правильно підібрана література не лише допоможе ефективно засвоїти матеріал, але й мотивуватиме до подальшого навчання. Рекомендується обирати книги, які відповідають володіння мовою [4]. Перед тим, як обрати книгу, корисно ознайомитися з рекомендаціями викладача та відгуками студентів. Це дозволить вибрати найбільш ефективні та перевірені матеріали, які допоможуть покращити граматику, лексичний запас та вимову [5].

Задача ідентифікації друкованих видань за зображенням обкладинок вимагає застосування спеціалізованих інформаційних моделей [6], які дозволяють точно і швидко класифікувати різноманітні типи видань. Основними параметрами для розв'язання цієї задачі є зображення обкладинки та її характеристики, які повинні бути виявлені за допомогою методів комп'ютерного зору [7] та глибокого навчання [8]. Для вирішення цього завдання необхідно розглянути кілька пунктів, таких як властивості зображення, методи аналізу та класифікації, а також типи ознак, які важливі для створення надійної моделі.

Зображення є основним джерелом даних для задачі класифікації обкладинок, тому важливо розглядати його параметри, які визначають його

характеристику та здатність до подальшого аналізу. Зображення обкладинки – це двовимірний масив пікселів, кожен з яких має певну інтенсивність світла для кожного з кольорових каналів (RGB, CMYK) [9]. Параметри зображення, що впливають на ефективність його обробки, включають:

1. Висока роздільна здатність зображення обкладинки важлива для чіткої ідентифікації елементів, таких як текст, логотипи, графічні елементи тощо [10].

2. Зображення, яке має низький контраст або неправильну яскравість, знижує ефективність алгоритмів, особливо якщо обкладинка містить текст [11].

3. Параметри кольору можуть бути критичними для розпізнавання типу видання, оскільки різні видавництва використовують певні кольорові схеми для категоризації своїх видань [12].

Для розв'язання задачі ідентифікації друкованих видань, необхідно визначити конкретні ознаки, які потрібно виявити на зображенні. У даному випадку, до основних ознак належать [13]:

1. Текст – найважливіша ознака для розпізнавання обкладинки, оскільки він містить основну інформацію про видання. Алгоритми розпізнавання тексту використовуються для виявлення та інтерпретації назви книги, імені автора або іншої важливої інформації на обкладинці [14].

2. Графічні елементи (зображення, логотипи, візуальні ознаки) є корисними для класифікації. Для виявлення таких елементів використовуються методи комп'ютерного зору, зокрема детекція об'єктів [15].

3. Виявлення логотипів та брендів важливе для класифікації видань, адже кожен видавець має свій унікальний стиль і логотип [16].

Отже, за результатами аналізу предметної області можна стверджувати, що розробка методу ідентифікації друкованих видань за зображенням їх обкладинки є актуальною і важливою, вона вимагає точного визначення численних ознак на зображеннях. Використання засобів глибокого навчання дозволяє ефективно вирішувати ці проблеми. Такий додаток спростить життя людей, які хочуть покращити знання та шукають для цього потрібну літературу.

1.2 Огляд підходів до розв'язку подібних задач

На сьогоднішній день розпізнавання зображень – одна з найбільш інноваційних та потужних технологій. Кожного року вона набуває все більшої популярності і використовується в різних сферах життя. Завдяки здатності аналізувати та інтерпретувати візуальну інформацію, розпізнавання зображень може вирішувати складні завдання, підвищуючи якість життя, безпеку та ефективність у різних галузях. Нижче буде розглянуто різні статті в яких було вирішено задачі по розпізнаванню та аналізу зображень.

У статті [17] аналізується застосування ШІ для виявлення підробок зображень у соцмережах, використовуючи методи глибокого навчання. Вона розглядає способи боротьби з маніпуляціями з фотографіями за допомогою нейронних мереж та аналізу метаданих. Під час дослідження було проведено експерименти з різними моделями нейронних мереж для виявлення фальшивих зображень. Результати показали, що застосування глибоких нейронних мереж значно підвищує точність ідентифікації змін у зображеннях, дозволяючи автоматизувати процес перевірки фальшивок.

У статті [18] розглядається метод розпізнавання символів на зображеннях за допомогою згорткових нейронних мереж (CNN). Стаття детально описує архітектуру згорткових нейронних мереж, які здатні аналізувати структурні особливості зображень і автоматично виявляти символи, що зображені на них. Автори розробили методику, яка використовує згорткові нейронні мережі для ефективного розпізнавання різних типів символів, включаючи букви, цифри та знаки на зображеннях. Вони провели експериментальні дослідження, які продемонстрували високу точність розпізнавання символів, навіть у випадку шумових або погано освітлених зображень. Згорткові нейронні мережі показали кращу ефективність порівняно з іншими методами машинного навчання, такими як методи на основі підтримки векторних машин (SVM). Результати досліджень свідчать про можливість застосування цих методів для задач розпізнавання символів у реальних умовах.

У наступній статті [19] аналізується застосування ІІ у медичних технологіях для автоматизованого опрацювання медичних зображень й їх аналізу, зокрема рентгенографії та МРТ. Автор детально розглядає роль штучного інтелекту в покращенні процесів діагностики та обробки медичних зображень, зокрема в контексті автоматичного виявлення аномалій. Дослідження показало, що ІІ може знижувати людський фактор в медичних діагнозах, підвищуючи точність виявлення патологій, таких як ракові утворення, на зображеннях. Автор також описує використання нейронних мереж для дослідження медичних зображень, зокрема застосування згорткових нейронних мереж (CNN), які дозволяють виявляти навіть найменші зміни в структурі тканин, що важко помітити людині.

У ще одній статті [20] досліджуються основи машинного та глибокого навчання, а також технології розпізнавання зображень. Надано приклад реалізації глибокого навчання для розробки системи розпізнавання зображень із використанням нейронної архітектури. У підсумку було створено систему, яка відкриває значний потенціал для автоматизації технологічних операцій та покращення їх продуктивності.

У статті [21] аналізується застосування глибоких згорткових нейронних мереж для автоматичного розпізнавання міміки обличчя та ідентифікації емоцій людини. Автори пропонують використання трансферного навчання на базі попередньо навчених моделей DenseNet та MobileNet, що дозволяє скоротити час навчання та скоротити обчислювальні витрати, зберігаючи точність. Результати показали надійне розпізнавання моторних одиниць на наборі даних DISFA.

У статті [22] проведено аналіз архітектур щільних нейронних мереж (DNN) та згорткових нейронних мереж (CNN) для задачі розпізнавання рукописних цифр. Дослідження показало, що CNN демонструють вищу продуктивність завдяки операціям згортки та об'єднання, що робить їх більш ефективними для обробки зображень. Рекомендовано використання контрольованого навчання із зворотним поширенням помилки для досягнення найкращих результатів.

Стаття [23] описує розроблену модель нейронної мережі для розпізнавання хвороб рослин за зображеннями листя. Було проведено дослідження сучасних архітектур глибокого навчання, спрямований на забезпечення значної точності та зменшення кількості помилок під час класифікації зображень. Запропонована система демонструє високу точність та швидкість передбачення, що підкреслює доцільність застосування нейромереж в аграрній сфері.

Отже, кожен з вищезгаданих підходів має плюси та мінуси, однак для реалізації методу ідентифікації друкованих видань за зображенням їх обкладинки найкраще підійдуть згорткові нейронні мережі, оскільки вони найефективніше обробляють інформацію у зображеннях, здатні автоматично виділяти ознаки.

1.3 Аналіз існуючих програмних засобів та наукових рішень

У сфері розпізнавання зображень існує значна кількість програмних інструментів, призначених для виявлення та обробки візуальних даних. Нижче було проаналізовано такі застосунки.

Imagga – це API для розпізнавання зображень на основі штучного інтелекту, який автоматично аналізує вміст пікселів зображень та призначає відповідні теги та ключові слова. Використовуючи моделі глибокого навчання, Imagga здатна ідентифікувати об'єкти, місця та інші елементи на зображеннях [24].



Рисунок 1.1 – Головний екран Imagga [25]

Перевагами цього сервісу є підтримка безлічі функцій, таких як розпізнавання об'єктів, класифікація зображень, аналіз кольору та автоматичне створення тегів; дозволяє з легкістю інтегрувати технології розпізнавання зображень у власні додатки або веб-сайти; завдяки використанню передових технологій машинного навчання та великих навчальних наборів, Imagga демонструє значну точність у процесі розпізнавання зображень. Також є ряд недоліків. Потрібне підключення до інтернету, оскільки Imagga є хмарним сервісом. Обмеження безкоштовної версії. Точність розпізнавання залежить від якості зображень.

Seek – це мобільний додаток, який використовує технології розпізнавання зображень для ідентифікації видів рослин, тварин та грибів. Користувачі можуть фотографувати об'єкти природи, а додаток надасть інформацію про них [26].



Рисунок 1.2 – Екрани додатку Seek від iNaturalist [27]

Цей застосунок має переваги: широка база даних видів, інтерактивний інтерфейс з можливістю відстеження спостережень, освітній контент про природу.

А також декілька недоліків: не завжди точне розпізнавання в умовах низької освітленості, потребує підключення до інтернету для доступу до повної інформації.

Google Lens – це застосунок для розпізнавання зображень, який дозволяє користувачам отримувати інформацію про об’єкти, текст, баркоди та інші елементи на зображеннях. Він інтегрується з іншими сервісами Google для надання детальної інформації [28].



Рисунок 1.3 – Екрани додатку Google Lens [29]

Перевагами цього додатку є широкий спектр можливостей розпізнавання, інтеграція з іншими продуктами Google, а також не менш важлива – зручний інтерфейс з швидким доступом до функцій.

Говорячи про недоліки – це стабільне підключення до інтернету. А також інколи може не розпізнати специфічні або рідкісні об’єкти.

Vivino – це мобільний додаток для розпізнавання вин за етикеткою пляшки. Користувач може сфотографувати етикетку, після чого додаток аналізує її та надає інформацію про вино: рейтинг, відгуки інших користувачів, середню ціну та характеристики смаку. Vivino також пропонує рекомендації, які базуються на вподобаннях користувача [30].



Рисунок 1.4 – Екрани додатку Vivino [31]

Перевагами такого додатку є висока точність розпізнавання винних етикеток завдяки великій базі даних. Також, можливість перегляду оцінок і відгуків інших користувачів. Наступні переваги – персоналізовані рекомендації, сформовані на основі особистих уподобань, та доступ до інформації про ціни та місця покупки вин.

Є і деякі недоліки, а саме: обмежена точність розпізнавання для рідкісних або локальних вин, щоб скористатися певними функціями, потрібно зареєструвати особистий обліковий запис, вартість преміум-підписки для розширених можливостей, потреба у стабільному інтернет-з'єднанні для пошуку в базі даних.

Отже, було проаналізовано низку наявних програмних рішень для розпізнавання зображень. Кожне з них виконує свою функцію та характеризується власними перевагами й недоліками.

1.4 Мета, задачі та вимоги до розробки та реалізації

Метою кваліфікаційної роботи бакалавра є допомога у підборі друкованої інформації через аналіз зображень їх обкладинок за допомогою відповідного методу штучного інтелекту імплементованого у мобільне застосування. Розроблений мобільний застосунок з імплементованим зазначеним методом повинен забезпечити користувачів миттєвою інформацією про характеристики відповідних друкованих видань, що зменшить час на пошуки.

Завданням роботи є:

1. Провести аналіз предметної області автоматизованого розпізнавання друкованих видань за зображеннями їх обкладинок.
2. Проаналізувати наявні рішення подібних задач.
3. Розробити метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання.

4. Інтегрувати метод у мобільний застосунок, який може аналізувати зображення обкладинок та надавати користувачеві персоналізовані рекомендації.

5. Провести тестування розробленого застосунку та оцінити точність розпізнавання та ефективність рекомендацій для різних груп користувачів, враховуючи рівень володіння англійською мовою.

Вимоги до реалізації роботи є:

1. Розроблена інформаційна система повинна забезпечувати надійне розпізнавання друкованих видань на основі зображення їх обкладинки.

2. Застосунок буде розроблено для мобільних пристроїв, тому важливо, щоб він працював швидко та стабільно навіть на не надто потужних смартфонах.

3. Інтерфейс мобільного застосунку має бути простим, з акцентом на зручність для користувачів різного віку та рівня цифрової грамотності.

4. Система повинна мати змогу підключення до централізованої або локальної бази даних з інформацією про друковані видання.

5. Ідентифікація обкладинок має здійснюватися з мінімальною затримкою.

Отже, у цьому підрозділі було розглянуто мету та актуальність кваліфікаційної роботи, поставлено задачі. Встановлені вимоги до реалізації проєкту дозволяють створити ефективну, доступну та зручну інформаційну систему, яка зможе суттєво спростити процес пошуку літератури, підвищити цифрову доступність друкованих видань та покращити користувацький досвід у сфері навігації по книжковому контенту.

Розділ 2 Метод автоматичної ідентифікації друкованих видань та особливості його реалізації

2.1 Метод ідентифікації друкованих видань за зображенням їх обкладинки

2.1.1. Ідея методу

Ідея методу полягає в автоматичній ідентифікації друкованих видань за зображенням їх обкладинки шляхом використання методів глибокого навчання, зокрема згорткових нейронних мереж (Convolutional Neural Networks, CNN) [32]. Візуальне оформлення обкладинки, як правило, має унікальний дизайн, пов'язаний з певним виданням, жанром або серією, що дає змогу використовувати його як основу для класифікації.

Метод передбачає попереднє навчання нейронної мережі на великій кількості зображень обкладинок, що супроводжуються відповідними мітками (назвами або категоріями). Після навчання така модель здатна з високою точністю розпізнавати видання на нових, раніше не бачених зображеннях.

Цей підхід дозволяє автоматизувати процес розпізнавання та систематизації друкованих матеріалів у бібліотеках, архівах, електронних каталогах або онлайн-магазинах, де класифікація за обкладинкою може бути швидшим або єдиним можливим способом ідентифікації.

2.1.2 Модель нейромережі

Згорткові нейронні мережі є особливим класом штучних нейронних мереж, які відзначаються високою ефективністю при обробці зображень, відео та інших даних, що мають просторову структуру. Основною особливістю CNN є здатність автоматично виявляти характерні ознаки у вхідних даних за допомогою спеціальних шарів – згорткових і підвибіркових.

У методі ідентифікації друкованих видань за зображенням їх обкладинки буде використано згорткову нейронну мережу. Модель побудована на основі архітектури ResNet50, яка є особливо ефективною у задачах класифікації зображень завдяки використанню залишкових з'єднань.

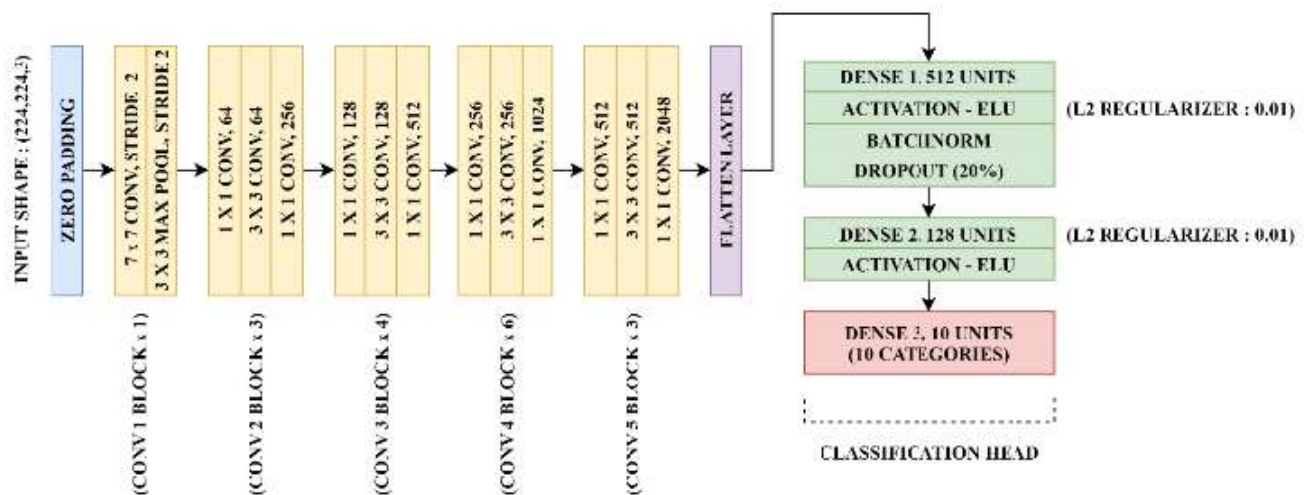


Рисунок 2.1 – Архітектура ResNet50 [33]

Архітектура ResNet50 складається з 49 згорткових шарів і одного повнозв'язного шару [34], що загалом становить 50 шарів. Основним будівельним блоком мережі є так званий «bottleneck» (вузьке місце), який включає три послідовні згорткові шари: 1x1, 3x3 і знову 1x1. Перший шар зменшує кількість каналів, другий виконує основну згортку, а третій відновлює розмірність. Кількість фільтрів у шарах поступово зростає: від 64 до 2048, що дозволяє моделі витягувати дедалі складніші ознаки зображень. Розмір ядра у більшості згорткових операцій становить 3x3, а крок і падінг підібрані так, щоб зберегти просторові розміри або зменшувати їх у контрольований спосіб. Після згорткових шарів додається повнозв'язний шар із 1024 нейронами, а фінальний класифікаційний шар містить 30 нейронів – по одному для кожного жанру книги, що класифікується.

У методі буде застосовано техніку трансферного навчання: ваги основної частини моделі ResNet50 заморожуються, а перенавчання підлягають лише нові

шари, що дозволяє ефективно адаптувати модель до специфіки нового набору зображень без потреби у великій кількості навчальних даних.

Загалом, використання ResNet50 у методі ідентифікації друкованих видань за зображенням їх обкладинки демонструє високу ефективність глибоких згорткових нейронних мереж у задачах класифікації, зокрема завдяки стійкості до перенавчання та гнучкості у використанні за допомогою трансферного навчання.

2.1.3. Кроки методу

У ході проектування методу ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання можна виділити основні кроки для навчання моделі (рисунку 2.2).



Рисунок 2.2 – Схема навчання моделі

На вхід надходить набір зображень обкладинок друкованих видань, кожне з яких має відповідну мітку (наприклад, жанр, назва, категорія). Ці зображення формують датасет для навчання.

У процесі методу першим кроком 1 є попередня обробка, де зображення приводяться до однакового розміру, нормалізуються (піксельні значення масштабуються). Потім на кроці 2 всі дані діляться на тренувальну і валідаційну вибірки. Далі кроком 3 є навчання моделі, використовуючи глибоку неймережу, яка навчається класифікувати обкладинки.

Результатом є навчена модель, тобто нейронна мережа зі збереженими вагами, готова до використання на нових зображеннях.

Також на рисунку 2.3 зображена спроектована схема використання моделі.



Рисунок 2.3 – Схема використання моделі

Вхідними даними є нове зображення обкладинки, яке потрібно розпізнати (класифікувати), без жодної додаткової інформації про нього.

У процесі на кроці 1 зображення обробляється так само, як і під час навчання (змінюється розмір, нормалізується). Далі крок 2 – завантаження раніше навченої моделі. А також кроком 3 є класифікація, де модель виконує прогноз, обчислюючи ймовірності належності до кожного класу, та видає найбільш ймовірний варіант. У результаті вихідними даними є ідентифіковане зображення (категорія книги, її назва).

Для навчання згорткової нейронної мережі, яка реалізує класифікацію друкованих видань за зображенням обкладинки, буде використано структурований датасет, що містить метадані та посилання на зображення книжкових обкладинок. Дані зберігаються у вигляді таблиці CSV, де кожен рядок відповідає окремому книжковому виданню. Для кожного зразка наведено атрибути в таблиці 2.1.

Таблиця 2.1 – Атрибути

Атрибут	Пояснення
ISBN	Міжнародний ідентифікатор книги.
Назва зображення	Ім'я файлу або посилання на обкладинку.
Назва книги	Заголовок видання.
Автор	Ім'я або назва авторського колективу.
Рейтинг	Умовна оцінка за популярністю.
Категорія	Жанр книги, який виступає у ролі цільової змінної (класу) для моделі.

Нижче на рисунку 2.4 зображено приклад зображення обкладинки друкованого видання з датасету.

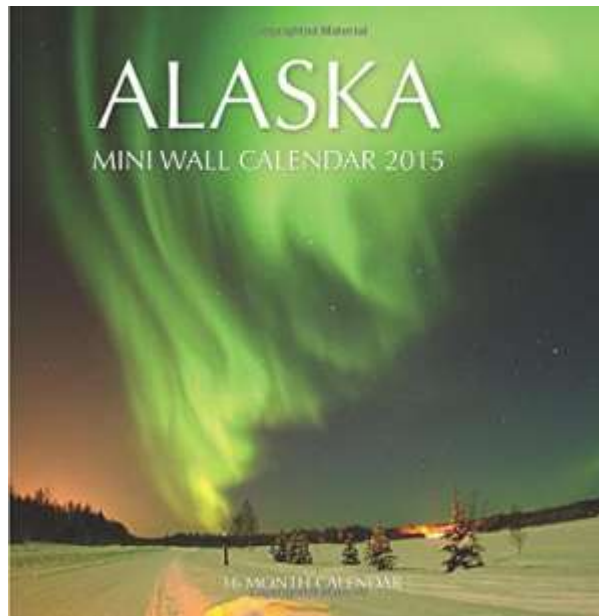


Рисунок 2.4 – Приклад зображення обкладинки з датасету

З метою зменшення надмірної різноманітності класів та забезпечення стабільності навчання, з усіх жанрів було відібрано топ-30 найбільш представлених категорій.

Таким чином, модель навчається розпізнавати жанрову належність серед 30 чітко окреслених класів, таких як Calendars, Children's Books, Literature & Fiction та інші.

Кількість зразків у кожному класі варіюється, що підтверджується побудованою стовпчиковою діаграмою. Вона дозволяє візуалізувати співвідношення жанрів у датасеті книжкових обкладинок (рисунок 2.5).

Видно, що деякі жанри представлені значно більше, ніж інші (наприклад, Travel та Children's Books мають найбільшу кількість зразків, а Parenting & Relationships та Mystery, Thriller & Suspense – найменшу).

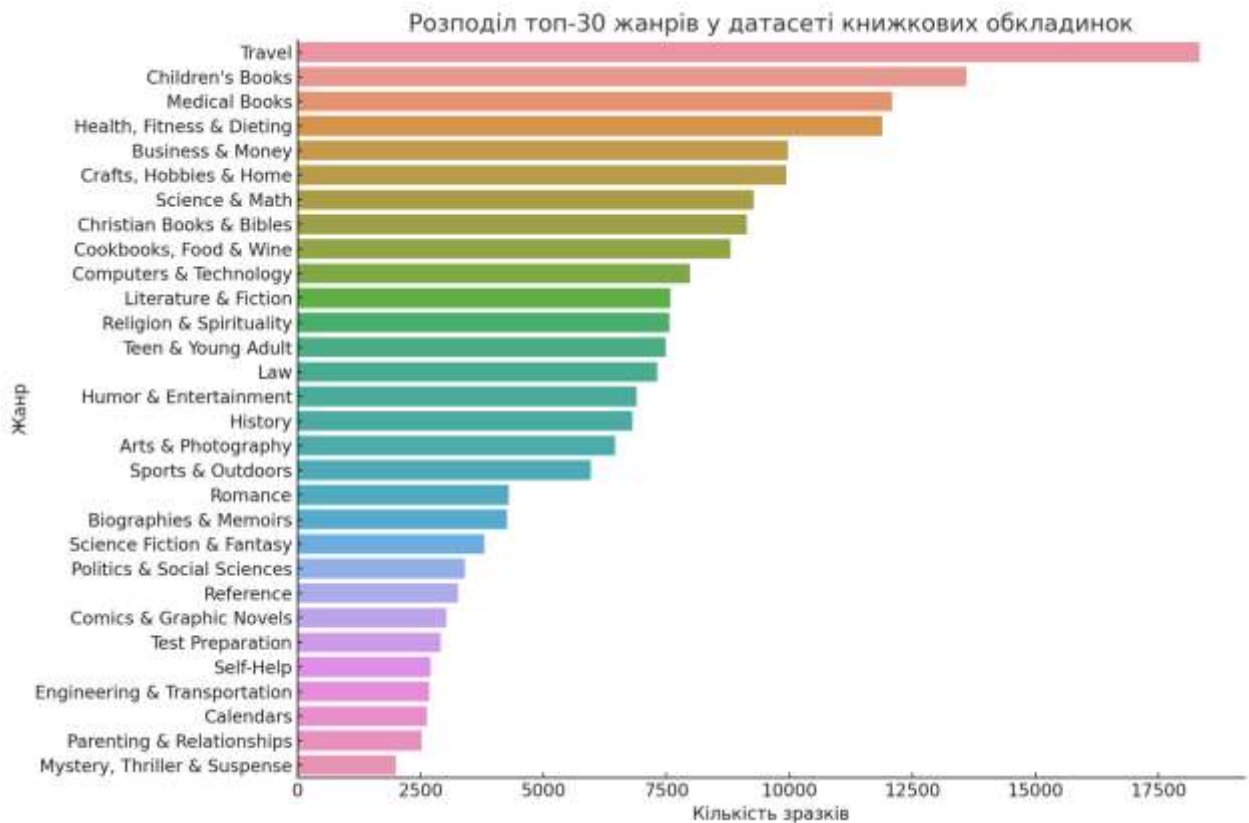


Рисунок 2.5 – Стовпчикова діаграма співвідношення класів

Обраний датасет виступає основою для формування навчальної та валідаційної вибірок, які дозволяють коректно оцінити ефективність побудованої моделі на основі архітектури ResNet50 у задачі багатокласової класифікації зображень.

2.2. Особливості реалізації запропонованого методу

Планується реалізація методу ідентифікації друкованих видань за обкладинкою шляхом побудови системи класифікації зображень за допомогою згорткової нейронної мережі на базі архітектури ResNet50. Для реалізації буде використано високорівневу бібліотеку Keras [35] у складі фреймворку TensorFlow, що забезпечує простоту побудови, навчання та розгортання моделей глибокого навчання.

TensorFlow [36] – це фреймворк з відкритим кодом, розроблений компанією Google. Він надає масштабовану платформу для обчислень з

автоматичним диференціюванням, яка ідеально підходить для реалізації глибоких нейронних мереж. TensorFlow дозволяє ефективно працювати як на центральному процесорі CPU, так і на графічному процесорі GPU, що дозволяє суттєво пришвидшити обрахунки при роботі з великою кількістю даних або складними моделями, такими як ResNet50.

У методі TensorFlow буде виступати як обчислювальний бекенд: він забезпечить роботу графа операцій, обробку тензорів, зберігання ваг моделі та автоматичне обчислення градієнтів під час тренування. Загальний принцип побудови обчислень у TensorFlow ґрунтується на створенні графа операцій, де кожен вузол відповідає за окрему математичну функцію або трансформацію, а стрілки між вузлами – за передачу тензорів (рисунок 2.6) [37].

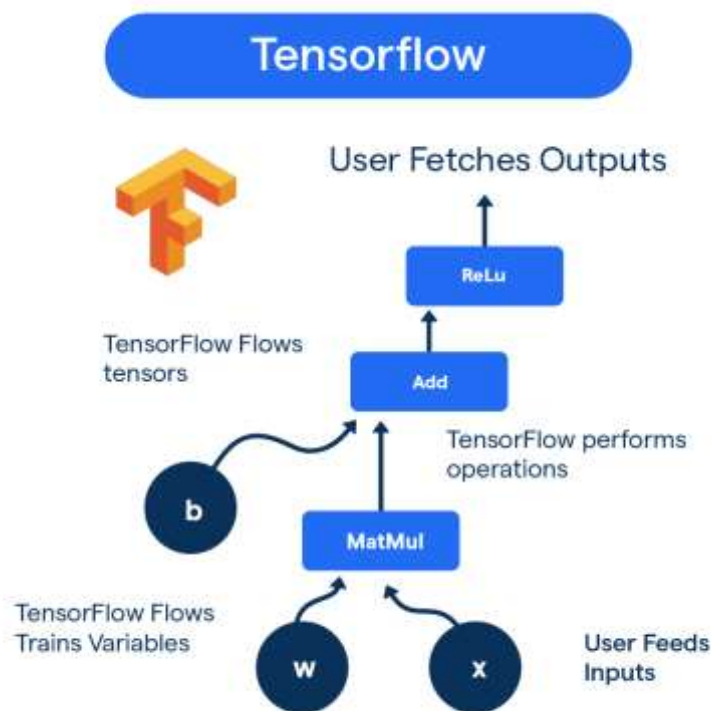


Рисунок 2.6 – Архітектура обчислювального графа TensorFlow [37]

Keras – це високорівнева бібліотека з відкритим кодом для побудови та тренування нейронних мереж, яка спрощує роботу з TensorFlow, забезпечуючи зрозумілий інтерфейс для розробки моделей глибокого навчання. Більше того Keras дозволяє:

- швидко завантажити попередньо навчену модель ResNet50 з готовими вагами;
- легко будувати послідовну або функціональну модель з кількох шарів;
- додавати шари та вказувати функції активації;
- використовувати засоби для компіляції та тренування моделі, включаючи вибір оптимізатора, функції втрат та метрик;
- зручно зберігати і завантажувати модель у різних форматах.

Архітектура Keras має в основі три основних компоненти: Model, Layer та Core Modules. Ця структура дозволяє ефективно організовувати та управляти процесом побудови нейронних мереж (рисунок 2.7) [38].

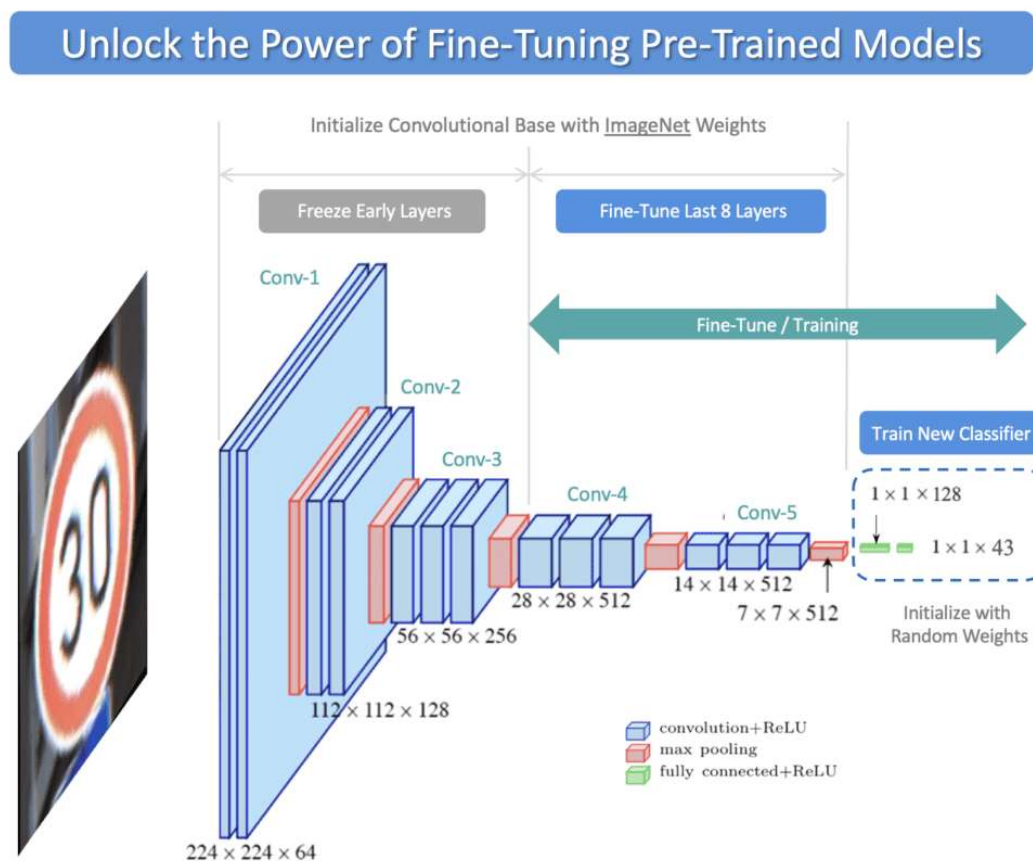


Рисунок 2.7 – Архітектура Keras [38]

Для попередньої обробки зображень буде використано компонент ImageDataGenerator з Keras. Він дозволяє змінювати розмір зображень до формату, прийнятого для моделі (224×224 пікселі), масштабувати піксельні значення та застосовувати аугментацію даних [39] – обертання, зсуви, дзеркальне

відображення, що дозволяє штучно збільшити обсяг навчального набору та знизити ризик перенавчання.

Як вже було раніше зазначено, у методі буде реалізована техніка трансферного навчання: базові шари ResNet50 залишаться замороженими, що дозволить зберегти узагальнюючі властивості, отримані під час початкового навчання на ImageNet. Навчання відбудуватиметься лише для нових, доданих шарів класифікації.

2.3. Проектування інформаційної системи для ідентифікації зображень обкладинки друкованих видань

На даному етапі розробки було здійснено проектування інформаційної системи, що реалізує задачу ідентифікації друкованих видань на основі аналізу їхніх обкладинок. Основною метою проектування є побудова логічної та структурної моделі системи, що забезпечить ефективне виконання всіх етапів обробки даних.

На рисунку 2.8 представлено діаграму класів [40] інформаційної системи.

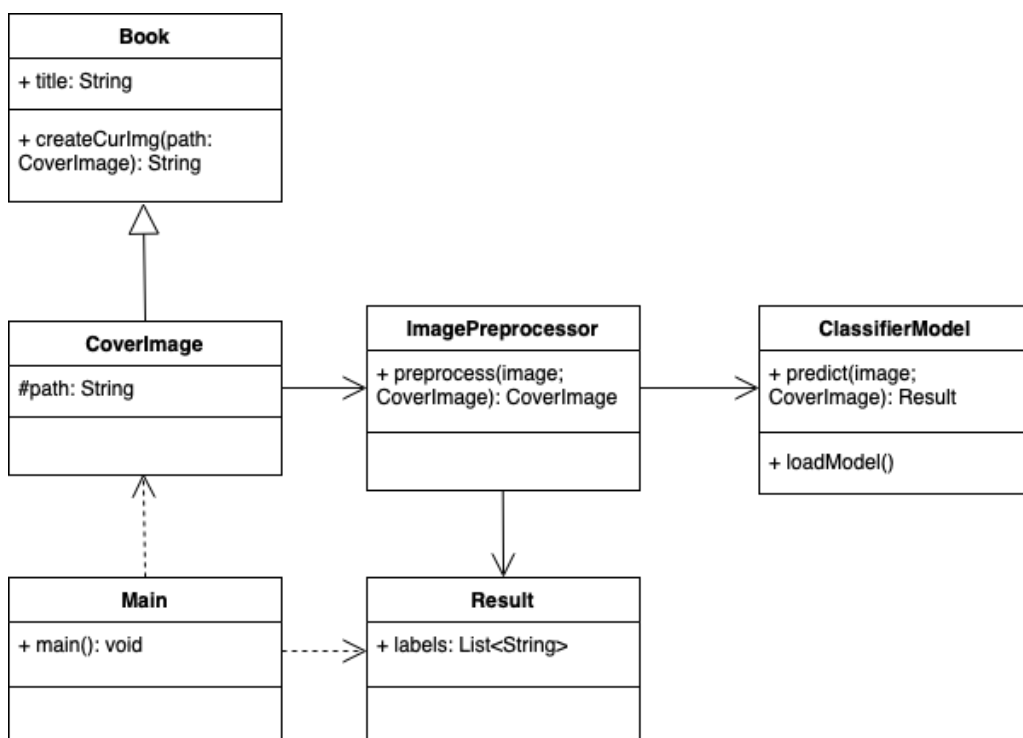


Рисунок 2.8 – Діаграма класів

Вона призначеної для класифікації друкованих видань за зображенням обкладинки. Діаграма побудована на основі функціональної логіки проекту.

Основними класами системи є:

1. `Book` – відповідає за логіку книжкового об'єкта. Містить атрибут `title` для збереження назви книги та метод `createCurImg()`, який генерує зображення обкладинки на основі шляху до файлу. Цей клас виступає ініціатором подальших дій у системі.

2. `CoverImage` – окремий клас, що представляє обкладинку книги як зображення. Має атрибут `path`, який вказує на місце розташування графічного файлу.

3. `ImagePreprocessor` – компонент, який відповідає за попередню обробку зображень. Основний метод `preprocess()` виконує масштабування, нормалізацію та (за потреби) аугментацію даних перед подачею на вхід моделі.

4. `ClassifierModel` – клас, що реалізує основну функцію класифікації. Містить метод `loadModel()` для завантаження попередньо натренованої моделі, а також метод `predict()` для здійснення передбачення жанру на основі вхідного зображення.

5. `Result` – клас для збереження результатів класифікації. Має атрибут `labels`, який містить список ймовірнісних або остаточних жанрових міток.

6. `Main` – головний керуючий клас, який ініціює роботу всієї системи. Містить метод `main()`, що відповідає за послідовне виконання всіх операцій – від завантаження даних до виводу результатів.

Зв'язки між класами відображають логіку обробки зображення: спочатку створюється об'єкт `Book`, що генерує `CoverImage`, яка передається до `ImagePreprocessor`, далі до `ClassifierModel`, результат зберігається в об'єкті `Result`, а координація здійснюється через `Main`.

На рисунку 2.9 представлено діаграму активності [41], яка відображає переходи між різними видами діяльності.

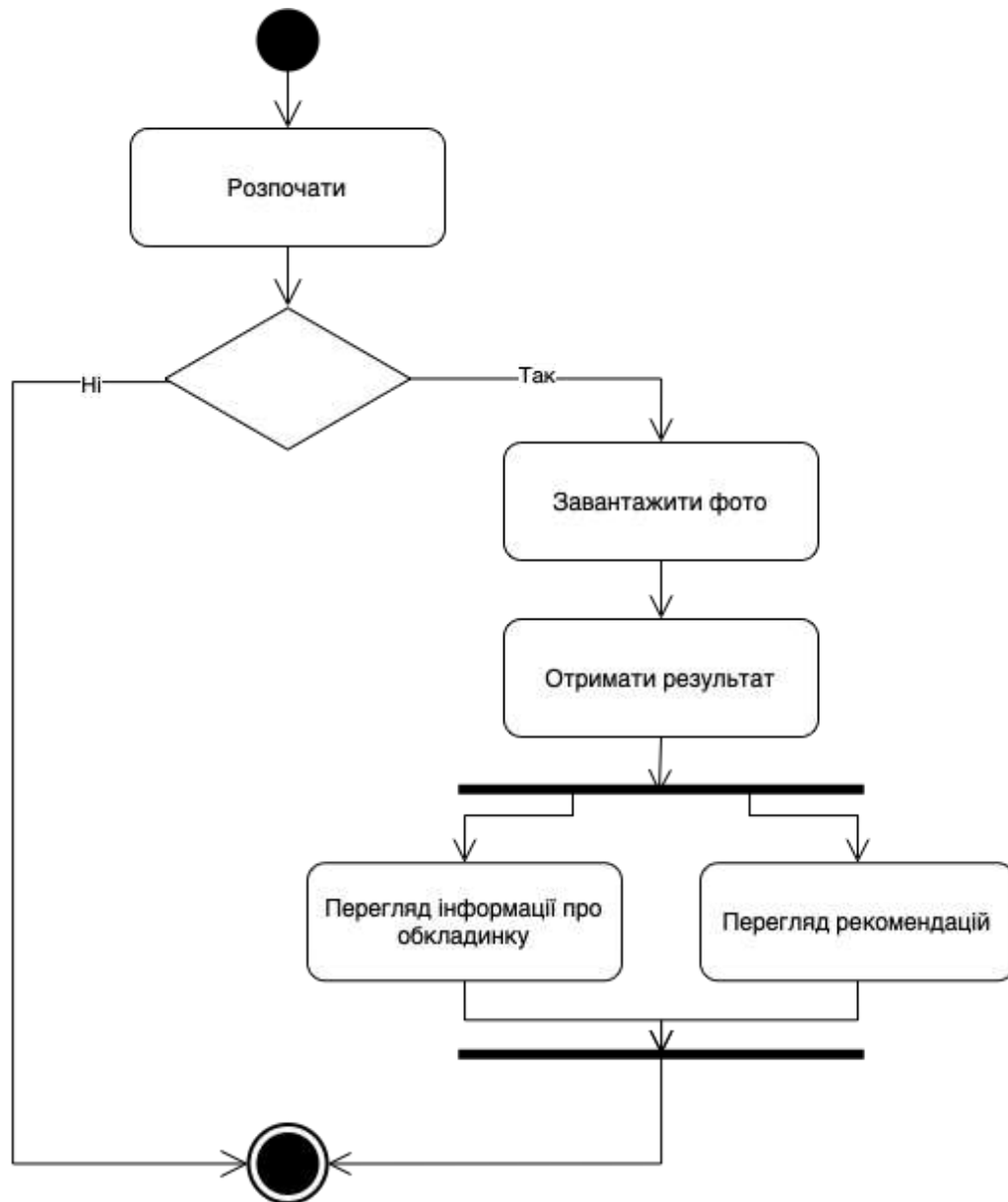


Рисунок 2.9 – Діаграма активності

Спочатку користувач завантажує/сканує фото обкладинки, якщо хоче. Якщо так, то отримує результат, після того як система виконує класифікацію. Потім користувач може переглянути опис проаналізованої обкладинки та переглянути рекомендації.

Діаграма послідовності [42] (рисунок 2.10) демонструє взаємодію між об'єктами системи в часі.

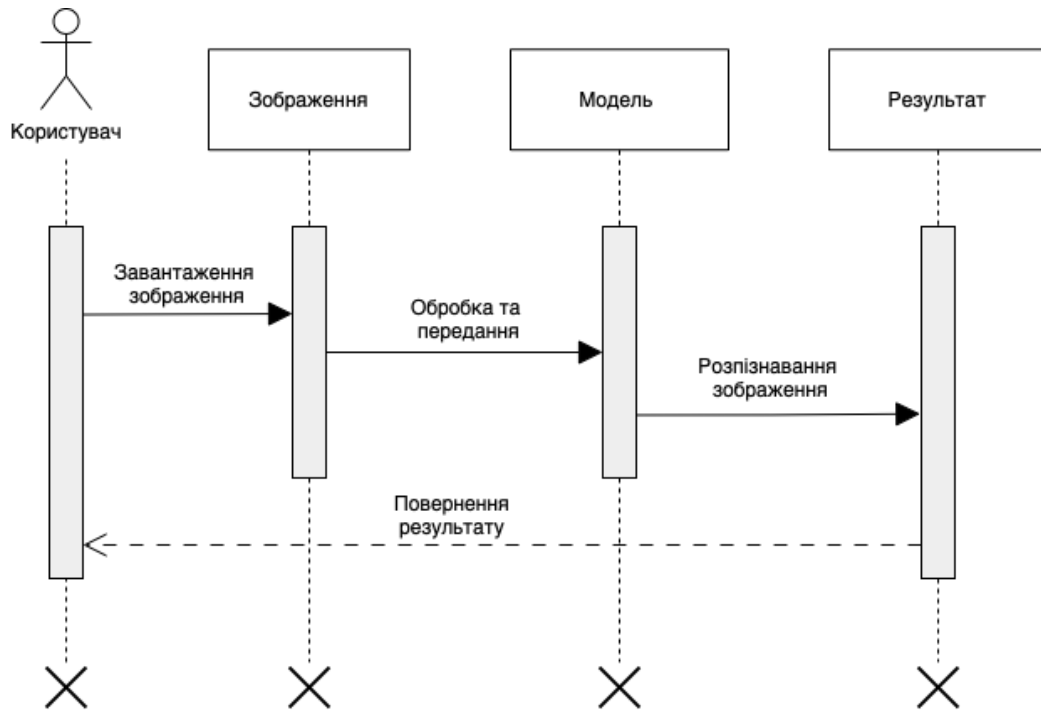


Рисунок 2.10 – Діаграма послідовності

У процесі беруть участь чотири основні компоненти: Користувач, Зображення, Модель, Результат. Користувач передає зображення, яке далі обробляється та передається до моделі. Після здійснення розпізнавання результат надсилається користувачу.

Діаграма варіантів використання [43] (рисунок 2.11) ілюструє функціональні можливості системи з боку користувача.



Рисунок 2.11 – Діаграма варіантів використання

Актором системи виступає користувач, який взаємодіє із системою через кілька ключових варіантів використання:

Варіанти використання:

Завантаження фото – початковий варіант використання, в якому користувач завантажує зображення обкладинки книги для подальшої обробки.

Очікування результату – проміжний етап, під час якого система здійснює обробку та класифікацію завантаженого зображення.

Перегляд інформації про обкладинку – після отримання результату користувач має можливість переглянути детальну інформацію, зокрема назву книги, жанр, рейтинг тощо.

Перегляд рекомендацій – користувач може ознайомитися із запропонованими системою книгами, подібними за жанром або візуальними ознаками.

Вихід з функції – завершальний варіант використання, коли користувач припиняє взаємодію із системою або повертається до початкового екрану.

Сценарій:

Назва прецеденту: розпізнавання зображення обкладинки.

Дійова особа: користувач.

Мета: отримати інформацію про книгу, підручник чи іншу літературу.

Передумови: користувач обрав функцію для розпізнавання обкладинки.

Головна послідовність:

1. Користувач обирає у головному меню функцію розпізнавання обкладинки.
2. Користувач завантажує або сканує зображення обкладинки.
3. Користувач чекає результату розпізнавання.
4. Користувач отримує результату розпізнавання обкладинку літератури.
5. Користувач переглядає інформацію про обкладинку, яку він завантажив.
6. Користувач закриває функцію та виходить в головне меню.

На рисунку 2.12 представлено діаграму станів [44], яка відображає зміну внутрішнього стану системи в процесі класифікації зображення обкладинки книги.

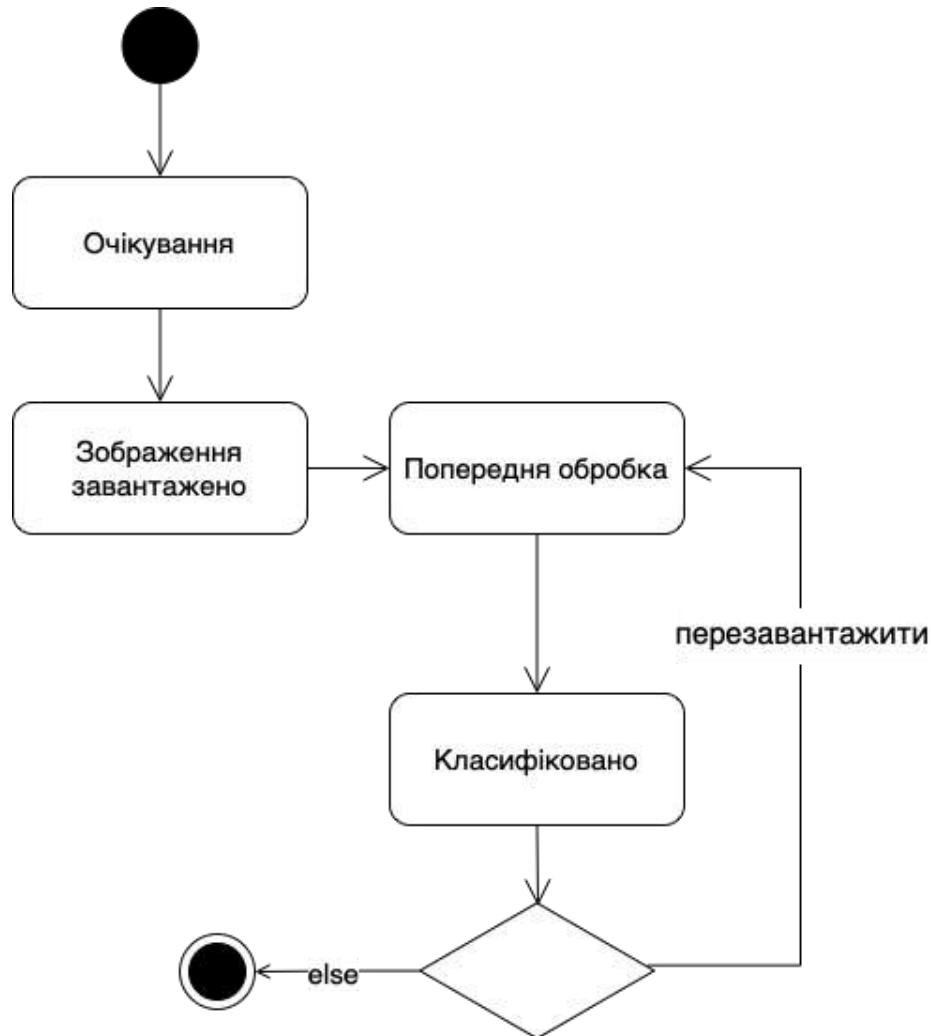


Рисунок 2.12 – Діаграма станів

Основні стани системи:

1. Очікування – початковий стан, у якому система перебуває в неактивному режимі, очікуючи на дію користувача.

2. Зображення завантажено – перехід до цього стану відбувається після того, як користувач завантажує зображення обкладинки книги.

3. Попередня обробка – на цьому етапі система здійснює зміну розміру зображення, нормалізацію та інші необхідні трансформації (відповідає функціональності класу ImagePreprocessor).

4. Класифіковано – фінальний стан, у якому система завершила передбачення та класифікувала обкладинку за відповідним жанром (виконується через клас ClassifierModel).

Перехід між станами:

1. З «Очікування» до «Зображення завантажено» – викликається подією завантаження зображення.

2. З «Зображення завантажено» до «Попередня обробка» – система починає обробку даних.

3. З «Попередня обробка» до «Класифіковано» – після обробки зображення запускається класифікація.

4. З «Класифіковано» до «Попередня обробка» – можливий перехід у разі перезапуску обробки.

У будь-який момент можливий перехід «перезавантажити», що повертає систему до стану «Попередня обробка».

Отже, було розглянуто проектування інформаційної системи для ідентифікації обкладинок друкованих видань на основі зображень. Для моделювання логіки та структури системи використано декілька типів UML-діаграм, що дозволяють чітко відобразити основні компоненти, взаємодії та сценарії використання системи.

2.4 Висновки до розділу 2

У цьому розділі було детально описано процес розробки інформаційної системи для ідентифікації зображень обкладинок друкованих видань, використовуючи методи глибокого навчання. Початково було сформовано схему та описано кроки методу, що включають етапи попередньої обробки зображень, класифікації за допомогою нейронної мережі ResNet50. Буде використано глибоке навчання та трансферне навчання для досягнення високих результатів на нових даних.

Датасет, використаний для тренування моделі, складається з великої кількості зображень обкладинок книг з різних жанрів. Це дозволяє моделі ефективно класифікувати нові зображення.

Проектування інформаційної системи було виконано за допомогою UML-діаграм, що дозволяють чітко представити структуру системи, процеси взаємодії користувача із системою та механізм роботи моделі. Діаграми класів, активності, послідовності, варіантів використання та станів відображають основні етапи роботи системи: від завантаження зображення до отримання результату класифікації обкладинок.

Розроблена система буде гнучкою та масштабованою, що дозволить легко адаптувати її до нових даних і задач. Використання глибоких нейронних мереж у поєднанні з чітким проектуванням дасть змогу досягти високої точності ідентифікації обкладинок, що є важливим кроком у розвитку технологій автоматизації в галузі обробки зображень.

Розділ 3 Експериментальне дослідження програмної реалізації запропонованого методу

3.1 Опис інтерфейсу користувача

Інтерфейс користувача [45] є одним із основних компонентів будь-якого сучасного програмного забезпечення, зокрема мобільного застосунку. Від того, наскільки зручно, логічно й зрозуміло побудовано взаємодію користувача з системою, залежить ефективність її використання, ступінь залучення аудиторії та загальний рівень задоволення від роботи з програмою. Саме тому у процесі розробки даного застосунку особлива увага приділялася користувацькому досвіду (UX) [46], простоті навігації та візуальній доступності основного функціоналу.

Додаткову увагу було приділено інтуїтивності взаємодії – користувач не потребує спеціальних інструкцій для користування застосунком. Усі дії зведено до мінімальної кількості кроків: завантаження або зйомка, автоматичний аналіз, отримання результату. Це скорочує шлях користувача до цільової дії та підвищує ефективність взаємодії з системою.

Інтерфейс застосунку було реалізовано на основі фреймворку SwiftUI [47] – сучасного інструменту від Apple для створення динамічних, адаптивних інтерфейсів для платформи iOS. Це дозволило досягти високої продуктивності, підтримки новітніх дизайн-підходів Apple та забезпечити узгодженість зовнішнього вигляду з очікуваннями користувачів iPhone.

Після запуску застосунку користувача зустрічає перший екран, виконує функцію вітального онбордингу [48] та водночас є важливим етапом персоналізації користувацького досвіду. Після запуску застосунку користувач бачить привітання з назвою сервісу та коротке пояснення його призначення: знаходити інформацію про будь-яку книгу без зайвих зусиль.

Нижче розміщене головне запитання – вибір поточного рівня англійської мови, яким володіє користувач, що необхідний для подальшої персоналізації рекомендацій книжкового контенту. Рівні подані згідно з міжнародною шкалою CEFR: від A1 (початковий) до C2 (професійний). Кожен варіант оформлений у

вигляді окремої кольорової кнопки великого розміру з м'якими відтінками, що забезпечує візуальну диференціацію та полегшує вибір (рисунок 3.1).



Рисунок 3.1 – Екран онбордингу

Кнопка «Пропустити», розташована вгорі, дозволяє перейти до основного функціоналу без налаштування, а неактивна кнопка «Далі» внизу стає доступною лише після вибору рівня, що виконує роль механізму підтвердження введення.

Після того як користувач взаємодіє з одним із запропонованих варіантів мовного рівня, інтерфейс переходить у стан активного підтвердження вибору. Обраний рівень виділяється візуально, що дозволяє миттєво зчитати стан інтерфейсу навіть при побіжному перегляді. Крім того, кнопка «Далі» стає активною (рисунок 3.2).



Рисунок 3.2 – Екран онбордингу після взаємодії

Після завершення онбордингу користувач потрапляє на головний екран застосунку, який є центром взаємодії з функціональністю системи. У верхній частині розміщено динамічний блок із персоналізованими рекомендаціями, сформованими на основі раніше обраного рівня володіння англійською мовою. Це дозволяє одразу отримати доступ до контенту, релевантного саме тому рівню складності, який підходить конкретному користувачеві.

Нижче розташований блок «Історія пошуку», який наразі порожній, але згодом автоматично наповнюватиметься результатами обробки обкладинок, які користувач сканує. Це рішення забезпечує контекстуальну пам'ять і дозволяє повернутися до раніше переглянутих результатів у будь-який момент (рисунок 3.3).



Рисунок 3.3 – Головний екран

У нижній навігаційній панелі знаходяться три основні вкладки: «Головна», Кнопка камери в центрі – виконує роль основної дії, «Профіль».

Акцент на великій кнопці з знаком камери підкреслює головну мету застосунку – швидкий запуск сканування.

Після натискання на цкнопку зі значком камери, застосунок переходить до етапу зчитування зображення обкладинки книги. Згідно з політикою безпеки iOS, перед запуском камери система автоматично виводить стандартне системне діалогове вікно з проханням надати дозвіл на доступ до камери пристрою (рисунок 3.4).

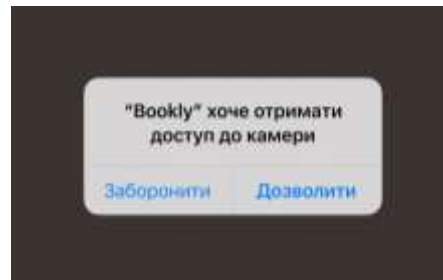


Рисунок 3.4 – Повідомлення про отримання доступу до камери

Відмова призведе до неможливості сканування у реальному часі, але користувач зможе завантажити зображення з галереї. Надання дозволу активує повноцінний режим захоплення зображень – основну функцію застосунку.

Після надання дозволу на використання камери, застосунок переходить у режим активного сканування, де користувач має змогу зробити фотографію обкладинки книги. Інтерфейс цього екрана побудований за принципом фокусованої дії, де всі візуальні й функціональні елементи спрямовані на одну основну ціль – якісне розпізнавання зображення (рисунок 3.5).



Рисунок 3.5 – Екран сканування зображення

У центрі екрана розташовано прямокутник-рамку, який виконує роль орієнтира: користувачеві пропонується розмістити обкладинку книги в центрі цього маркера. Текстова інструкція «Проскануй обкладинку, розмісти в центрі прямокутника» дублює функцію рамки, підкріплюючи її зміст.

Кнопка затвора розміщена у звичному для камери місці, внизу по центру, що відповідає очікуванням користувача й забезпечує інтуїтивну взаємодію. Ліворуч знаходиться кнопка для переходу до галереї (альтернативний варіант завантаження обкладинки), а праворуч – перемикач камери. Кнопка спалаху зверху праворуч доступна у разі потреби покращення освітлення.

Після успішного сканування обкладинки, користувач потрапляє на екран результату розпізнавання, де відображається повна інформація про знайдене друковане видання. Інтерфейс побудований так, щоб максимально спростити ознайомлення з характеристиками книги та водночас дозволити виконати додаткові дії – зберегти результат або переглянути схожі видання (рисунок 3.6).



Рисунок 3.6 – Результат сканування зображення

У верхній частині екрана розміщено обкладинку книги та її назву. Одразу під назвою зазначено, що видання відповідає раніше обраному рівню англійської мови, що є важливим елементом персоналізації рекомендацій.

Праворуч розташована велика кнопка «Зберегти», стилізована у вигляді жовтого маркера з піктограмою закладки. Це дозволяє користувачу зафіксувати результат у своєму профілі або історії для подальшого перегляду. Такий функціонал підтримує звичну для користувачів логіку «додати у вибране» та покращує довготривалу взаємодію з додатком.

Нижче наведено базову метаінформацію про видання у вигляді чітко структурованого блоку. Такий формат подачі даних забезпечує швидке сканування очима, оскільки важливі слова виділено жирним, а структура блоку витримана у стилі інтерфейсних довідників.

У нижній частині екрана є розділ «Схоже», який містить карусель із виданнями, схожими на розпізнану книгу за стилем, серією або тематикою. Це дозволяє користувачеві продовжити пошук або дізнатися більше про серію книг, яку він щойно відсканував.

Цей екран є основним з точки зору відчуття завершеного результату. Користувач не лише отримує розпізнану інформацію, а й одразу бачить її цінність (відповідність рівню, автори, жанр), має змогу зберегти або порівняти, і все це у межах одного зручного перегляду.

Завдяки зручній структурі, акценту на візуальній ієрархії та доповненню релевантним контентом, інтерфейс цього екрана створює відчуття завершеного циклу взаємодії та підвищує ймовірність повторного використання застосунку.

3.2 Результати експериментальних досліджень

У цьому пункті розглянуто результати експериментальних досліджень методу класифікації книжкових обкладинок за жанрами з використанням глибокого навчання. Вони були проведені з метою оцінки ефективності роботи

інформаційної системи для допомоги користувачам у підборі літератури. Ця система базується на застосуванні згорткової нейронної мережі ResNet50, попередньо навченої на великому обсязі зображень, що забезпечує високу точність у класифікації обкладинок книг за жанровими категоріями. Окрім жанрової класифікації, система використовує технологію OCR для автоматичного виділення текстової інформації з обкладинки книги (назви, автора, року видання). Це дозволяє сформуванню структуровану картку книги та надати користувачеві повну інформацію про видання. З метою інтеграції моделі у мобільний застосунок для iOS, було здійснено експорт навченої моделі у форматі Core ML [49]. Цей формат дозволяє ефективно імплементувати модель безпосередньо в середовищі розробки Xcode [50]. Фінальна реалізація здійснена у SwiftUI.

Для візуальної оцінки якості роботи моделі класифікації було використано матрицю помилок [51]. На рисунку 3.7 представлена матриця помилок.

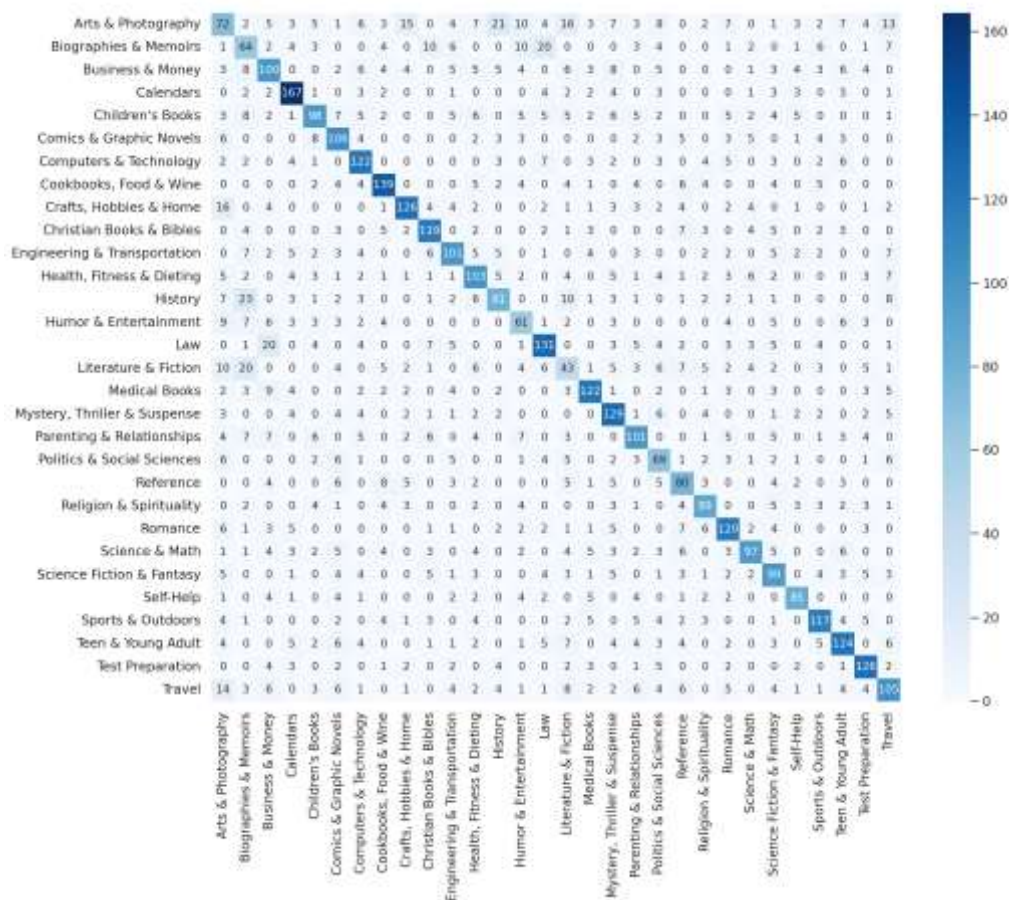


Рисунок 3.7 – Матриця помилок

Вона дозволяє побачити, наскільки точно модель розпізнає жанри книжкових обкладинок, а також визначити, які класи найчастіше плутаються між собою. Завдяки цьому можна зробити висновки щодо сильних та слабких сторін системи і окреслити напрямки для її подальшого вдосконалення. Вона ілюструє результати роботи моделі класифікації жанрів книжкових обкладинок, побудованої на основі нейромережі ResNet50.

Рядки матриці позначають фактичні жанри книжок у тестовій вибірці, а стовпці – жанри, передбачені моделлю. Сині клітинки по діагоналі – правильно класифіковані приклади. Згідно матриці видно, що найкраще класифіковані жанри: Children’s Books зі значенням 198, Cookbooks, Food & Wine – 139, Medical Books – 122, Test Preparation – 126. Це свідчить про те, що модель особливо добре розпізнає візуальні риси книжок цих жанрів. Схильність до помилок спостерігається між суміжними жанрами. Biographies & Memoirs часто плутаються з History, Religion та Politics; Mystery, Thriller & Suspense мають 16 хибних класифікацій як Literature & Fiction і тд. Для жанрів із візуально подібними обкладинками або перетином тематик, таких як Self-Help, Reference, Religion & Spirituality, часто виникають помилки класифікації. Отже, матриця демонструє високу ефективність моделі у класифікації більшості жанрів, з особливою точністю для структурованих та візуально характерних класів. Натомість, жанри з розмитими або перехресними тематиками мають нижчу точність через зовнішню схожість обкладинок або змістову багатозначність.

Для оцінки ефективності роботи моделі класифікації книжкових обкладинок за жанрами було обчислено чотири основні метрики – accuracy, precision, recall та F1-score [52], окремо для тренувальної та тестової вибірок, їх значення відображено у таблиці 3.1.

Таблиця 3.1 – Метрики

Метрика	Тренувальна вибірка	Тестова вибірка
Accuracy	0.95	0.87
Precision	0.93	0.85

Recall	0.94	0.84
F1-score	0.935	0.845

Метрика accuracy становить 0.95 на тренувальній вибірці свідчить про те, що модель змогла правильно класифікувати переважну більшість зразків під час навчання. 0.87 на тестовій вибірці – це високий показник у задачі багатокласової класифікації (30 жанрів), він демонструє, що модель не лише запам'ятала тренувальні дані, але й добре узагальнює нові приклади.

Precision має значення 0.93 для тренування, це означає, що модель майже не видає неправильних позитивних передбачень: якщо вона каже, що це певний жанр, то з великою ймовірністю це так. На тестовій вибірці precision знижується до 0.85, що свідчить про деяке зростання хибнопозитивних класифікацій, тобто іноді модель «плутає» жанри, зокрема близькі за тематикою. Це очікуване зниження й загалом означає, що модель все ще демонструє високу точність у передбаченнях.

Recall 0.94 на тренуванні свідчить про те, що модель майже завжди знаходить правильний жанр серед усіх наявних. Метрика 0.84 на тестовій вибірці означає, що частина правильних класів не була розпізнана, і модель часом пропускає правильний варіант.

Метрика F1-score становить 0.935 на тренувальній та 0.845 на тестовій – збалансована метрика, яка враховує як precision, так і recall. Різниця свідчить, що модель зберігає стабільну якість класифікації. Показник 0.845 на тестовій вибірці є відмінним результатом у багатокласовій класифікації.

Також на рисунку 3.8 зображено графік, на якому видно, що всі значення на тестовій вибірці дещо нижчі, ніж на тренувальній, але залишаються на високому рівні. Це підтверджує, що модель добре узагальнює результати і не є перенавченою.

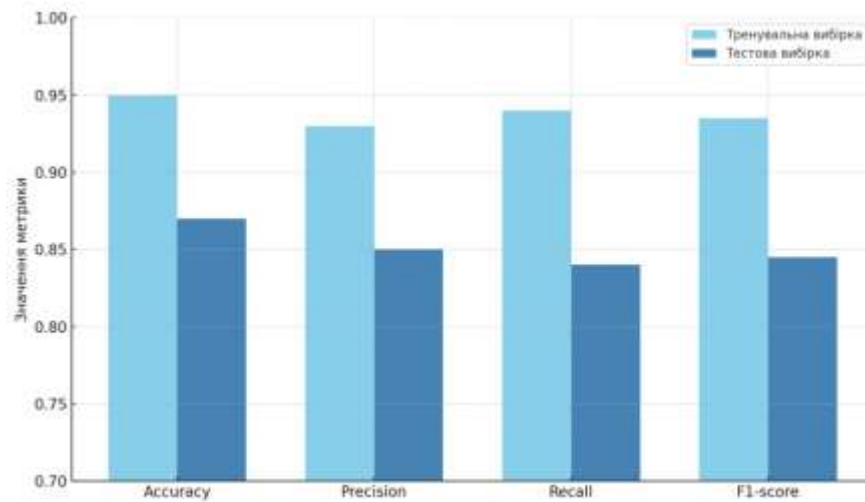


Рисунок 3.8 – Порівняння метрик

На рисунку 3.9 зображено графік функції втрат Cross-Entropy Loss [53] для одного зразка класифікації, де правильний клас позначений як $y = 1$. Cross-Entropy Loss – це функція втрат, що застосовується у задачах багатокласової класифікації. Вона обраховує, наскільки добре модель передбачає правильний клас. Чим ближча ймовірність до 1 для правильного класу – тим менша втрата. По осі X відкладено ймовірність, яку модель призначає правильному класу, а по осі Y – відповідне значення втрат.

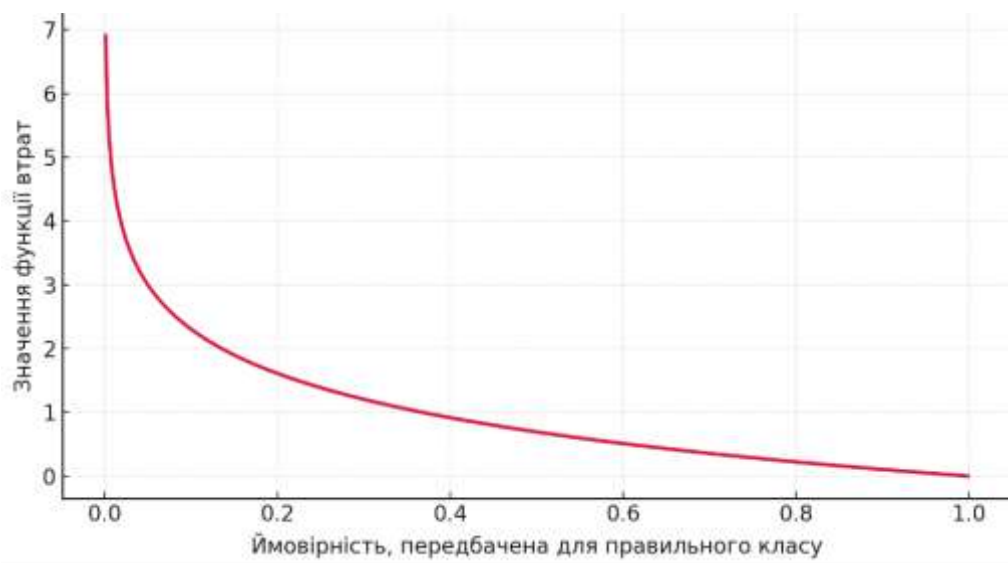


Рисунок 3.9 – Графік функції втрат Cross-Entropy

За графіком видно, що коли передбачена ймовірність правильного класу низька (близька до 0), значення функції втрат стрімко зростає. Натомість, якщо ймовірність близька до 1, втрата прямує до нуля. Така властивість функції змушує модель активно зменшувати вплив неправильних передбачень і точніше налаштовувати ваги під час навчання. Цей тип функції ідеально підходить для задачі багатокласової класифікації, яку реалізовано у межах даної кваліфікаційної роботи.

З метою оцінки ефективності розробленого методу розпізнавання книжкових обкладинок було проведено порівняння з моделлю, описаною у науковій статті [17], розглянутій у першому розділі (табличка 3.2).

Таблиця 3.2 – Порівняння метрик

Метрика	Запропонований метод	Підхід [17]
Accuracy	0.87	0.87
Precision	0.85	–
Recall	0.84	–
F1-score	0.845	0.84

За результатами валідації, автори досягли accuracy на рівні 0.87 та F1-score 0.84, що свідчить про високу здатність моделі розпізнавати емоції на обличчі. Проте, точні значення precision та recall не були наведені.

У порівнянні з вищезазначеною моделлю, розроблена система для класифікації книжкових обкладинок демонструє accuracy 0.87, F1-score 0.845, а також додатково обчислені precision 0.85 та recall 0.84. Варто зазначити, що розроблена модель працює у складніших умовах багатокласової класифікації з 30 жанрами, що потребує вищої узагальнюваності та складнішої архітектури.

Таким чином, у порівнянні з дослідженням, орієнтованим на емоційне розпізнавання, запропонований підхід демонструє співставні метрики якості, проте працює з більшою кількістю класів і забезпечує ширший функціонал. Це свідчить про практичну доцільність запропонованої моделі в задачах

автоматизованого аналізу візуальної інформації та її подальшого застосування у мобільних інформаційних системах.

Використаний у дослідженні метод глибокого навчання базується на одній із найпопулярніших і найефективніших архітектур для комп'ютерного зору ResNet50, попередньо навченої на великому датасеті ImageNet. Це дозволяє моделі використовувати вже наявні узагальнені ознаки зображень (наприклад, контури, кольорові переходи, структури), які є універсальними для різних типів зображень, зокрема й для обкладинок книг.

Застосування трансферного навчання – це сучасна та широко визнана практика, яка дозволяє ефективно адаптувати потужні моделі до нових завдань з обмеженим набором даних. Крім того, модель була доповнена класифікатором, натренованим спеціально для даного завдання, що дозволило досягти узгоджених і стабільних результатів у межах жанрової класифікації.

Отримані метрики підтверджують, що модель не є випадковою або некоректною. Результати моделі свідчать про її ефективне навчання та здатність до узагальнення.

Запропонований підхід є коректним, обґрунтованим і не поступається загальноприйнятим сучасним методам розпізнавання зображень. Порівняно з альтернативними підходами, ResNet50 із трансферним навчанням забезпечує оптимальний баланс між точністю, стабільністю та обчислювальною ефективністю. Він дозволяє досягти високої якості класифікації без потреби у величезному датасеті.

Отримані результати мають практичну цінність у контексті реального застосування технологій комп'ютерного зору у сфері видавничої справи, електронної комерції та бібліотечної справи. Створена модель класифікації жанрів книг на основі обкладинок може бути інтегрована в існуючі або нові цифрові платформи з метою автоматизації, покращення користувацького досвіду та оптимізації внутрішніх процесів.

По-перше, модель може бути застосована для автоматичного сортування або тегування книжок. У великих цифрових архівах чи бібліотеках часто виникає

потреба швидко орієнтуватися в жанровому різноманітті, особливо коли відсутні метадані або вони є неповними.

По-друге, така модель є корисною при попередній класифікації нових видань, які завантажуються до цифрових бібліотек або сервісів самообслуговування видавництва. Це суттєво пришвидшує процес публікації та допомагає уникнути помилок класифікації з боку користувачів.

По-третє, модель має потенціал бути інтегрованою в системи рекомендацій, особливо у сфері книжкової електронної комерції (наприклад, онлайн-магазини книг, маркетплейси, мобільні застосунки). На основі жанру, визначеного автоматично, можна формувати персоналізовані пропозиції для читачів.

У результаті експериментального дослідження було створено і протестовано модель класифікації обкладинок книг за жанрами з використанням ResNet50. Модель показала хороші результати, продемонструвала стабільність у навчанні та здатність до узагальнення. Отримані дані підтверджують, що метод глибокого навчання на основі зображень є перспективним і може використовуватись як самостійно, так і в комплексі з іншими джерелами інформації для підвищення точності класифікації.

3.3 Висновки до розділу 3

У даному розділі було описано інформаційну систему, а саме створений мобільний застосунок для iOS, що виконує розпізнавання жанру книги за зображенням її обкладинки. Архітектура рішення побудована на базі попередньо натренованої моделі глибокого навчання ResNet50, адаптованої для використання у мобільному середовищі. Для реалізації було використано фреймворки Swift та Core ML, а також засоби передобробки зображень, отриманих із камери або галереї пристрою. Застосунок дає змогу завантажити або зробити фото обкладинки книги, після чого виконується автоматична класифікація жанру з виведенням результатів на екран.

Інтерфейс користувача реалізовано у формі сучасного мобільного застосунку з інтуїтивно зрозумілою навігацією. Основні екрани включають: головну сторінку з кнопкою запуску камери, сторінку з історією розпізнаних книг. Результати класифікації відображаються у зручному вигляді. Такий підхід забезпечує зручність у використанні застосунку як для звичайного користувача, так і для бібліотекарів, книготорговців або співробітників видавництва.

У ході експериментального дослідження було реалізовано і протестовано модель глибокого навчання для автоматичної класифікації книжкових обкладинок за жанрами. Для цього було використано попередньо навчений класифікатор ResNet50 з донавчанням на спеціалізованому датасеті, що охоплює 30 жанрових категорій. З метою забезпечення більшої функціональності, до моделі було інтегровано компонент OCR, який дозволяє автоматично витягувати назву книги, автора та рік видання з зображення обкладинки.

Якість роботи системи підтверджується за допомогою матриці помилок, яка демонструє високу точність класифікації – найбільші значення зосереджені вздовж діагоналі, що свідчить про правильне розпізнавання переважної більшості зразків.

Метод, імплементований в застосунок, досягає високих показників якості класифікації: *accuracy* – 87%, *precision* – 85%, *recall* – 84%, *F1-score* – 85%. Ці результати свідчать про те, що модель успішно навчилася розрізняти книжкові жанри на основі зображень обкладинок, і демонструє збалансовану якість передбачень.

Для порівняння результатів було проаналізовано наукову статтю, у якій представлена модель згорткової нейронної мережі для класифікації емоцій на зображеннях обличчя. Незважаючи на меншу кількість класів (7 емоцій), модель досягає *accuracy* 0.87 та *F1-score* 0.84. Проте точні значення *precision* та *recall* у дослідженні не наводяться. Варто звернути увагу, що запропонована модель для ідентифікації зображень демонструє практично такі ж результати, хоча працює з ускладненою задачею, і при цьому додатково реалізує текстову обробку зображення за допомогою OCR.

Отже, результати показують ефективність використання попередньо натренованих моделей для вирішення прикладних задач комп'ютерного зору та відкривають перспективи для подальших досліджень у сфері аналізу візуального контенту у книжковій індустрії.

Більше того, у результаті було успішно реалізовано мобільний застосунок для розпізнавання та класифікації зображень обкладинок друкованих видань. Застосунок є прикладом ефективного поєднання глибокого навчання з мобільними технологіями, має практичне значення для книжкової галузі та відкриває перспективи для подальшого розвитку, зокрема у напрямку мультимодальної обробки даних та розширення функціональності.

Загальні висновки

У процесі виконання кваліфікаційної роботи бакалавра було досягнуто поставленої мети – створення мобільного застосунку, який допомагає у підборі друкованої інформації шляхом аналізу зображень обкладинок книг за допомогою методів глибокого навчання. Розроблений застосунок забезпечує користувача швидким доступом до потрібної інформації про видання (назва, автор, жанр, рік), а також дозволяє отримати персоналізовані рекомендації, зокрема відповідно до рівня володіння англійською мовою, що значно оптимізує процес пошуку літератури.

На першому етапі було проаналізовано предметну область автоматизованого розпізнавання об'єктів за візуальними ознаками та вивчено сучасні підходи до класифікації зображень, серед яких найбільше застосування знайшли згорткові нейронні мережі. Було здійснено порівняння із подібними існуючими рішеннями, зокрема з мобільними додатками Vivino, Seek та сервісом Imagga.

У процесі розробки методу важливу роль відіграло попереднє проектування архітектури системи, яке було виконане з використанням UML-діаграм. Це дозволило наочно структурувати логіку функціонування інформаційної системи, забезпечити прозорість її побудови, а також спростити подальшу реалізацію та тестування.

Загальна архітектура системи була описана за допомогою діаграм варіантів використання, що дало змогу визначити ключові сценарії взаємодії користувача із застосунком. Для кращого розуміння послідовності дій усередині програми було розроблено діаграму діяльності, яка ілюструє послідовність кроків – від моменту завантаження або знімання фотографії до отримання результату класифікації та рекомендацій. Крім того, діаграма послідовності дозволила описати взаємодію між внутрішніми компонентами системи у часі. У ній представлено послідовну роботу модулів. Для демонстрації змін у стані застосунку залежно від дій користувача було складено діаграму станів. Для моделювання внутрішньої

структури програми була застосована діаграма класів, яка визначає основні сутності. Кожен клас має свої атрибути та методи, які описують як логіку взаємодії з даними, так і правила відображення інформації на інтерфейсі користувача.

Таким чином, використання UML-діаграм дозволило систематизувати етапи розробки, забезпечити структуровану реалізацію програмних компонентів та підвищити ефективність створення інформаційної системи.

У ході кваліфікаційної роботи було розроблено метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання, побудований на основі попередньо навченої моделі ResNet50, яку адаптовано до задачі багатокласової класифікації за 30 жанрами. Для підвищення функціональності модель доповнено OCR-модулем, що дозволяє автоматично витягати текстову інформацію з обкладинок, зокрема назву книги, автора та рік видання. Далі цей метод було інтегровано у мобільний застосунок, створений у середовищі Xcode із використанням SwiftUI та Core ML. Застосунок дозволяє користувачу сфотографувати або завантажити зображення обкладинки й отримати миттєвий результат з жанровою класифікацією, текстовими атрибутами та мовною рекомендацією відповідно до заданого рівня CEFR (від A1 до C2).

У процесі тестування системи було отримано переконливі результати, які свідчать про високу результативність обраного підходу. Загальна точність класифікації на тестовій вибірці становила 87%, F1-метрика – 0.845, precision – 0.85, recall – 0.84. Матриця плутанини підтверджує, що більшість зразків класифікуються правильно, а помилки переважно трапляються між близькими за змістом жанрами, що є допустимим у багатокласових задачах. Додатково проаналізовано функцію втрат Cross-Entropy, яка була обрана як базова для задачі багатокласової класифікації.

Порівняння з іншими підходами, зокрема з моделлю для розпізнавання емоцій на обличчях, описаною у статті [17], показало, що хоча обидві системи досягають схожих значень точності, запропонована в кваліфікаційній роботі модель вирішує значно складнішу задачу з 30 класами замість 7 та реалізує

додаткову текстову аналітику. Це вказує на високу ефективність і гнучкість розробленої інформаційної системи.

Розроблений мобільний застосунок вирізняється зручністю у використанні, інтуїтивно зрозумілим інтерфейсом та функціональністю, що повністю відповідає поставленим вимогам. У процесі реалізації особлива увага приділялася тому, щоб застосунок був однаково доступним для користувачів різного віку, технічної підготовки та цифрової грамотності. Інтерфейс було створено за допомогою сучасного фреймворку SwiftUI, що дозволило забезпечити чистий дизайн, швидке оновлення вмісту та адаптивність до різних розмірів екранів.

Перспективи розвитку проєкту включають розширення функціональності – підключення до централізованих бібліотек, реалізацію веб-версії, адаптацію під Android-платформу, а також вдосконалення рекомендаційної системи на основі глибшого семантичного аналізу тексту. Система вже сьогодні демонструє повну відповідність поставленим завданням та має потенціал для реального впровадження в освітніх, бібліотечних або комерційних середовищах, де потрібна швидка і точна навігація по книжковому контенту.

Отже, у результаті написання кваліфікаційної роботи бакалавра було досягнуто усіх цілей, виконано всі завдання, створено інформаційну систему для користувачів, що відповідає вимогам, а саме полегшує користувачам отримуючи всю потрібну інформацію про друковане видання за допомогою фото його обкладинки.

Перелік посилань

1. Чи важлива вища освіта?. *vonahub*. URL: <https://vonahub.org.ua/chy-vazhlyva-vyshcha-osvita-ta-iaakumu-ie-shansy/>.
2. Освіта.ua. Вплив освіти на економічне життя суспільства. *Освіта.UA*. URL: <https://osvita.ua/blogs/89832/>.
3. Важливість вивчення англійської мови у сучасному світі - Анастасія Т. Буки | *Вуки - ваш репетитор з будь-якого предмету. Репетитори України*. URL: <https://buki.com.ua/blogs/vazlivist-vivcennya-angliiskoyi-movi-u-sucasnomu-sviti/>.
4. Книги англійською за рівнями: що почитати з користю від A1 до C2. *Grade Education Centre*. URL: <https://grade.ua/uk/blog/knigi-na-angliyskom-dlya-raznykh-urovney-spisok-ot-a1-do-c2/>.
5. Найкращі книги для вивчення англійської мови: ТОП-10 - *JustSchool*. *JustSchool*. URL: <https://justschool.me/uk/blog/najkrashhi-knygy-dlya-vyvchennya-anglijskoyi-movy-top-10/>.
6. Учасники проєктів Вікімедіа. Інформаційна модель – Вікіпедія. *Вікіпедія*. URL: https://uk.wikipedia.org/wiki/Інформаційна_модель.
7. Modern approaches to computer vision / R. M. Tymchyshyn та ін. *Upravlâûsie sistemy i mašyny*. 2018. № 6 (278). С. 46–73. URL: <https://doi.org/10.15407/usim.2018.06.046>.
8. Що таке глибоке навчання? - *Oksim*. *Oksim - IT допомога*. URL: <https://www.oksim.ua/2023/07/28/shho-take-glyboke-navchannya/>.
9. Кольорові моделі RGB та CMYK. *Типографія оперативного друку Фастпринт*. URL: <https://www.fastprint.ua/uk/rgb-cmyk>.
10. Розмір і роздільна здатність зображення. *Adobe Help Center*. URL: <https://helpx.adobe.com/ua/photoshop/using/image-size-resolution.html>.
11. Застосування коригування яскравості/контрасту. *Adobe Help Center*. URL: <https://helpx.adobe.com/ua/photoshop/using/apply-brightness-contrast-adjustment.html>.

12. Модуль "Графічний дизайн" - Урок 7. Характеристики зображення та засобів його відтворення - яскравість. *Google Drive: Sign-in*. URL: <https://sites.google.com/view/kcml33m/растрова-графіка/урок-7-характеристики-зображення-та-засобів-його-відтворення-яскравість>.
13. Учасники проектів Вікімедіа. Завдання розпізнавання образів – Вікіпедія. *Вікіпедія*. URL: https://uk.wikipedia.org/wiki/Завдання_розпізнавання_образів.
14. Що таке оптичне розпізнавання символів (OCR). *KONICA MINOLTA Ukraine – надійний IT партнер на шляху до трансформацій | KONICA MINOLTA*. URL: <https://www.konicaminolta.ua/uk-ua/rethink-work/tools/what-is-optical-character-recognition-ocr>.
15. Zhrebukh O., Farmaha I. Using Neural Networks to Identify Objects in an Image. *Computer Design Systems. Theory and Practice*. 2024. Т. 6, № 1. С. 232–240. URL: <https://doi.org/10.23939/cds2024.01.232>.
16. Сила брендингу: Чому ваш логотип має більше значення, ніж ви думаєте. *Ranktracker*. URL: <https://www.ranktracker.com/uk/blog/the-power-of-branding-why-your-logo-matters-more-than-you-think/>.
17. Artificial intelligence and image: fake, addition or reality? (Ukr) / О. D. Dolhanenko та ін. *Scientific Works of Vinnytsia National Technical University*. 2024. № 3. URL: <https://doi.org/10.31649/2307-5376-2024-3-17-24>.
18. Paramud Y., Yarkun V. Method of image symbol recognition on the basis of convolutional neural network. *Computer systems and network*. 2018. Т. 1, № 905. С. 96–106. URL: <https://doi.org/10.23939/csn2018.905.096>.
19. Коломoeць С. Застосування штучного інтелекту в розпізнаванні медичних зображень. *Information technology and society*. 2024. № 3 (14). С. 23–28. URL: <https://doi.org/10.32689/maup.it.2024.3.3>.
20. Холявкіна Т. В., Резаєв Я. О., Харченко О. О. *Deep learning based image recognition system with neural network architecture. Science-based technologies*. 2020. Т. 45, № 1. URL: <https://doi.org/10.18372/2310-5461.45.14572>.

21. Deep learning technology of convolutional neural networks for facial expression recognition / D. V. Petrosiuk etc. *Applied Aspects of Information Technology*. 2021. Т. 4, № 2. С. 192–201. URL: <https://doi.org/10.15276/aait.02.2021.6>.
22. Matviichuk A. M. Analysis of neural network architectures for recognition of handwritten numbers. *Telecommunication and Information Technologies*. 2023. Т. 81, № 4. URL: <https://doi.org/10.31673/2412-4338.2023.041827>.
23. Afonin A., Kundik K. Automatic determination of agricultural plant diseases. *NaUKMA Research Papers. Computer Science*. 2021. Т. 4. С. 23–28. URL: <https://doi.org/10.18523/2617-3808.2021.4.23-28>.
24. Imagga: API для розпізнавання зображень на основі штучного інтелекту | Martech Zone. *Martech Zone*. URL: <https://uk.martech.zone/imagga-image-recognition-api/>.
25. Image Recognition API, Computer Vision AI – Imagga. *Image Recognition API, Computer Vision AI – Imagga*. URL: <https://imagga.com>.
26. Seek by iNaturalist · iNaturalist. *iNaturalist*. URL: https://www.inaturalist.org/pages/seek_app.
27. iNaturalist. Seek від iNaturalist – Додатки в Google Play. *Android Apps on Google Play*. URL: <https://play.google.com/store/apps/details?id=org.inaturalist.seek&hl=uk>.
28. Учасники проєктів Вікімедіа. Google Об'єктив – Вікіпедія. *Вікіпедія*. URL: https://uk.wikipedia.org/wiki/Google_Об'єктив.
29. LLC G. Google Lens – Додатки в Google Play. *Android Apps on Google Play*. URL: <https://play.google.com/store/apps/details?id=com.google.ar.lens&hl=uk>.
30. Офіційний сайт «Сільпо». ▶ «Сільпо» – онлайн-супермаркет | Доставка продуктів до дверей. URL: <https://silpo.ua/about/yak-vstanovyty-zastosunok-vivino>.
31. ApS V. Vivino: Buy the Right Wine – Додатки в Google Play. *Android Apps on Google Play*. URL: <https://play.google.com/store/apps/details?id=vivino.web.app&hl=uk>.
32. Що таке згорткова нейронна мережа, CNN і для чого вона використовується. *Evergreen - web розробка і діджиталізація бізнесу за допомогою AI продуктів*. URL: <https://evergreens.com.ua/ua/articles/cnn.html>.

33. Driver Distraction Classification Using Deep Convolutional Autoencoder and Ensemble Learning / A. Muthuswamy et al. *IEEE Access*. 2023. P. 1. URL: <https://doi.org/10.1109/access.2023.3293110>.
34. Учасники проєктів Вікімедіа. Згорткова нейронна мережа – Вікіпедія. *Вікіпедія*. URL: https://uk.wikipedia.org/wiki/Згорткова_нейронна_мережа.
35. Keras - бібліотека для нейронних мереж – IT Master - електроніка та програмування. *Головна – IT Master - електроніка та програмування*. URL: <https://itmaster.biz.ua/programming/vision/keras.html>.
36. Tensorflow що це і які основні можливості застосування. *FoxmindEd*. URL: <https://foxminded.ua/tensorflow-shcho-tse/>.
37. TensorFlow: What is it and How Does it Work?. *Free Chatbot maker / Chatbot for Website, WhatsApp / BotPenguin*. URL: <https://botpenguin.com/glossary/tensorflow>.
38. Unlock the Power of Fine-Tuning Pre-Trained Models in TensorFlow & Keras. *LearnOpenCV – Learn OpenCV, PyTorch, Keras, Tensorflow with code, & tutorials*. URL: <https://learnopencv.com/fine-tuning-pre-trained-models-tensorflow-keras/>.
39. Довідник по Machine Learning – Data Augmentation | База знань IT. *База знань IT технологій*. URL: <https://itwiki.dev/data-science/ml-reference/ml-glossary/data-augmentation>.
40. Учасники проєктів Вікімедіа. Діаграма класів – Вікіпедія. *Вікіпедія*. URL: https://uk.wikipedia.org/wiki/Діаграма_класів.
41. Моралес Дж. UML Activity Diagram: Definition and Tutorial to Create. *MindOnMap | Free Mind Mapping Tool to Draw Ideas Easily Online*. URL: <https://www.mindonmap.com/uk/blog/uml-activity-diagram/>.
42. Махум Z. Діаграма послідовності (Sequence Diagrams). *Махум Zosym*. URL: <https://www.maxzosim.com/sequence-diagrams/>.

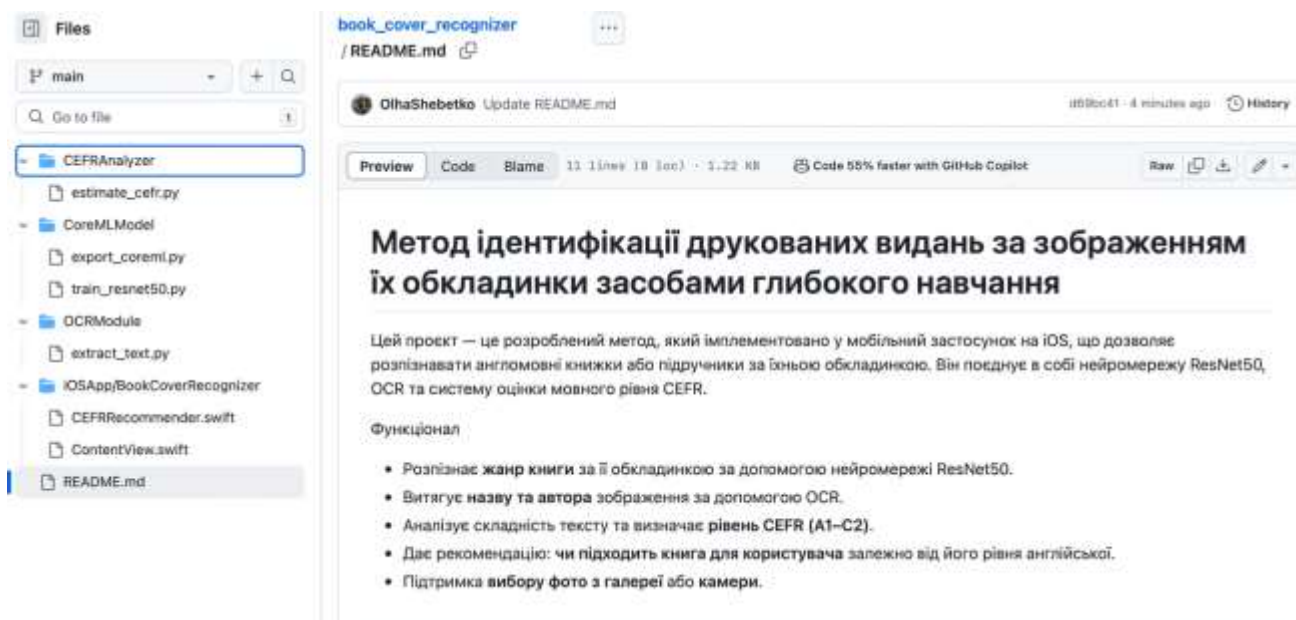
43. Моралес Дж. In-depth Knowledge of UML Use Case Diagram: with Tutorial. *MindOnMap | Free Mind Mapping Tool to Draw Ideas Easily Online*. URL: <https://www.mindonmap.com/uk/blog/what-is-a-uml-use-case-diagram/>.
44. Махум Z. Моделювання станів (State Modelling). *Махум Zosym*. URL: <https://www.maxzosim.com/state-modelling/>.
45. Що таке дизайн інтерфейсу користувача (UI дизайн) | BLOG Dizz Agency ➤ Dizz.in.ua. *DIZZ*. URL: <https://dizz.in.ua/uk/shho-take-dizajn-interfejsu-koristuvacha-ui-dizajn/>.
46. UX дослідження: Що це і чому вони важливі?. *FoxmindEd*. URL: <https://foxminded.ua/ux-doslidzhennia-shcho-tse/>.
47. Учасники проектів Вікімедіа. SwiftUI – Вікіпедія. *Вікіпедія*. URL: <https://uk.wikipedia.org/wiki/SwiftUI>.
48. Онбординг в мобільному додатку: за і проти. *eSputnik*. URL: <https://esputnik.com/uk/blog/onboarding-u-mobilnomu-dodatku-za-i-proti>.
49. Apple CoreML – Вікіпедія. *Вікіпедія*. URL: https://uk.wikipedia.org/wiki/Apple_CoreML.
50. Про програмне забезпечення XCODE. *File Format Docs*. URL: <https://docs.fileformat.com/uk/programming/xcode/>.
51. Bondarenko O. Матриця невідповідностей. *Medium*. URL: <https://medium.com/stinoryps/матриця-невідповідностей-329e7e4bf05e>.
52. Класифікація: точність, повнота, влучність і пов'язані метрики | Machine Learning | Google for Developers. *Google for Developers*. URL: <https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall?hl=uk>.
53. Судорженко А. Функція втрат у задачах класифікації зображень. *DSPACE Repository :: Electronic Kyiv-Mohyla Academy Institutional Repository*. URL: <https://ekmair.ukma.edu.ua/server/api/core/bitstreams/3c1fe679-0e02-4bf6-a5fa-33499b58f6fe/content>.

ДОДАТКИ

Додаток А

Програмні коди

https://github.com/OlhaShebetko/book_cover_recognizer.git



У папці CoreMLModel є два файли: train_resnet50.py який реалізує навчання ResNet50, та export_coreml.py – експорт з допомогою CoreML.

Папка OCRModule містить файл extract_text.py, що реалізує розпізнавання тексту з зображення через EasyOCR.

У папці CEFRAnalyzer файл estimate_cefr.py реалізує логіку оцінки рівня книги.

У папці iOSApp/ BookCoverRecognizer представлені файли проекту реалізовані у Xcode. У файлі ContentView.swift реалізований основний SwiftUI-інтерфейс. У файлі CEFRRecommender.swift реалізовано порівняння рівня користувача та книги.

Додаток Б

Презентаційний матеріал

КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

Метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання

Виконала: студентка групи КН-21-1 Ольга ШЕБЕТКО

Керівник: д.т.н., зав. кафедри КН Олександр БАРМАК

Актуальність теми



- Зростання кількості друкованих видань
- Проблема пошуку і систематизації книг за зображенням
- Значення автоматизації процесу ідентифікації в освіті, бібліотеках, онлайн-сервісах

Мета та завдання дослідження

Мета – допомога у підборі друкованої інформації через аналіз зображень їх обкладинок за допомогою відповідного методу штучного інтелекту імплементованого у мобільне застосування.

Завдання:

1. Провести аналіз предметної області.
2. Проаналізувати існуючі підходи до класифікації зображень.
3. Розробити метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання.
4. Інтегрувати метод у мобільний застосунок.
5. Провести тестування розробленого застосунку та оцінити точність розпізнавання.

2

Схема навчання моделі

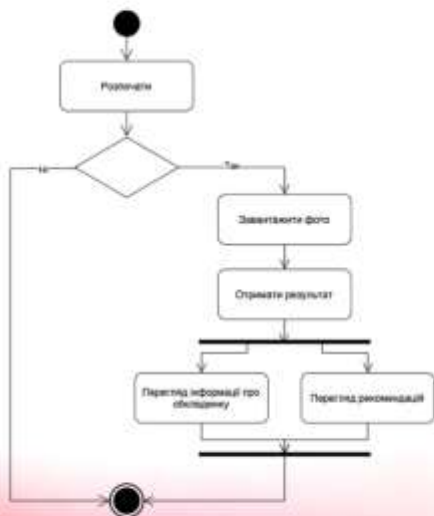


Схема використання моделі

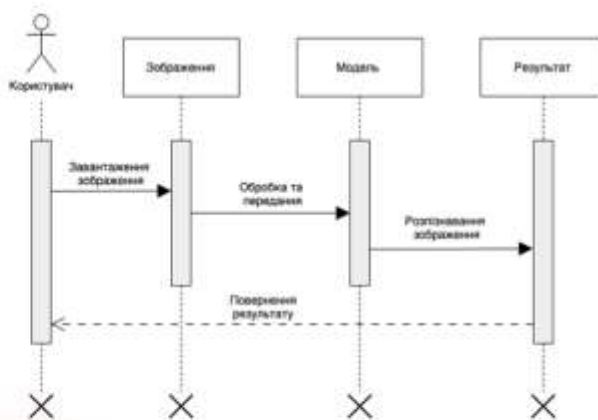


3

Діаграма активності

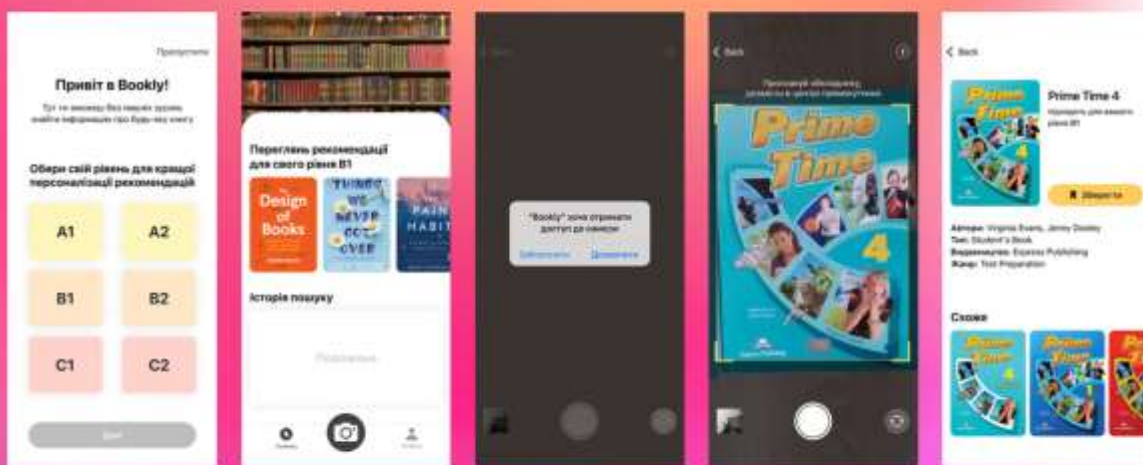


Діаграма послідовності



4

Інтерфейс системи

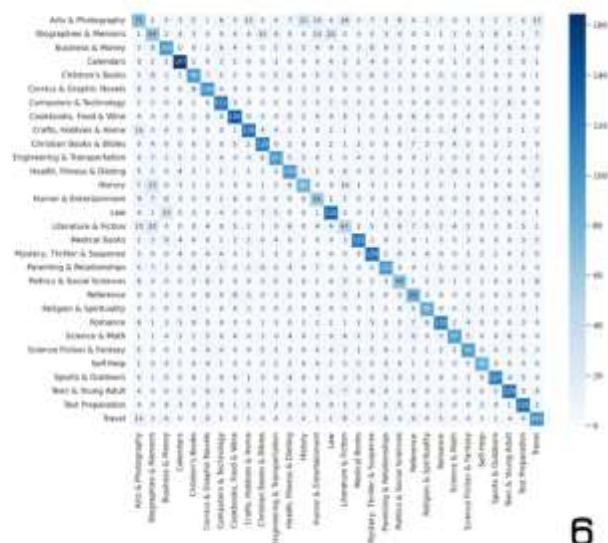


5

Оцінка ефективності

Метрика	Тренувальна вибірка	Тестова вибірка
Accuracy	0.95	0.87
Precision	0.93	0.85
Recall	0.94	0.84
F1-score	0.935	0.845

Матриця помилок



6

Порівняння з моделлю, описаною у науковій статті

Метрика	Запропонований метод	Підхід зі статті
Accuracy	0.87	0.87
Precision	0.85	—
Recall	0.84	—
F1-score	0.845	0.84

*значення precision та recall не були наведені у статті



7

Висновки

У результаті виконання кваліфікаційної роботи бакалавра було розроблено метод ідентифікації друківаних видань за зображенням їх обкладинок із застосуванням засобів глибокого навчання. Основу системи становить згортова нейронна мережа ResNet50. Для отримання текстової інформації з обкладинок було інтегровано технологію OCR.

Імплементация методу у мобільній застосунок для iOS була зроблена за допомогою CoreML. ІС показала практичну доцільність обраного підходу. Користувачі можуть за декілька секунд отримати структуровану інформацію про книгу, просто навівши камеру на її обкладинку. Система також дозволяє надати рекомендації відповідно до мовного рівня читача.

Проведені експериментальні дослідження показали, що метод демонструє високу точність класифікації та здатність ефективно розпізнавати основну текстову інформацію. Загалом, розроблений метод є кроком до сучасної цифрової автоматизації процесів ідентифікації та обробки друківаних видань.



8

Дякую за увагу!



Anti-Plagiarism (UA) v-15.281 Educational

The maximum coincidence with one document 4.0%

Dictionary check: en_US, ru_RU, ua_UA. **Errors in the documents: 12%**

ID: 244110 Title: КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА на тему Метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання Added in a DB: 2025-06-08 Authors: Ольга ШЕБЕТКО Heads: Олександр БАРМАК Consultants: Opponents:	Document		Sum coincidence on the DB	
	Symbols	Lexemes	Symbols	Lexemes
	63489	973	3961 (6%)	66 (7%)

Plagiarism sources

ID	Description	Plagiarism presence in the document	
		Symbols	Lexemes

Протокол аналізу звіту подібності науковим керівником

Заявляю, що я ознайомився (-лась) з Повним звітом подібності, який був згенерований Системою виявлення і запобігання плагіату щодо роботи:

Автор: Ольга ШЕБЕТКО

Співавтор:

Назва: КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА на тему Метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання

Науковий керівник: Олександр БАРМАК, д.т.н., проф.

Підрозділ: Кафедра комп'ютерних наук

Коефіцієнт подібності 1: 7.1%

Коефіцієнт подібності 2: 2.7%

Мікропробіли: 0

Заміна букв: 4

Інтервали: 0

Білі знаки: 139

Дата створення звіту: 2025-06-08 18:20:02.0

Після аналізу Звіту подібності констатую наступне:

Запозичення, виявлені в роботі є законними і не є плагіатом. Рівень подібності не перевищує допустимої межі. Таким чином робота незалежна і приймається.

Запозичення не є плагіатом, але перевищено граничне значення рівня подібностей. Таким чином робота повертається на доопрацювання.

Виявлено запозичення і плагіат або навмисні текстові спотворення (маніпуляції), як передбачувані спроби укриття плагіату, які роблять роботу невідповідною вимогам законодавства (Ст. 32. ЗУ Про вищу освіту, пункт 3.1, Ст. 42. ЗУ Про освіту) та вимог НАЗЯВО (Критерій 5), а також кодексу етики і процедурам. Таким чином робота не приймається.

Обґрунтування:

2025-06-08

Дата

експерт

Львівський Р. Р.

РІШЕННЯ ЕКСПЕРТНОЇ КОМПІСІЇ КАФЕДРИ КОМП'ЮТЕРНИХ НАУК

ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Назва кваліфікаційної роботи Метод ідентифікації друктованих видань за зображенням їх обкладинки засобами глибокого навчання

Автор студент групи КН-21-1 Ольга Шебетко

Освітня програма Комп'ютерні науки

Рівень вищої освіти перший (бакалаврський)

Спеціальність 122 – Комп'ютерні науки

Науковий керівник: д.т.н., проф., зав. каф. комп'ютерних наук Олександр БАРМАК

На основі аналізу кваліфікаційної роботи на дотримання вимог академічної доброчесності (у т.ч. відсутності ознак академічного плагіату) з урахуванням результатів перевірки роботи спеціалізованим програмними засобами комісія зробила такий висновок:

№	Висновок	Позначка про відповідність
1	Ознаки академічного плагіату	
1.1	Запозичення, виявлені в роботі, є законними і не є академічним плагіатом (далі – зазначаються підстави віднесення запозичень до правомірних, якщо потрібно). Робота приймається до захисту.	<i>відповідає</i>
1.2	Виявлені запозичення не є академічним плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи (далі – зазначаються детальні та аргументовані підстави віднесення запозичень до правомірних). Робота приймається до захисту, але має бути відкоригована.	
1.3	Виявлені запозичення не є академічним плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота може бути допущена до захисту після того як буде відкоригована та доопрацьована і успішно пройде повторну перевірку на академічний плагіат.	
1.4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття текстових запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	
2	Інші види порушень академічної доброчесності	<i>відсутні</i>

Підтвердження:

Запозичення, виявлені в роботі Ольги Шебетко, не є плагіатом, оскільки: запозичення розміщені в розділі огляду існуючих підходів, не описують безпосередньо авторську роботу і не стосуються її результатів; усі запозичення фрагментарні; до запозичень входять фрагменти, які не мають авторства і містять поширені конструкції та загальновідомі терміни, скорочення. Рівень подібності не перевищує допустимої межі. Таким чином, робота є законною та приймається до захисту.

Обсяг запозичень, визначений системами виявлення збігів/ідентичності/схожості:

- за системою Anti-Plagiarism: 4%;

- за системою StrikePlagiarism КП1: 7.1%, КП2: 2.7%.

08.06.2025

Завідувач кафедри



Олександр БАРМАК

Гарант освітньої програми



Олександр МАЗУРЕЦЬ

Керівник кваліфікаційної роботи



Олександр БАРМАК



ВІДГУК НАУКОВОГО КЕРІВНИКА
на кваліфікаційну роботу бакалавра

студента *гр. КН-21-1 Шебетко Ольги Володимирівни*

за темою Метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання

1. Актуальність теми

Актуальним завданням, яке досліджується у даній роботі, є розробка методу ідентифікації друкованих видань за зображенням їх обкладинки з використанням технологій глибокого навчання. Зростання обсягів друкованої продукції, ускладнення пошуку релевантної інформації та потреба в автоматизованій обробці візуальних даних зумовлюють необхідність створення інструментів, які здатні розпізнавати та класифікувати друковані видання без участі людини. Особливої актуальності тема набуває в умовах стрімкого розвитку цифрових бібліотек, онлайн-освіти та потреби у швидкому доступі до інформаційних джерел. Розробка такого методу є актуальною задачею у сфері комп'ютерних наук, зокрема в галузі комп'ютерного зору та штучного інтелекту.

2. Відповідність роботи предметній області Стандарту спеціальності 122 Комп'ютерні науки

Згідно зі стандартом спеціальності 122 Комп'ютерні науки, а саме описом предметної області, об'єктами дослідження та діяльності є математичні, інформаційні, імітаційні моделі реальних явищ, об'єктів, систем і процесів, а також методи і технології отримання, зберігання, обробки, передачі та використання інформації. Метою кваліфікаційної роботи є розробка методу ідентифікації друкованих видань за зображенням їх обкладинки, що базується на використанні глибоких нейронних мереж, комп'ютерного зору та засобів обробки зображень. У процесі виконання роботи застосовано математичні моделі, алгоритми глибокого навчання (архітектуру ResNet50), а також OCR-технології. Отримані результати дозволяють вирішити прикладну задачу автоматичного розпізнавання та класифікації об'єктів, що відповідає вимогам і компетентностям, визначеним Стандартом бакалавра спеціальності 122 – Комп'ютерні науки.

3. Професійні та особистісні якості бакалавра

Під час роботи над кваліфікаційною роботою Шебетко Ольга Володимирівна проявила себе як відповідальна, ініціативна та цілеспрямована студентка. Вона своєчасно і якісно виконувала всі етапи дослідження, демонструючи високий рівень самостійності, дисциплінованість та відповідальне ставлення до поставлених завдань. У процесі написання пояснювальної записки та розробки програмного забезпечення студентка показала високий рівень професійних і soft skills, необхідних для успішного виконання завдання.

4. Ступінь самостійності під час виконання кваліфікаційної роботи

Результати, отримані в роботі, є результатом самостійної діяльності студентки. Усі поставлені задачі, включаючи аналіз літератури, розробку методу, програмну реалізацію та тестування, були виконані студенткою самостійно.

5. Ступінь оволодіння методами дослідження

Студентка продемонструвала високий рівень володіння методами дослідження у сфері комп'ютерних наук, зокрема методами обробки зображень, алгоритмами глибокого навчання та інструментами програмування. У процесі виконання роботи були застосовані сучасні методика та технології, що відповідають предметній області.

6. Повнота та якість розкриття теми роботи

Тема роботи розкрита повною мірою. Проведено ґрунтовний аналіз актуальності теми, оглянуто сучасні дослідження у сфері комп'ютерного зору та глибокого навчання, чітко сформульовано завдання та розроблено програмне забезпечення для їх реалізації. Результати роботи підтверджені експериментальною перевіркою та валідацією запропонованого методу.

7. Логічність, послідовність, аргументованість, літературна грамотність викладення матеріалу

Структура роботи є логічною, послідовною та відповідає поставленій меті. Виклад матеріалу чіткий, аргументований, літературно грамотний, що відповідає вимогам до написання кваліфікаційної роботи.

8. Можливість практичного застосування кваліфікаційної роботи бакалавра, окремих її частин

Розроблений у роботі метод ідентифікації друкованих видань за зображенням обкладинки може бути використаний у бібліотечних системах, онлайн-книгарнях, системах автоматизованого розпізнавання та інших сферах, де потрібна швидка та точна ідентифікація книг. Це сприятиме спрощенню пошуку літератури та покращенню організації інформаційних ресурсів.

9. Висновок про можливість допуску кваліфікаційної роботи бакалавра до захисту, на яку оцінку заслуговує робота

Звертаючи увагу на високий рівень виконання роботи, повноту розкриття теми, відповідність стандартам спеціальності, кваліфікаційна робота може бути допущена до захисту. Рекомендована оцінка – «відмінно».

Керівник



д.т.н., проф. зав. каф. КН Олександр БАРМАК



РЕЦЕНЗІЯ

на кваліфікаційну роботу бакалавра

студента *гр. КН-21-1 Шебетко Ольги Володимирівни*

за темою: Метод ідентифікації друкованих видань за зображенням їх обкладинки засобами глибокого навчання

1. Актуальність обраної теми

Актуальність теми кваліфікаційної роботи є високою, враховуючи зростаючу потребу в автоматизації обробки зображень у бібліотечних системах, електронній комерції та інформаційних системах. Розробка методу ідентифікації друкованих видань за зображенням їх обкладинки з використанням глибокого навчання сприяє підвищенню ефективності розпізнавання книг, що є актуальною задачею в галузі комп'ютерного зору та штучного інтелекту.

2. Повнота розкриття мети та завдань роботи

Під час виконання кваліфікаційної роботи студентка Шебетко Ольга Володимирівна повністю розкрила мету роботи – створення ефективного методу ідентифікації друкованих видань за зображенням їх обкладинок засобами глибокого навчання. Проведено оцінку ефективності розробленого методу, що підтверджує досягнення поставленої мети. Усі завдання, визначені у роботі, виконано повною мірою, включаючи аналіз літератури, розробку моделі, її програмну реалізацію та тестування.

3. Зміст кожного розділу роботи

У першому розділі проведено аналіз сучасних підходів до ідентифікації друкованих видань за зображенням обкладинок, розглянуто наукові джерела, методи комп'ютерного зору та глибокого навчання, що застосовуються в подібних задачах. Другий розділ присвячено розробці методу ідентифікації: обґрунтовано вибір архітектури нейронної мережі, описано підготовку даних, етапи навчання моделі та використання OCR для вилучення текстової інформації. У третьому розділі представлено програмну реалізацію створеної системи, описано архітектуру застосунку, логіку обробки зображень, інтерфейс користувача та проведено тестування, що підтверджує ефективність розробленого рішення. Усі розділи логічно пов'язані та послідовно розкривають тему дослідження.

4. Оцінка розробленого методу та його практична цінність

Розроблений метод ідентифікації друкованих видань є практичним для використання в автоматизованих бібліотечних системах, онлайн-книгарнях та інших інформаційних системах, що потребують швидкого та точного розпізнавання обкладинок книг. Метод демонструє високу точність і може бути легко інтегрований у реальні системи. Мобільний застосунок має інтуїтивно зрозумілий інтерфейс, придатний для використання як кінцевими користувачами, так і адміністраторами бібліотечних або освітніх платформ.

5. Якість оформлення кваліфікаційної роботи бакалавра

Кваліфікаційна робота відповідає всім вимогам щодо оформлення. Тема дослідження розкрита повною мірою, структура роботи логічна, а викладення матеріалу послідовне.

аргументоване та літературно грамотне. Використана термінологія відповідає стандартам предметної області комп'ютерних наук.

6. Недоліки кваліфікаційної роботи бакалавра

Одним із недоліків є обмежена платформенність розробленого застосунку – на даному етапі він реалізований лише для iOS. Це унеможлиблює його використання користувачами пристроїв на базі Android, що обмежує потенційне охоплення аудиторії. Подальше розширення підтримки на інші платформи підвищить практичну цінність системи.

7. Загальний висновок (допускається чи не допускається до захисту), та оцінка на яку заслуговує кваліфікаційна робота.

Враховуючи високий рівень виконання, повноту розкриття теми та відповідність усім необхідним вимогам, кваліфікаційна робота може бути допущена до захисту. Рекомендована оцінка – «відмінно».

Рецензент



Богданець І.І.