


КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

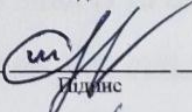
на тему Метод нейромережевого розпізнавання жестів у відеопотоці

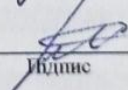
Галузь знань 12 – Інформаційні технології
Шифр і назва галузі знань

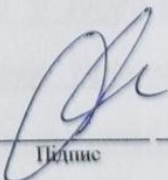
Спеціальність 122 – Комп'ютерні науки
Шифр і назва спеціальності

Освітня програма Комп'ютерні науки
Назва освітньої програми

Виконала: студентка групи КН-20-1  Ангеліна БОГДАНОВА
Група виконавця Ідентифікаційне Ім'я, ПРІЗВИЩЕ

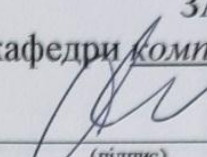
Керівник: к.т.н., доц. каф. КН  Олександр МАЗУРЕЦЬ
Науковий ступінь, посада Ідентифікаційне Ім'я, ПРІЗВИЩЕ

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

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

18 червня 2024 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
« 16 » шостою 2024 року

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ БАКАЛАВРА

1. Тема кваліфікаційної роботи бакалавра: «Метод нейромережевого розпізнавання жестів у відеопотоці»

2. Завдання видано студентці Ангеліні БОГДАНОВІЙ
(Ім'я, прізвище)

3. Керівник роботи доцент кафедри КН Олександр МАЗУРЕЦЬ
(посада, ім'я, прізвище)

4. Затверджено наказом університету від « 15 » шостою 2024 р. № 8


5. Дата видачі завдання студенту: « 16 » шостою 2024 р.

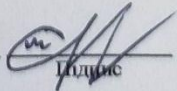
6. Зміст пояснювальної записки (перелік задач) та вихідні дані:

Мета роботи – покращення розпізнавання жестів у відеопотоці. Слід виконати дослідження предметної області розпізнавання жестів у відеопотоці, розробити метод нейромережевого розпізнавання жестів у відеопотоці, на основі розробленого методу виконати проектування структури системи нейромережевого розпізнавання жестів у відеопотоці, здійснити програмну реалізацію системи розпізнавання жестів, провести тестування розробленої системи нейромережевого розпізнавання жестів у відеопотоці та здійснити дослідження ефективності розробленого методу з використанням розробленої системи нейромережевого розпізнавання жестів.

7. Календарний план виконання кваліфікаційної роботи бакалавра:

№	Назва етапів (розділів) кваліфікаційної роботи бакалавра	Термін виконання	Примітка
1	Вибір напряму дослідження та узгодження тематики кваліфікаційної роботи бакалавра з керівником, складання календарного графіка виконання роботи	січень 2024	виконано
2	Ознайомлення з предметною областю, формулювання мети та задач дослідження, визначення об'єкта та предмета дослідження	лютий 2024	виконано
3	Проектування та розробка загальної архітектури програмного забезпечення, інтерфейсу користувача, вибір засобів реалізації програмного забезпечення	березень 2024	виконано
4	Створення та тестування програмного забезпечення	квітень 2024	виконано
5	Написання пояснювальної записки, урахування зауважень керівника, оформлення згідно вимог	травень 2024	виконано
6	Розробка презентаційних матеріалів та попередній захист кваліфікаційної роботи	травень 2024	виконано
7	Отримання відгуку керівника, рецензії, перевірка на плагіат, нормоконтроль	червень 2024	виконано
8	Підготовка до захисту та захист кваліфікаційної роботи бакалавра	червень 2024	виконано

Виконавець: студентка групи КН-20-1  Ангеліна БОГДАНОВА
Група виконавця Підпис Ім'я, ПРІЗВИЩЕ

Керівник: к.т.н., доц. каф. КН  Олександр МАЗУРЕЦЬ
Науковий ступінь, посада Підпис Ім'я, ПРІЗВИЩЕ

Анотація

Тема кваліфікаційної роботи бакалавра: «Метод нейромережевого розпізнавання жестів у відеопотоці»

Виконавець кваліфікаційної роботи бакалавра: студентка групи КН-20-1 Ангеліна БОГДАНОВА

Керівник кваліфікаційної роботи бакалавра: к.т.н., доцент кафедри КН Олександр МАЗУРЕЦЬ

Кваліфікаційна робота бакалавра містить:

Пояснювальна записка				Кількість додатків
Сторінок	Рисунків	Таблиць	Джерел інформації	
68	34	3	32	5

Метою кваліфікаційної роботи бакалавра є покращення розпізнавання жестів у відеопотоці, для чого була виконана розробка та програмна реалізація нейромережевого методу розпізнавання жестів у відеопотоці.

Для розробки інформаційної системи було використано мову програмування Javascript, бібліотеку для машинного навчання Tensorflow.js, бібліотеку React для побудови інтерфесу користувача, бібліотеку Bootstrap для верстки веб-застосунку, бібліотеку Three.js для створення об'єктів анімації в інтерфейсі користувача та бібліотеку Jest для розробки та проведення тестування складових частин веб-застосунку. Реалізована автоматизація розпізнавання жестів у відеопотоці може бути імплементована у сферах охорони здоров'я, сфері розваг, автомобільній галузі.

Ключові слова: нейромережа, розпізнавання жестів, розпізнавання у відеопотоці, інформаційна система, жест.

Виконавець: студентка групи КН-20-1
Група виконавця

AB
Підпис

Ангеліна БОГДАНОВА
Ім'я, ПРІЗВИЩЕ

Зміст

Перелік скорочень	4
Вступ.....	5
Розділ 1 Характеристика предметної області розпізнавання жестів: аналіз моделей, методів та реалізацій.....	7
1.1 Аналіз інформаційних моделей предметної області	7
1.2 Огляд теоретичних підходів до нейромережевого розпізнавання жестів.....	9
1.3 Аналіз існуючих програмних засобів та наукових рішень щодо розпізнавання жестів у відеопотоці	13
1.4 Мета, задачі та вимоги до реалізації інформаційної системи	18
Розділ 2 Проектування системи нейромережевого розпізнавання жестів у відеопотоці	20
2.1 Схема методу нейромережевого розпізнавання жестів у відеопотоці	20
2.2 Функціональна структура системи нейромережевого розпізнавання жестів у відеопотоці	23
2.3 Аналіз та автоматизація обробки потоків даних для нейромережевого розпізнавання жестів.....	26
2.4 Архітектура нейронної мережі для розпізнавання жестів у відеопотоці.....	29
2.5 Проектна архітектура системи нейромережевого розпізнавання жестів та взаємозв'язок компонентів.....	32
2.6 Підготовка робочих вхідних даних для системи нейромережевого розпізнавання жестів.....	35
2.7 Особливості використання спеціалізованих програмних компонентів для нейромережевого моделювання	36
2.8 Визначення шляхів дослідження нейромережевого розпізнавання жестів у відеопотоці	37
2.9 Висновки до розділу 2	39
Розділ 3 Експериментальне дослідження методу нейромережевого розпізнавання жестів у відеопотоці.....	41

3.1 Структура та функціональне призначення програмних складових системи нейромережевого розпізнавання жестів	41
3.2 Особливості реалізації програмних складових системи нейромережевого розпізнавання жестів у відеопотоці	44
3.3 Тестування системи розпізнавання жестів у відеопотоці	48
3.4 Аналіз функціональності системи нейромережевого розпізнавання жестів	53
3.5 Результати досліджень з розпізнавання жестів у відеопотоці	56
3.6 Висновки до розділу 3	60
Загальні висновки.....	62
Перелік посилань.....	65
Додатки	

Перелік скорочень

Скорочення, термін, позначення	Пояснення
CCTV	Closed circuit television
LSTM	Long short-term memory
CNN	Convolutional neural network
DNN	Deep neural network
RGB	Red-Green-Blue
ЗНМ	Згорткові нейронні мережі
CPU	Central processing unit
GPU	Graphic processing unit
4D	4-dimensional spacetime
API	Application programming interface
HD	High-definition

Вступ

Кваліфікаційна робота бакалавра присвячена покращенню розпізнавання жестів у відеопотоці, для чого була виконана розробка та програмна реалізація нейромережевого методу розпізнавання жестів у відеопотоці у вигляді вебресурсу.

Актуальність. Дослідження мови жестів стали важливою сферою для розробки та імплементації програмного забезпечення в останні роки. Провідні науковці, розробники та аналітики створюють та презентують відповідні рішення на різноманітних виставках та конференціях з метою автоматизації розпізнавання жестів у відеопотоці. Провідні компанії світу, такі як Google, Microsoft, Intel та інші зацікавлені у фінансуванні досліджень у сфері розпізнавання жестів у відеопотоці.

За даними Всесвітньої організації охорони здоров'я 2005 року близько 278 мільйонів осіб у світі мали помірні або важкі порушення слуху, 80% з них мешкають у країнах з низьким і середнім рівнем доходу. Водночас в Україні станом на 2008 рік налічується близько 300 тисяч дітей та 1 мільйон дорослих з порушенням слуху, які потребують слухопротезування. За статистикою близько 1.1 мільярду молодих людей віком від 12 до 35 років загрожує втрата слуху через вплив шуму в рекреаційних умовах [1]. Зважаючи на статистику захворювань одним із рішень боротьби з проблемами слуху є створення програмних засобів, що дозволяють автоматизувати та спростити процес міжлюдської комунікації.

Проте вади слуху не є єдиними сферами для застосування технологій розпізнавання. Зокрема, дані підходи були успішно імplementовані в різних галузях, зокрема таких, як кіберполіція, автомобільна промисловість, медицина, тощо. Технології орієнтовані на розпізнавання жестів користувачів допомагають зокрема людям з різноманітними хворобами, такими як хвороба Паркінсона, порушення мовлення або ж слуху, тунельний синдром, полегшенням взаємодії з

комп'ютером, проведення ефективного навчання в школах та університетах і успішної соціалізації у суспільстві.

Окрім того, дослідження і розробки у даній сфері посилюють розвиток інформаційних технологій. Так, розробляються нові методи, архітектури, підходи та технології для вирішення поставлених завдань. Саме тому, дослідження даної проблеми є актуальним в наш час і необхідне для багатьох сфер життя і діяльності людини.

Об'єкт дослідження – процес нейромережевого розпізнавання жестів у відеопотоці.

Предмет дослідження – методи та технології машинного навчання для розпізнавання об'єктів у відеопотоці.

Мета кваліфікаційної роботи бакалавра – покращення розпізнавання жестів у відеопотоці.

Завдання кваліфікаційної роботи бакалавра – виконати дослідження предметної області розпізнавання жестів у відеопотоці; в рамках дослідження предметної області виконати огляд теоретичних підходів щодо розпізнавання жестів у відеопотоці; виконати аналіз наукових надбань й існуючих програмних рішень в області розпізнавання жестів у відеопотоці; розробити метод нейромережевого розпізнавання жестів у відеопотоці; на основі розробленого методу виконати проєктування структури системи нейромережевого розпізнавання жестів у відеопотоці; здійснити вибір засобів розробки для системи нейромережевого розпізнавання жестів у відеопотоці; здійснити програмну реалізацію системи розпізнавання жестів; провести тестування розробленої системи нейромережевого розпізнавання жестів у відеопотоці; здійснити дослідження ефективності розробленого методу з використанням розробленої системи нейромережевого розпізнавання жестів.

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

1.1 Аналіз інформаційних моделей предметної області

Зображення є відтворенням виду, форми, кольору, текстури предмета на картині, малюнку, фотографії, цифровому носії, тощо. Найлегший спосіб створення зображення – це малювання вручну, малювання на дисплеї, використання графічних засобів або верстки. Зображальні засоби поділяються на 3 види:

- основні елементи образотворчої мови (лінія, пляма, крапка, штрих, тощо);
- композиція зображення (поєднання основних елементів образотворчої мови);
- властивості поверхні, що слугує основою зображення [2].

Розпізнавання зображення – це процес аналізу графічних даних в реальному часі. Тобто кожний пристрій чи систему, що має «можливість» розпізнавати графічне зображення можна запрограмувати певним чином автоматично рефлексувати. Такі пристрої та системи можуть виконують дії абсолютно без участі людини [3].

Основна мета розпізнавання зображень має на меті класифікацію зображень базуючись на завчасно визначених категоріях та класах з подальшою інтерпретацією графічних елементів для одержання значущої інформації [4].

Людина, як зображальний об'єкт, є дуже складною фігурою для зображення, адже людина це істота, що постійно перебуває в динаміці, таким чином її обриси змінюються відповідно до її рухів: м'язів, міміки, тощо. Саме цей аспект є складним не тільки для відображення людини на малюнку, але й для розпізнавання її характеристик і особливостей. Ідентифікація людини відбувається за наступними зовнішніми ознаками:

- Анатомічні, що характеризують зріст, статуру, форму голови, риси обличчя, тощо.

– Функціональні, що характеризують рухові, мовні ознаки, а також включають в себе ходу, поставу, жестикуляцію, міміку, голос, звички людини.

– Прикмети одягу [5].

Функціональні характеристики людини є дуже важливими для розпізнавання, адже саме вони допомагають розрізнити людей, якщо анатомічні ознаки дуже схожі. Як наведено вище, однією з таких функціональних ознак є жестикуляція. Проте, жести є не лише певною характеристикою, що може бути використана для розпізнавання людини, але також є одним із способів комунікації.

Жест – це дія або рух людини з метою вираження певної емоції або інформації, що може виникати під час розмови, або замість використання вербальних засобів комунікації і становить важливу частину міжперсональної комунікації людського виду. Кожен жест несе в собі смислове навантаження і виступає певним знаком, символом або позначенням. Деякі жести використовуються людиною свідомо, несучи необхідну інформацію, деякі ж є несвідомим проявами людини. В різних культурах використання жестів відрізняється, так само як і частота жестикуляції загалом [6].

У науковій літературі та дослідженнях немає одностайної думки з приводу класифікації жестів.

В загальному жести поділяють на умовні та безумовні, знакові та незнакові.

Умовні жести є жестами, що використовуються певними соціокультурними групами і мають для них певне значення (військові вітання та віддавання честі, кивання головою, що в українців означає – «так», у болгар – «ні»). Безумні жести – це жести, що не потребують пояснень і сприймаються незалежно від культурної приналежності та інших соціальних факторів. До безумовних жестів відносять указівні жести; описові жести; експресивні жести; ритмічні жести; жести наслідування, що відображають жести іншої людини.

Знакові жести мають на меті передачу інформації від однієї людини до іншої. До знакових жестів відносять вказівні жести; образотворчі (іконографічні)

жести, які мають реальний прототип і особливості якого намагається передати той, хто жестикулює; символічні жести, що є досить абстрактними і зазвичай сприймаються в межах одного колективу, соціально-культурної групи, тощо [6].

Таксомонія, як наука, що класифікує складні явища та об'єкти, виділяє наступні жестові системи взаємодії людини та комп'ютера:

- маніпулятивна;
- сигнальна;
- комунікативна [7].

Жести є складною структурою, мають багато різних класифікацій, але є важливою функціональною характеристикою людини і універсальним способом спілкування, що не вимагає залучення голосу, мовлення та слуху, як засобів для комунікації. Враховуючи невіддільність жестів від людини, як її функціональної характеристики, широке використання жестів для спрощення спілкування, а також забезпечення комунікації у разі відсутності інших способів, автоматизація розпізнавання жестів є вкрай актуальною галуззю. Відповідно, основними сутностями предметної області є жести, розпізнавання жестів, а властивостями є вигляд жестів, частота жестів, смислове навантаження жестів.

1.2 Огляд теоретичних підходів до нейромережевого розпізнавання жестів

З розвитком інформаційних технологій на основі ідеї розпізнавання зображень виникла нова галузь сучасного штучного інтелекту – комп'ютерне бачення або комп'ютерний зір. За допомогою комп'ютерного бачення стає можливим розпізнавання та ідентифікація зображень, шаблонів та об'єктів на цифрових носіях [4].

І розпізнавання зображень, і комп'ютерне бачення мають один і той самий принцип ідентифікації об'єкта, але мають певні відмінності, що полягають у обсязі, цілях, рівню аналізу даних та використаними методами. На відміну від розпізнавання зображень, комп'ютерне бачення дозволяє не лише виконувати

безпосередній процес розпізнавання, але й інтерпретувати отримані дані, реагувати на класифікований об'єкт та створювати прогнози [4]. Застосування технологій та засобів комп'ютерного зору відбувається у таких сферах: автомобільна, охорона здоров'я, роздрібна торгівля, сільське господарство, тощо.

В автомобільній галузі система замкнутого телебачення CCTV використовує комп'ютерний зір для відстеження і класифікації транспорту для зменшення заторів, визначення вільних ділянок для проїзду, контролю паркування транспортного засобу, тощо.

Комп'ютерне бачення у галузі охорони здоров'я використовується для поліпшення виявлення ракових клітин, виявлення потенційних або вже наявних захворювань опорно-рухового апарату, виявлення наявності неврологічних захворювань, а також діагностики COVID.

В сфері роздрібної торгівлі з метою підрахунку трафіку через магазини комп'ютерний зір широко використовується і забезпечує відстеження тенденцій та преференцій покупців, а також дозволяє запобігати втратам та крадіжкам у місцях роздрібної торгівлі.

Використання програмного забезпечення в сфері сільського господарства дозволяє покращити способи ідентифікації проблем з обладнанням, виявлення погіршення якості продуктів, покращити продуктивність та ефективність працівників у сфері сільського господарства за рахунок розпізнавання обличчя, тощо [8].

Розпізнавання образів є нетривіальним завданням, адже існує багато змінних параметрів та факторів. Проте за допомогою штучних нейронних мереж подібні задачі можливо вирішувати.

Нейромережа – це математична модель у вигляді програмного та апаратного втілення, що будується на принципах функціонування біологічних нейромереж.

Виділяються кілька різних архітектур штучних нейромереж, у тому числі нейромереж для розпізнавання зображень:

– Багатошаровий перцептрон, який складається з 3 або більше шарів, а також використовує нелінійну функцію активації для класифікації даних [9].

– Рекурсивна нейронна мережа, яка є класом глибоких нейронних мереж, що створені рекурсивним застосуванням одного і того ж набору ваг до структури, щоб здійснювати структурне передбачення вхідних даних динамічних розмірів. Цей вид архітектури нейронної мережі гарно зарекомендував себе в обробці природної мови, а саме – передбачення цілих фраз і речень на основі векторного представлення слів [10].

– Рекурентна нейронна мережа, яка є класом штучних нейронних мереж, де з'єднанні між собою вузли утворюють орієнтований у часі граф. Саме ця властивість дозволяє мережі проявляти динамічну поведінку в часі, використовувати свою внутрішню пам'ять для обробки довільних послідовностей вхідних даних, що робить її хорошим варіантом для вирішення таких задач, як розпізнавання неперервного рукописного тексту чи розпізнавання мовлення [11].

– Sequence-to-sequence є видом нейронної мережі, що складається з двох рекурентних нейронних систем, що виконують функції кодувальника і розкодовувача [12].

– LSTM (long short-term memory) це архітектура рекурентних нейронних мереж. Мережа LSTM добре підходить для навчання з досвіду з метою класифікації, обробки або передбачення часових рядів в умовах, коли між важливими подіями існують часові затримки невідомої тривалості. Є досить популярною, її використовують такі компанії-гіганти, як Google, Apple, Microsoft [13].

– Згорткова нейромережа (CNN, convolutional neural network) є класом глибоких штучних нейронних мереж прямого поширення, що означає поширення сигналу лише в одному напрямку – від вхідного шару далі до прихованих і до вихідного шару. Тобто це різновид багатошарового перцептрона з використанням операції згортки. Успішно застосовується для аналізу візуальних зображень [14].

Задача розпізнавання зображень ускладнюється необхідністю розпізнавання в реальному часі. Зображення в реальному часі може мати різні розміри, орієнтацію в просторі, освітлення, контраст об'єкту з фоном, тощо, тож нейронній мережі необхідно реагувати на зміну багатьох параметрів багатьох зображень і оперативно опрацьовувати їх. Тому нейромережа повинна задовольняти вимогу швидкості і точності на великих обсягах даних. Додатково потрібно враховувати затримки мережі, збої та інші фактори для коректної ідентифікації зображень або їх частин.

Для розпізнавання об'єктів у відеопотоці часто використовуються наступні методи і алгоритми: DNN, CNN, R-CNN, Fast-R-CNN, Faster R-CNN, алгоритм YOLO.

DNN є видом глибоких нейронних мереж, що добре зарекомендували себе в розпізнавання зображень, але вони потребують значних обчислювальних ресурсів і великого набору даних для тренування

CNN є видом згорткових нейронних мереж, які є хорошим вибором для розпізнавання зображень у відеопотоці, що дозволяє швидко розпізнавати об'єкти та їх характеристики. Дані нейронні мережі є дуже швидкими порівняно з іншими видами, особливо за використання додаткової оптимізації.

R-CNN, Fast-R-CNN, Faster R-CNN оснований на згорткових нейронних мережах та використовуються для розпізнавання зображень, але потребує багато ресурсів для навчання.

Алгоритм YOLO – це метод створений на основі CNN, та є дуже швидким у розпізнавання зображень у відеопотоці, так як використовує метод зворотного поширення, що дозволяє здійснювати розпізнавання навіть на пристроях з обмеженими ресурсами [15].

Враховуючи доступні методи для створення програмного засобу з метою розпізнавання жестів у відеопотоці, було обрано CNN як класичну нейромережу, яка відмінно зарекомендувала себе як для розпізнавання зображень, так і для забезпечення швидкодії в умовах роботи в реальному часі.

1.3 Аналіз існуючих програмних засобів та наукових рішень щодо розпізнавання жестів у відеопотоці

Одним із прикладів застосування штучних нейронних мереж для розпізнавання зображень у відеопотоці є робота «Розпізнавання жестів за допомогою згорткової нейронної мережі» авторів Проніної О.І., Остапенка П.Б., Альошина С.Б [16]. В роботі автори використали алгоритм YOLOv5 з представленням об'єктів розпізнавання у вигляді векторів для забезпечення однорідності між зображеннями одного класу та найбільшої відмінності для зображень різних класів. Авторам вдалося досягти високої точності моделі при різних видах освітлення, положення рук при демонстрації жестів. Навчання моделі зайняло більше 50 годин, а точність складає 70% (Рисунок 1.1).

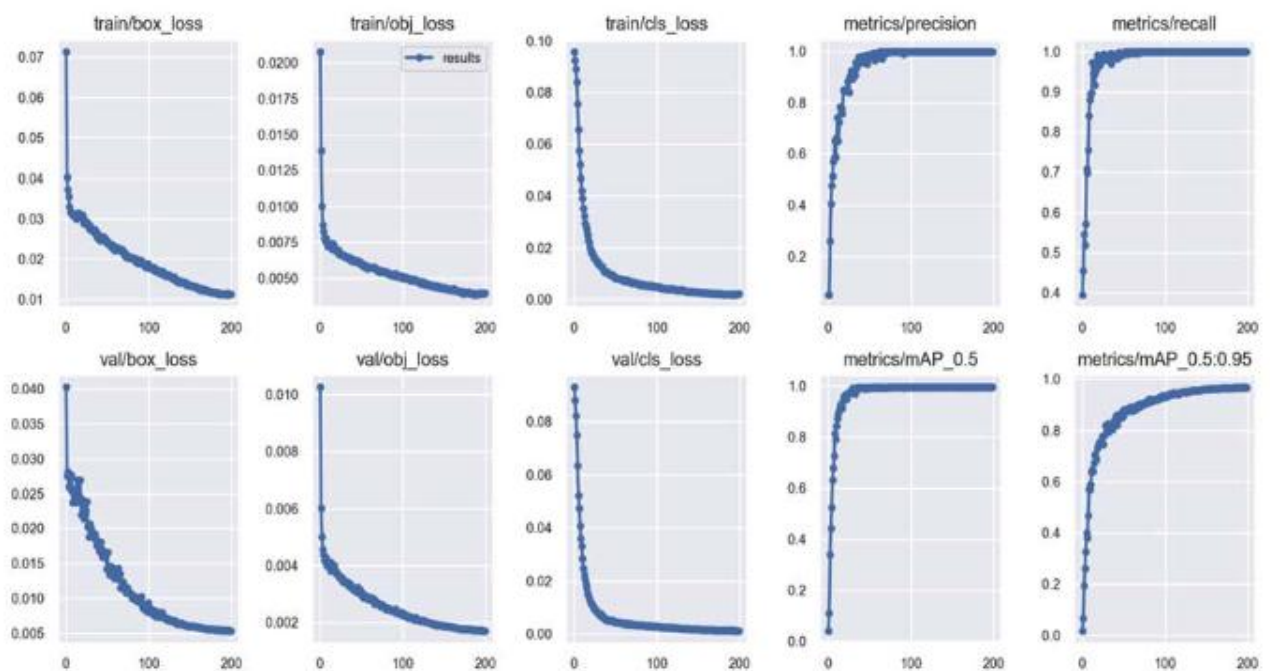


Рисунок 1.1 – Результат тренування моделі YOLOv5 [16]

У 2017 році корейська компанія Masgon презентувала свій застосунок для розпізнавання жестів FingerTalk [17]. Щоб запустити технологію розпізнавання жестів програми, користувач піднімає свою долоню в повітря, що призводить до автоматичної появи певного об'єкта на долоні користувача. Потім цей об'єкт

можна кидати або переміщати іншими способами. Цей застосунок захоплює зображення в реальному часі та відображає його на екрані смартфона. До переваг програмного застосунку можна віднести відсутність обов'язкової вимоги до наявності 3D-камери або додаткових сенсорів, а також те, що додаток створює можливість «взаємодіяти» з іншими об'єктами. Жести розпізнаються без використання допоміжних засобів, таких як спеціальні рукавички, для того, щоб контролювати девайс, що робить застосунок доступним для різних верств населення. Інтерфейс є простим і інтуїтивно зрозумілим, що значно полегшує користуванням додатком (Рисунок 1.2).



Рисунок 1.2 – Програмний застосунок FingerTalks [17]

Ще одним прикладом застосунку для розпізнавання жестів можна є створений французькою компанією по виробництву електроніки Terabee однойменний застосунок [18]. Це не лише програмний продукт, але також ціла система камер, сенсорів і датчиків для розпізнавання жестів і їхнього практичного застосування (Рисунок 1.3).

RELATED PRODUCTS

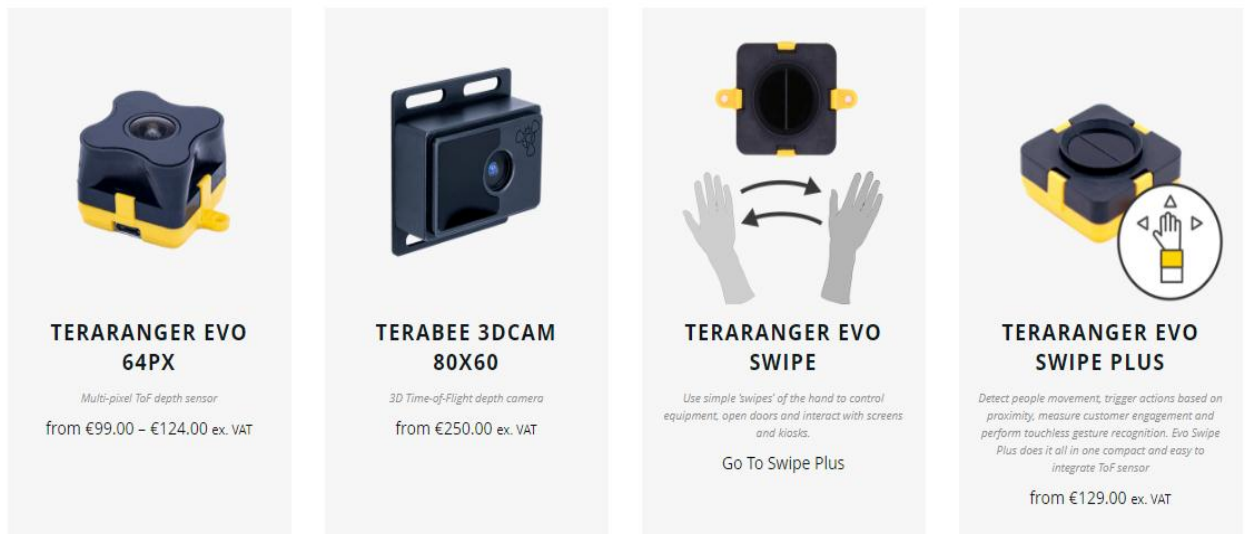


Рисунок 1.3 – Лінійка девайсів для використання розпізнавання жестів від компанії Terabee [18]

Компанія презентувала безконтактний девайс, який зчитує жест людини, що знаходиться поряд, і виконує певну команду, що «прив’язана» до цього жесту. Для людей з певними обмеженнями, яким важко пересуватись постійно чи тимчасово, літніх людей або ж тих, хто має проблеми з маніпуляцією дрібними предметами, такими як смартфон або пульт, застосунок дозволяє максимально ефективно керувати різноманітними пристроями, прикладаючи найменші зусилля.

Серед переваг можна зазначити практичне застосування створеного продукту. На презентації було продемонстровано застосування технології для переключення картинок на телевізорі, керування пунктами меню, тощо (Рисунок 1.4).

До недоліків цього застосунку можна віднести обов’язкову взаємодію із камерами і сенсорами для того, щоб керувати об’єктами інтерфейсу, за відсутності відповідних девайсів керування за допомогою жестів буде неможливим, а ціни на відповідні девайси не є доступними кожному користувачу комп’ютера.



Рисунок 1.4 – Демонстрація використання сенсорів і камер для практичного застосування технології розпізнавання жестів [18]

Іншим прикладом застосування технологій для розпізнавання жестів є застосунок випущений у 2019 році компанією Conena під назвою Gesture Control, доступний для завантаження в Google App [19].

Застосунок пропонує своїм користувачам полегшити навігацію у смартфоні, відмовившись від дотику до екрану смартфона і пошуку потрібної навігаційної панелі, а користуючись лише жестами, що також є перевагою для людей з різноманітними хворобами, такі як артрит і артроз, що характеризуються чутливістю і болями при взаємодії з мілкими девайсами а також швидку втомлюваність. Це ж стосується людей, що страждають на тунельний синдром.

Додаток має безліч налаштувань, простий і інтуїтивно зрозумілий інтерфейс. Застосунок дає можливість використовувати 20 різних жестів, кожен з яких відповідає певній команді, а також гнучкі можливості для налаштування додатку під себе (Рисунок 1.5).

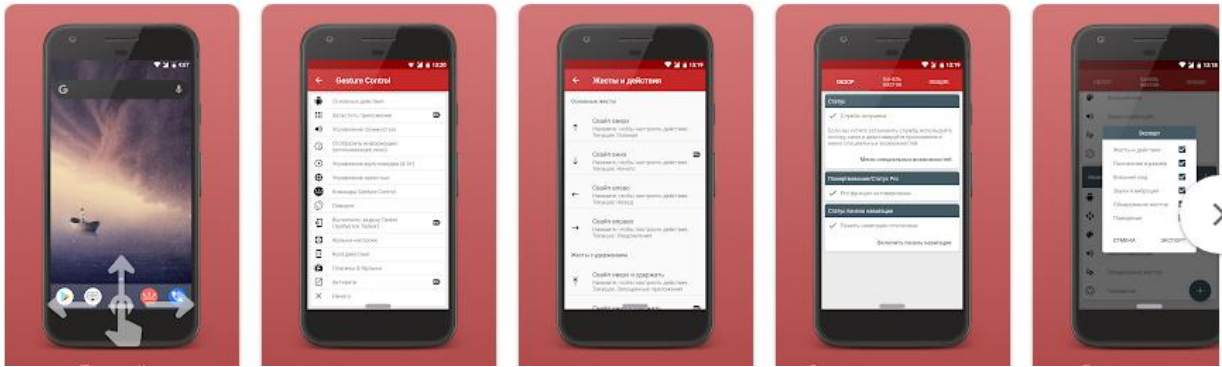


Рисунок 1.5 – Додаток для керування за допомогою жестів Gesture Control [19]

Проте, користувачі скаржаться, що є певні затримки при відкритті додатку на смартфонах на android 10. Також користувачі зазначили наявність реклами та певні проблеми з бічною анімацією в додатку, яка є не досить плавною, як недоліки додатку Gesture Control.

Розпізнавання жестів широко використовується для створення протезів людям, що втратили кінцівки. Так, американські вчені застосували штучний інтелект для того, щоб забезпечити високу точність керування біонічною рукою. Коли пацієнт із біонічною рукою, пов'язаною з периферійними нервами, бажає поворухнути нею, сигнал про це з мозку доходить до тих самих периферійних нервів, де зчитується декодером. Завданням декодера є в реальному часі правильно інтерпретувати сигнал та передати до протеза у вигляді команди виконати саме бажану людиною дію. Цей «переклад» сигналу від нервів до протеза вимагає опрацювання значної кількості інформації, тому далеко не завжди він відбувається безпомилково, і для вирішення цього питання розробники скористались штучним інтелектом [20].

Експерименти показали, що технологія має певні недоліки і працює нестабільно. Один із учасників експерименту завдяки цій розробці навіть зміг подумки пограти у відеогру. Та все ж біонічна рука реагує на рухи пальцями кисті руки, і найкраще – на рухи безіменним пальцем і мізинцем. Технологія ще буде проходити етапи вдосконалення і має великі перспективи на майбутнє (Рисунок 1.6).



Рисунок 1.6 – Демонстрація роботи біонічної руки за використанням штучного інтелекту для розпізнавання жестів [20]

Звісно, на ринку представлено багато варіантів розпізнавання жестів для різних потреб. Здебільшого з вільним доступом, безкоштовно можна знайти додатки, що мають на меті розважити користувача або полегшити побутові питання, такі як скрол новин під час сніданку або ж просто спрощення навігації по смартфону. Але враховуючи актуальність предметної області, необхідність у нових сучасних рішеннях для вирішення проблеми розпізнавання різних видів зображень, метод неймережевого розпізнавання жестів у відеопотоці є досить актуальним напрямом для дослідження та розробки. Багато користувачів бажають полегшити взаємодію з комп'ютерами і телефонами, тому попит на подібні додатки лише зростатиме.

1.4 Мета, задачі та вимоги до реалізації інформаційної системи

Мета кваліфікаційної роботи бакалавра – покращення розпізнавання жестів у відеопотоці, для чого потрібно розв'язати наступні задачі:

- виконати дослідження предметної області розпізнавання жестів у відеопотоці;
- в рамках дослідження предметної області виконати огляд теоретичних підходів щодо розпізнавання жестів у відеопотоці;
- виконати аналіз наукових надбань й існуючих програмних рішень в області розпізнавання жестів у відеопотоці;
- розробити метод неймережевого розпізнавання жестів у відеопотоці;
- на основі розробленого методу виконати проектування структури системи неймережевого розпізнавання жестів у відеопотоці;
- здійснити вибір засобів розробки для системи неймережевого розпізнавання жестів у відеопотоці;
- здійснити програмну реалізацію системи розпізнавання жестів;
- провести тестування розробленої системи неймережевого розпізнавання жестів у відеопотоці;
- здійснити дослідження ефективності розробленого методу з використанням розробленої системи неймережевого розпізнавання жестів.

Для дослідження ефективності розробленого методу слід виконати його програмну реалізацію, що містить наступні основні функції:

- запитування дозволу браузера ввімкнути камеру на девайсі користувача;
- створення додаткового навчального датасету користувачем;
- вибору кількості вхідних даних для подальшого тренування моделі користувачем;
- запуску навчання моделі у реальному часі;
- перегляду статусу навчання моделі у реальному часі;
- початку процесу розпізнавання зображень;
- зупинки розпізнавання зображень
- демонстрація результату розпізнавання жестів у відеопотоці.

Розділ 2 Проєктування системи нейромережевого розпізнавання жестів у відеопотоці

2.1 Схема методу нейромережевого розпізнавання жестів у відеопотоці

Схема методу нейромережевого розпізнавання жестів у відеопотоці передбачає 4 етапи, вхідні дані та вихідні дані (Рисунок 2.1).



Рисунок 2.1 – Схема методу нейромережевого розпізнавання жестів у відеопотоці

В якості вхідних даних є перелік актуальних класів для розпізнавання образів (кількість, назва), параметри навчання нейромережевої моделі (розмірність вхідних даних, оптимізатор, функція витрат, кількість епох, функція зворотного виклику), параметри донавчання моделі (кадр./мсек., кадр./серія), відеопотік для донавчання нейромережевої моделі та розпізнавання жестів.

Перелік актуальних класів для розпізнавання образів являє собою масив даних різних класів зображень для подальшої категоризації зображень під час розпізнавання жестів у відеопотоці.

Розмірність вхідних даних, як параметр навчання нейромережевої моделі, задає розміри блоку захоплення відеокамери для подальшого коректного опрацювання кадрів.

В даному методі в якості функції активації використовуються функції Relu та Softmax, що дозволяють нормалізувати вихідні дані. Функція Relu представлена наступною формулою [21]:

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

де z – аргумент, переданий у функцію – вхідний вектор.

Функція активації Softmax, яка перетворює вектор z розмірності K в вектор σ тієї ж розмірності, де кожна координата σ_i отриманого вектора представлена дійсним числом в інтервалі $[0, 1]$ і сума координат дорівнює 1. Координати σ_i вираховуються наступним чином [22]:

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^K e^{z_k}},$$

де z – вхідний вектор, K – розмірність вектора, σ – інший вектор тієї ж розмірності.

Координати σ_i отриманого вектора трактуються як ймовірність того, що об'єкт належить класу i . Вектор-стовбець z при цьому розраховується наступним чином [22]:

$$z = w^T x - \theta,$$

де x – це вектор-стовбець ознак об'єкта розмірності $M \times 1$, w^T – транспонована матриця вагових коефіцієнтів ознак об'єкта, що має розмірність $K \times M$, θ –

вектор-стовбець з пороговим значенням розмірності $K \times 1$, K – кількість класів об'єкта, M – кількість ознак об'єкта.

Оптимізатором в моделі виступає функція Adam з визначеним коефіцієнтом швидкості навчання для оптимізації функції витрат під час навчання нейромережевої моделі, що представлено наступними формулами [23]:

$$w_{t+1} = w_t - \alpha w_t,$$

$$m_t = \beta w_{t-1} + (1 - \beta) \left[\frac{\delta L}{\delta w_t} \right],$$

де m_t – сукупність градієнтів у момент часу t [поточний час] (спочатку $m_t = 0$), m_{t-1} сукупність градієнтів у момент часу $t-1$ [попередній час], w_t – ваги в момент часу t , w_{t+1} – ваги в момент часу $t+1$, α_t – швидкість навчання в момент t , δL – похідна функції похибки, δw_t – похідна ваг у момент t , β – параметр середнього рухомого (const, 0,9).

В якості функції витрат використано функцію категоріальної крос-ентропії categoricalCrossentropy, що використовується для порівняння прогнозованих і фактично розпізнаних класів.

Кількість епох, як параметр навчання нейромережевої моделі, визначає кількість ітерацій для навчання моделі.

Функція зворотного виклику використовується для відстеження втрати під час навчання і викликається на кожній партії даних.

На першому етапі моделі відбуваються такі кроки, як ініціація нейромережевої моделі, підбір актуальних класів зображень для донавчання, налаштування параметрів моделі.

На другому етапі для створення повного датасету зображень відбувається збір зображень за допомогою відеокамери користувача та обробка даного відеопотоку для формування множини кадрів для кожного вибраного раніше класу.

На третьому етапі відбувається навчання нейромережевої моделі, яке супроводжується остаточним формуванням кінцевого датасету (базова + персоналізована частини), коригування нейромережевої моделі згідно параметрів задачі розпізнавання (додавання необхідних шарів, функцій

активації, функції втрати) та безпосереднє навчання нейромережевої моделі навчальним датасетом.

На кінцевому четвертому етапі відбуваються такі процеси, як обробка відеопотоку та виокремлення кадру для розпізнавання, розпізнавання жесту на відеокадрі з відеопотоку та повернення до моніторингу відеопотоку. Вихідними даними є розпізнанні жести користувача з відеопотоку.

2.2 Функціональна структура системи нейромережевого розпізнавання жестів у відеопотоці

Для розпізнавання жестів у відеопотоці використовується модель MobileNet архітектури CNN. Функціональна структура моделі включає в себе основні етапи, що має виконати модель для досягнення поставленої мети – нейромережевого розпізнавання жестів у відеопотоці. Для подальшої роботи було побудовано блок-схему, що показує основні етапи роботи моделі, а також описує необхідні для роботи моделі вхідні дані та результат нейромережевого розпізнавання жестів у відеопотоці – жести як вихідні дані нейромережевої моделі (Рисунок 2.2).

Спочатку в модель входять дані розміром 244x244 з трьома кольоровими каналами RGB.

На етапі попередньої обробки відбувається нормалізація зображень до розміру, який модель здатна прийнята – в даному випадку зображення розміром 244x244 пікселів та нормалізація пікселів зображення для подальшої коректної роботи моделі.

Далі нормалізовані зображення проходять крізь наступні шари моделі mobileNet:

- Конволюційний шар – це шар, де відбувається витягування ознак класів зображень.

- Глибока згортка відповідає за обробку кожного шару (каналу) зображення.

- Точкова згортка відповідає за об'єднання інформації з усіх каналів, отриманих на попередньому етапі.
- Функція активації Relu дозволяє моделі навчатись розпізнавати комплексні патерни в даних.
- Повнозв'язний шар відповідає за об'єднання всіх даних для класифікації зображення.
- Функція активації Softmax відповідає за визначення ймовірності для кожного класу для визначення найбільш точного жесту.

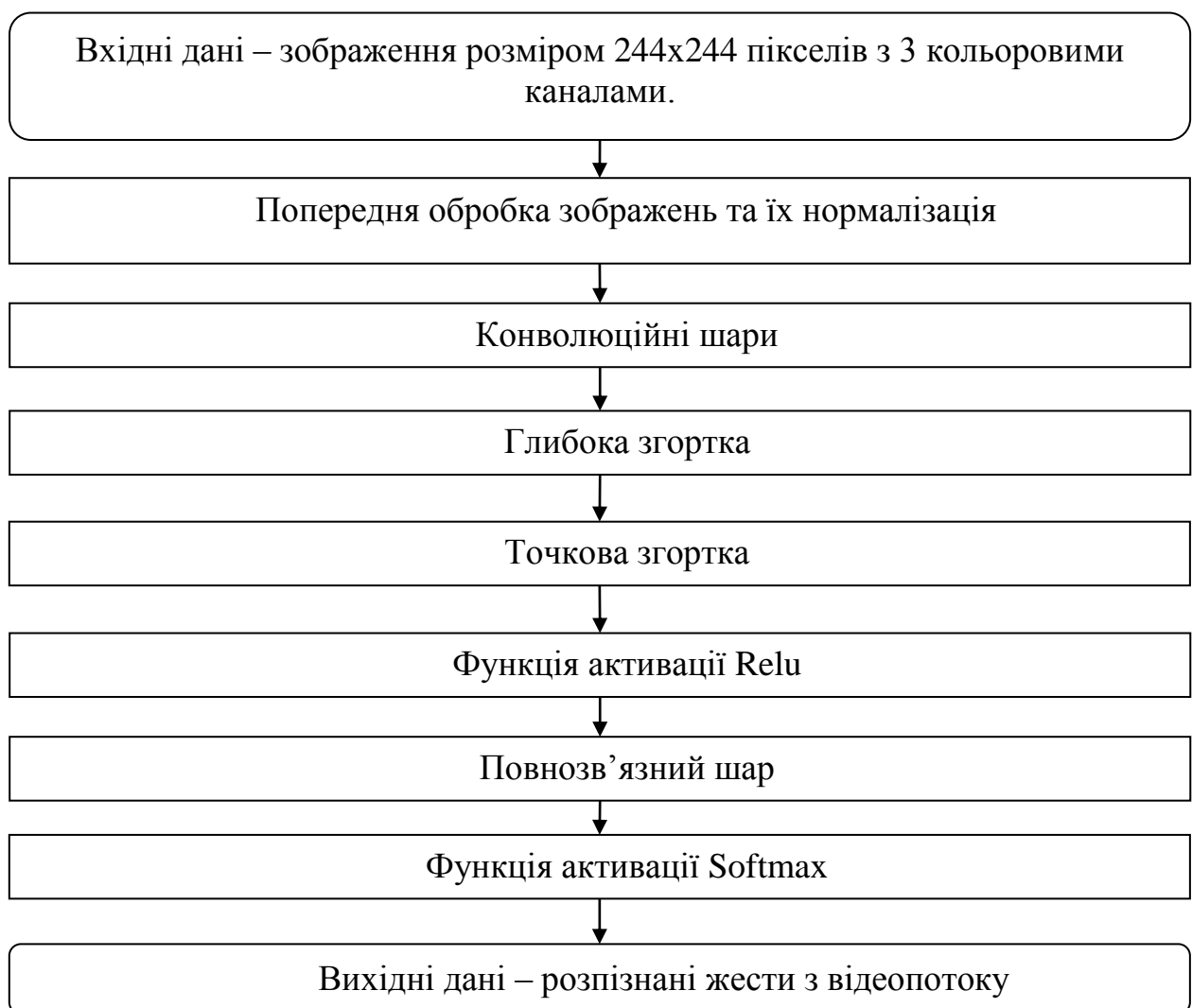


Рисунок 2.2 – Блок-схема основних етапів моделі MobileNet

Для створення системи, що виконує поставлену задачу і є придатною для використання користувачем, було побудовано діаграму послідовностей, яка

відображає послідовність використання системи та способи, якими користувач зможе взаємодіяти із застосунком (Рисунок 2.3).

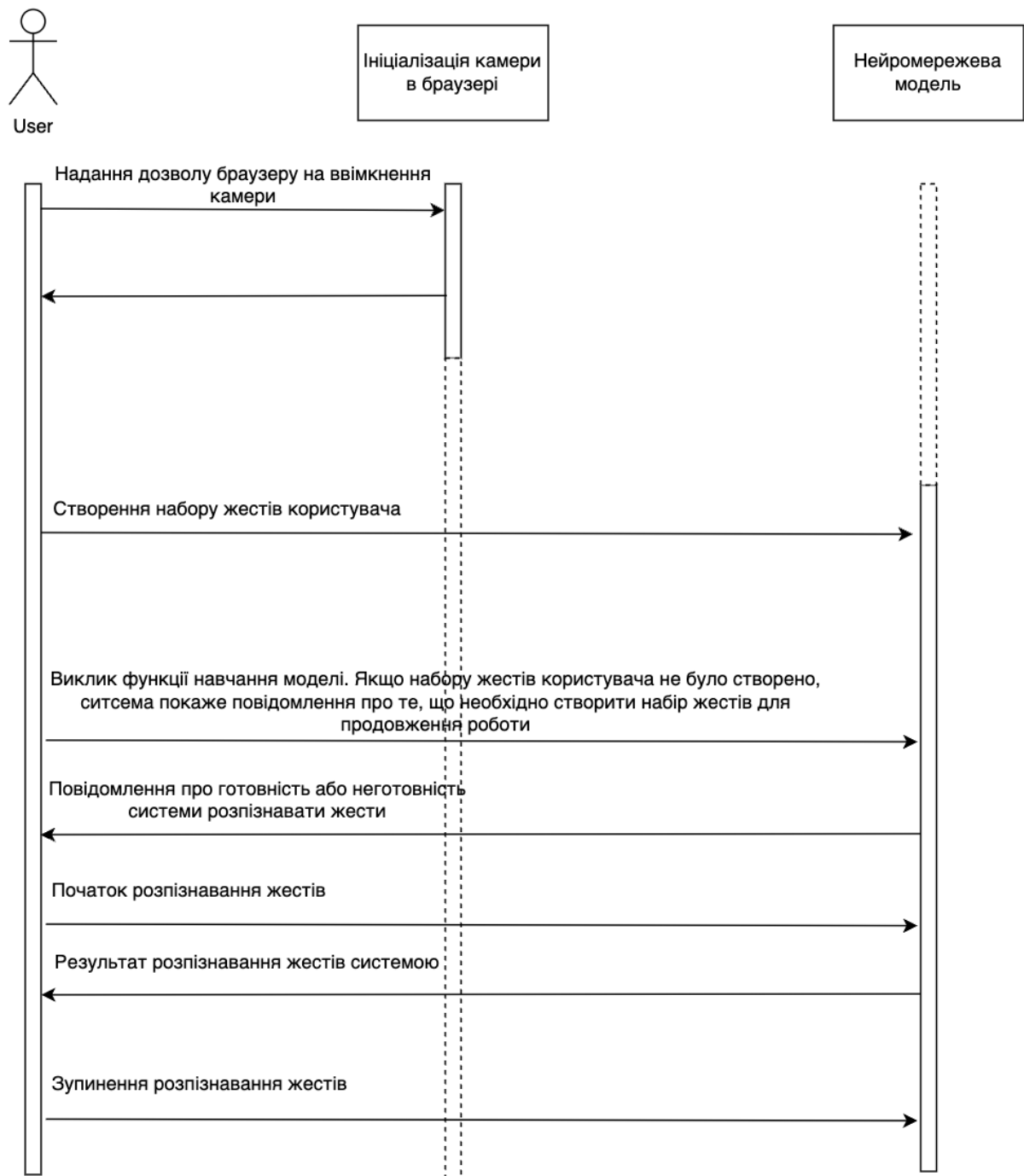


Рисунок 2.3 – Діаграма послідовностей системи

Відповідно до діаграми можна прослідкувати послідовність роботи системи та взаємодії всіх компонентів.

Першим кроком є отримання дозволу браузера ввімкнути відео користувача для подальшої активації можливості розпізнавання жестів у відеопотоці.

Наступним кроком є створення вибірки жестів користувача. Цей крок є необхідним для початку роботи моделі, адже система має пройти донавчання на актуальних зображеннях жестів поточного користувача. Після демонстрації модель збереже зображення для їх обробки, нормалізації та донавчання моделі.

При виклику функції, що відповідає за початок тренування моделі користувачем, відбувається процес донавчання. При відсутності вибірки зображень жестів користувача система не зможе перейти до цього кроку і необхідно буде повернутись до попереднього кроку – створення вибірки жестів користувача. За успішного донавчання моделі користувач побачить повідомлення в інтерфейсі системи про те, що можна починати процес розпізнавання жестів у відеопотоці.

Для початку і завершення розпізнавання жестів необхідно запустити відповідну функцію, що викликається з інтерфейсу користувача і дозволяє керувати етапами розпізнавання жестів.

2.3 Аналіз та автоматизація обробки потоків даних для нейромережевого розпізнавання жестів

На діаграмі послідовностей системи було продемонстровано, що користувач має змогу взаємодіяти з моделлю та керувати певними етапами її роботи. Для розуміння способів використання моделі користувачем було побудовано діаграму варіантів використання системи (Рисунок 2.4).

Сценарій прецеденту «Надання дозволу браузеру на ввімкнення камери»:

1. Назва прецеденту: надання дозволу браузеру на ввімкнення камери.
2. Діюча особа: користувач.
3. Мета: дозволити ввімкнути камеру для подальшого захоплення жестів користувача.

4. Передумови: відкрити застосунок у браузері.

5. Головна послідовність: відкрити застосунок у браузері > натиснути на кнопку «Дозволити» як відповідь на запит браузера > ввімкнення камери.

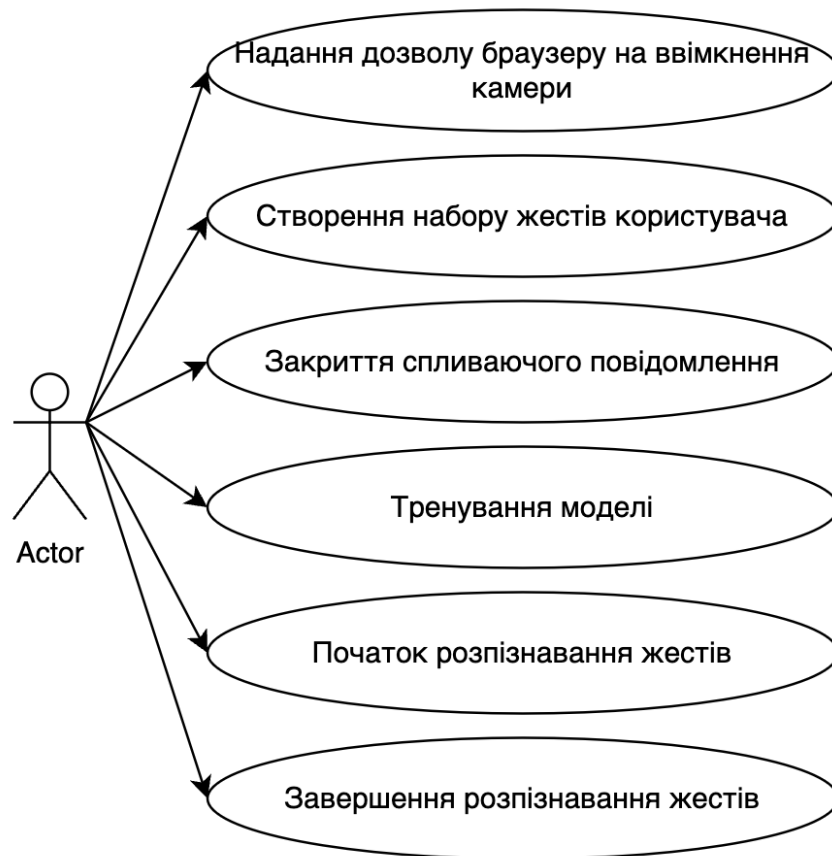


Рисунок 2.4 – Діаграма варіантів використання системи

Сценарій прецеденту «Створення набору жестів користувача»:

1. Назва прецеденту: створення набору жестів користувача.
 2. Діюча особа: користувач.
 3. Мета: створити набір жестів користувача, для подальшого навчання моделі.

4. Передумови: відкрити застосунок у браузері, дозволити ввімкнути камеру.

5. Головна послідовність: виконати прецедент «надання дозволу браузеру на ввімкнення камери» > продемонструвати жест (в даному випадку від 0 до 5) і натиснути на відповідну кнопку обраного класу жесту.

Сценарій прецеденту «Закриття спливаючого повідомлення»:

1. Назва прецеденту: закриття спливаючого повідомлення.
2. Діюча особа: користувач.
3. Мета: закрити інформаційне спливаюче повідомлення.
4. Передумови: відкрити застосунок у браузері, дозволити ввімкнути камеру.
5. Головна послідовність: виконати прецедент «надання дозволу браузеру на ввімкнення камери» > натиснути на кнопку «Training» або «Start training» > після появи інформаційного спливаючого повідомлення натиснути на кнопку хрестик для його закриття.

Сценарій прецеденту «Тренування моделі»:

1. Назва прецеденту: тренування моделі.
2. Діюча особа: користувач.
3. Мета: почати тренування моделі для подальшого розпізнавання жестів користувача.
4. Передумови: відкрити застосунок у браузері, дозволити ввімкнути камеру.
5. Головна послідовність: виконати прецедент «надання дозволу браузеру на ввімкнення камери» > виконати прецедент «створення набору жестів користувача» > натиснути на кнопку «Training»

Сценарій прецеденту «Розпізнавання жестів»:

1. Назва прецеденту: розпізнавання жестів.
2. Діюча особа: користувач.
3. Мета: почати розпізнавання жестів користувача.
4. Передумови: відкрити застосунок у браузері, дозволити ввімкнути камеру.
5. Головна послідовність: виконати прецедент «надання дозволу браузеру на ввімкнення камери» > виконати прецедент «створення набору жестів користувача» > Натиснути на кнопку «Start predicting».

Сценарій прецеденту «Завершення розпізнавання жестів»:

1. Назва прецеденту: завершення розпізнавання жестів.

2. Діюча особа: користувач.
3. Мета: завершення розпізнавання жестів користувача.
4. Передумови: відкрити застосунок у браузері, дозволити ввімкнути камеру.
5. Головна послідовність: виконати прецедент «надання дозволу браузеру на ввімкнення камери» > виконати прецедент «створення набору жестів користувача» > натиснути на кнопку «Stop predicting»

Для взаємодії користувача з системою було надано певний діапазон можливостей, як саме користувач може керувати етапами роботи моделі за допомогою інтерфейсу для розпізнавання жестів у відеопотоці.

2.4 Архітектура нейронної мережі для розпізнавання жестів у відеопотоці

Метод нейромережевого розпізнавання жестів у відеопотоці базується на використанні моделі MobileNet, що в свою чергу базується на використанні архітектури CNN – згорткової нейронної мережі.

CNN складається з вхідних та вихідних шарів, а також прихованих шарів. Така архітектура складається зазвичай з згорткових шарів, агрегувальних шарів, повноз'єднаних шарів та шарів нормалізації. Згорткові шари слугують входом до операції згортки, яка імітує реакцію нейрона на зоровий стимул. Агрегувальні шари є шарами, які дозволяють об'єднати виходи кластерів нейронів одного шару, щоб повноцінно передати їх до наступного шару. Повноз'єднані шари використовуються для поєднання кожного нейрону одного шару з кожним нейроном наступного шару. А ваги в згортковій нейронній мережі є спільними в згорткових шарах, що дозволяє зменшити обсяг необхідної пам'яті та оптимізувати швидкодію нейронної мережі [24].

MobileNet – це клас ЗНМ з відкритим кодом Google, яка використовує роздільні згортки по глибині, що значно зменшує кількість параметрів у порівнянні з мережею зі звичайними згортками з однаковою глибиною [25].

Загальна архітектура моделі MobileNet дозволяє змінювати різноманітні параметри і таким чином підлаштовувати модель для різного виду проблем і задач (Рисунок 2.5).

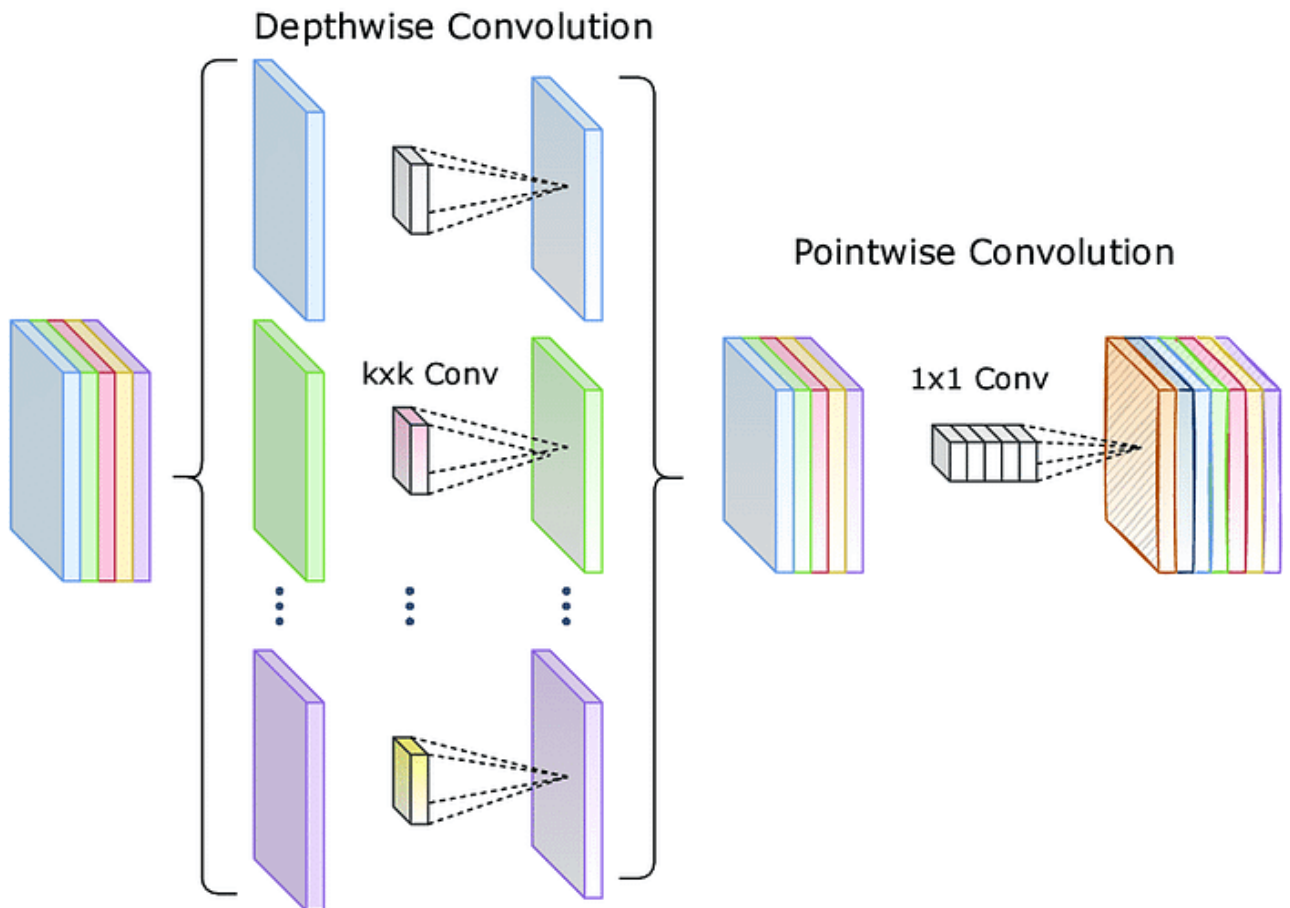


Рисунок 2.5 – Загальна діаграма роботи моделі MobileNet [26]

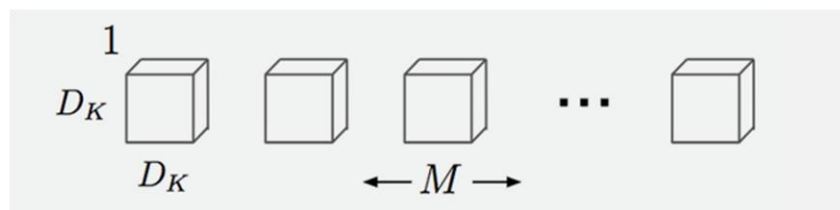


Рисунок 2.6 – Глибинна згортка для одного вхідного каналу [25]

Згортка – це частина моделі з роздільною глибиною, що складається з двох операцій:

- глибинна згортка;
- точкова згортка.

Глибинна згортка виникла з ідеї, що глибина і просторовий вимір фільтра можуть бути розділені – звідси і назва роздільна (Рисунок 2.6).

Точкова згортка – це згортка з розміром каналів 1×1 , тобто вона поєднує шари, створені за допомогою глибинної згортки (Рисунок 2.7).

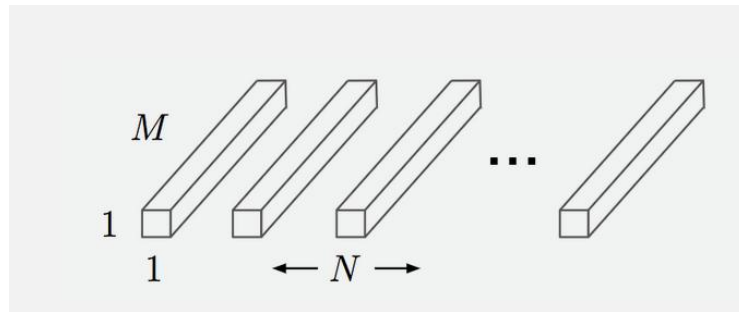


Рисунок 2.7 – Точкова згортка [25]

До прихованих шарів моделі MobileNet архітектури CNN відносяться шари Flatten і Dense.

Шар Flatten використовується для конвертації вхідних даних у меншу розмірність, тобто в одномірний вектор [27] (Рисунок 2.8).

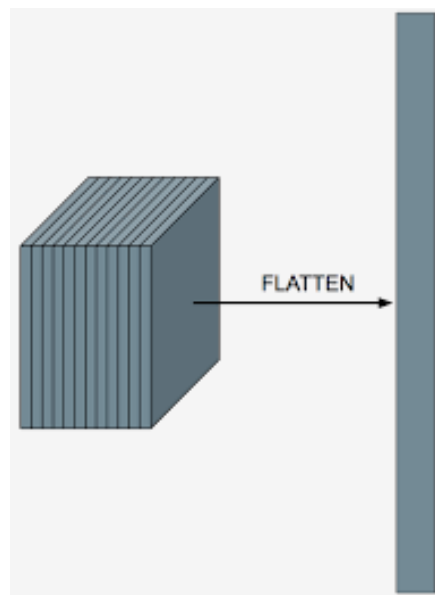


Рисунок 2.8 – Демонстрація роботи Flatten шару [27]

Dense – це наступний шар після шару Flatten, в якості функції активації шару Dense виступають Relu і Softmax функції.

В якості функції-оптимізатора виступає функція Adam, а для визначення похибки між очікуваними вихідними даними і реальними використовується функція categoricalCrossentropy.

Така архітектура дозволяє досягти високої точності для розпізнавання образів, при тому зберігаючи швидкодію та забезпечуючи економію виробничих ресурсів.

2.5 Проектна архітектура системи нейромережевого розпізнавання жестів та взаємозв'язок компонентів

Відповідно до методу нейромережевого розпізнавання жестів у відеопотоці, було спроектовано відповідну структуру інформаційної системи (Рисунок 2.9).

Інформаційна система нейромережевого розпізнавання зображень жестів у відеопотоці включає у себе наступні 4 підсистеми:

- підсистема «Налаштування нейромережевої моделі MobileNet»;
- підсистема «Формування персоналізованої частини датасету»;
- підсистема «Навчання нейромережевої моделі для розпізнавання жестів»;
- підсистема «Нейромережеве розпізнавання жестів у відеопотоці».

Підсистема налаштування нейромережевої моделі MobileNet включає в себе наступні функції:

- Ініціація нейромережевої моделі. На цьому етапі відбувається першочерговий запуск нейромережевої моделі для подальшого процесу розпізнавання образів у відеопотоці.

- Підбір актуальних класів зображень для донавчання. Для дослідження роботи моделі з метою нейромережевого розпізнавання жестів було обрано 6 класів зображень.

- Налаштування параметрів моделі. Розмірність вхідних даних становить 244x244x3 пікселів, активаційними функціями виступають функція Relu та

функція Softmax, в якості оптимізатора використовується оптимізатор Adam з коефіцієнтом швидкості навчання 0.0001, функція витрат є функцією категоріальної крос-ентропії categoricalCrossentropy, кількість епох становить 10, функція зворотного виклику викликається на кожній партії даних.

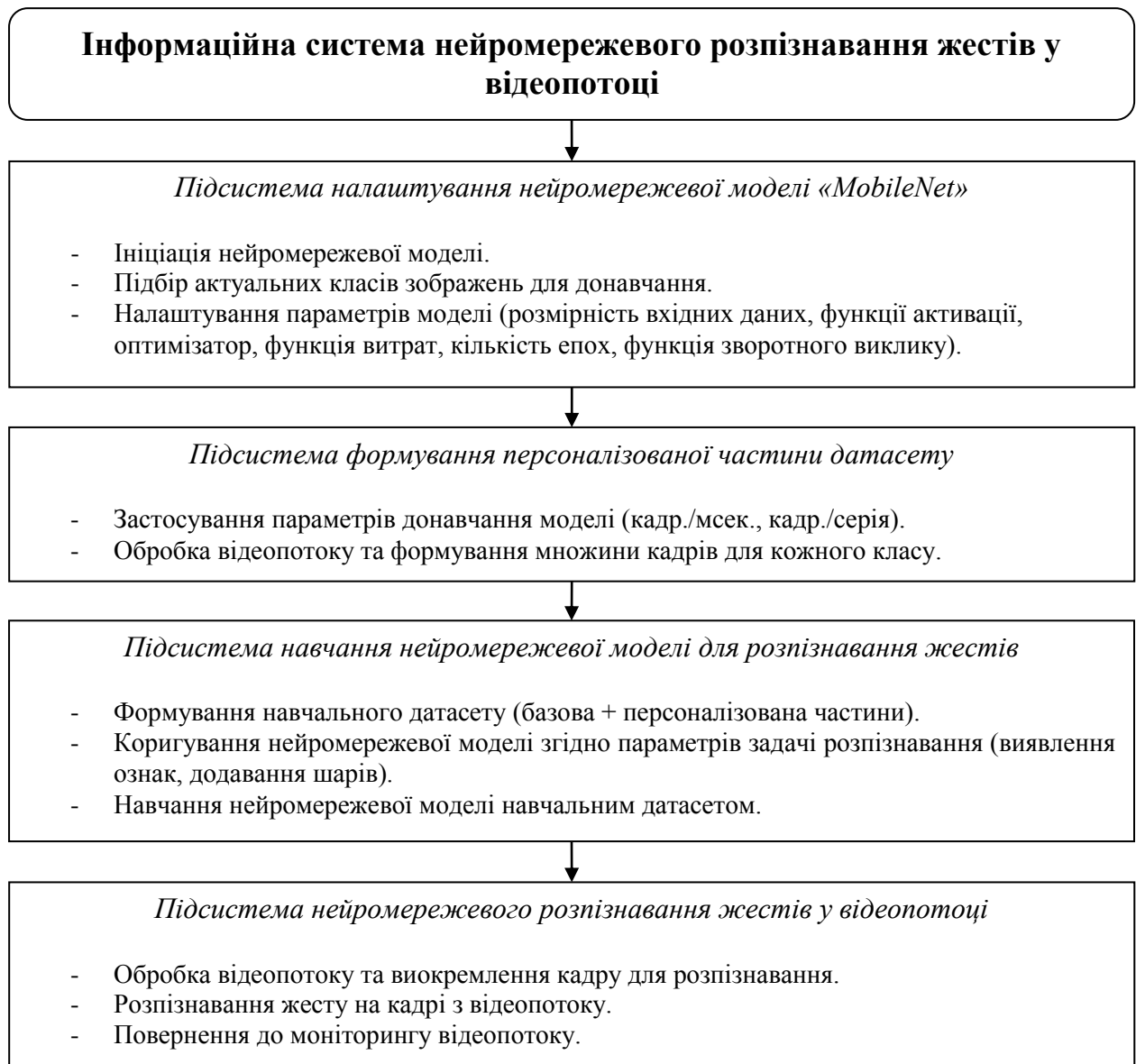


Рисунок 2.9 – Схема інформаційної системи нейромережевого розпізнавання жестів у відеопотоці

Підсистема формування персоналізованої частини включає в себе наступні функції:

– Застосування параметрів донавчання моделі. Для навчання моделі нового образу користувача необхідно додати до наявного навчального сету зображень новий образ у розмірі 244x244x3 пікселів.

– Обробка відеопотоку та формування множини кадрів для кожного класу. Для кожного класу (в даній роботі 6 класів, що відповідають жестам на позначення цифр 0, 1, 2, 3, 4, 5) потрібно додати захоплені відеокамерою користувача продемонстровані ним образи і створити повний навчальний датасет.

Підсистема навчання нейромережевої моделі для розпізнавання жестів включає в себе наступні функції:

– Формування навчально датасету (базова + персоналізована частини). Враховуючи те, що модель MobileNet є попередньо навченою моделлю і містить близько 1000 різноманітних класів, необхідно сформувати повний навчальний датасет шляхом збору нових образів продемонстрованих користувачем і додавання їх до відповідних класів.

– Коригування нейромережевої моделі згідно параметрів задачі розпізнавання, а саме розмірність вхідних даних становить, активаційні функції Relu і Softmax, коефіцієнт швидкості навчання оптимізатора Adam, функцію витрат категоріальної крос-ентропії categoricalCrossentropy, кількість епох, функцію зворотного виклику.

– Навчання нейромережевої моделі навчальним датасетом. Провести навчання нейромережевої моделі відповідно вибраної моделі та архітектури.

Підсистема нейромережевого розпізнавання жестів у відеопотоці включає в себе наступні функції:

– Обробка відеопотоку та виокремлення кадру для розпізнавання. Зображення захоплюється відеокамерою у розмірі 244x244x3 пікселі, його необхідно опрацювати і привести до такого стану, щоб нейронна мережа мала змогу проводити розпізнавання зображень у відеопотоці.

– Розпізнавання жесту на кадрі з відеопотоку. Зображення потрібно захопити, опрацювати, привести до належного виду, щоб нейромережева модель

могла додати його до попередньо наявного датасету та пізніше почати процес розпізнавання зображення.

– Повернення до моніторингу потоку. Після завершення розпізнавання образів потрібно мати можливість призупинити процес безпосереднього розпізнавання і повернутись до тренування чи дотреновування моделі.

Кожна підсистема моделі розпізнавання зображень у відеопотоці налічує список функцій, що дозволяють побудувати зв'язок між компонентами та створити програмну архітектуру методу нейромережевого розпізнавання зображень у відеопотоці.

2.6 Підготовка робочих вхідних даних для системи нейромережевого розпізнавання жестів

Використана модель MobileNet на основі архітектури CNN здатна приймати вхідні зображення різних розмірів (Рисунок 2.10).

Table 1. MobileNet Body Architecture

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
5× Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

Рисунок 2.10 – Параметри моделі MobileNet [28]

Для даного дослідження роботи моделі MobileNet було вирішено обрати розмір вхідних даних – 224x224x3. В свою чергу 224x224 означає, що ширина і висота кожного зображення, захопленого відеокамерою, має становити по 224 пікселі. Цей розмір забезпечує працездатність на екранах мобільних і десктопних пристроїв, а також забезпечує можливість обробки великої кількості зображень.

Позначення x3 позначає кількість каналів у кожному пікселі. У зображення RGB (Red-Green-Blue) кожен піксель має три значення кольору – червоний, зелений та синій. Отже, позначення x3 означає, що кожен піксель зображення має три значення, що відображають три канали кольору.

2.7 Особливості використання спеціалізованих програмних компонентів для нейромережевого моделювання

Для дослідження методу нейромережевого розпізнавання зображень було використано спеціалізовано програмну бібліотеку Tensorflow.js.

Бібліотеку Tensorflow було розроблено компанією Google, а саме її департаментом – Google Brain Team для внутрішнього використання, але згодом було випущене як відкрите програмне забезпечення у 2015 році. У січні 2019 року розробники компанії випустили Tensorflow.js – реалізацію Tensorflow на мові Javascript із забезпеченням основних функцій оригінальної бібліотеки, що була написана на мові Python [29].

Бібліотека надає широкий спектр різноманітних функцій та методів для роботи навчання і розгортання машинного навчання у браузері, а також додавати будь-які функції машинного навчання до будь-якої веб-програми [29].

До переваг бібліотеки Tensorflow.js можна віднести:

- Універсальність, за якої бібліотека дозволяє розробляти програмні засоби з використанням машинного навчання на різних платформах, має безліч функцій та методів, може використовуватись в браузері, та багато іншого.

- Масштабованість, внаслідок якої бібліотека є адаптованою до обчислювальних середовищ CPU та GPU та забезпечує ефективні обчислення і є ефективною для різного штибу завдань.

- Висока підтримка користувачів та спільноти.

- Інтеграція з Google Cloud Service, що дозволяє швидко і якісно інтегрувати сервіси між собою і використовувати потужну інфраструктуру Google середовища [30].

Отож, бібліотека Tensorflow є потужним інструментом, що добре підходить для вирішення задач пов'язаних з машинним навчанням, що пропонує потужний функціонал для створення різного виду рішень для різних видів проблем та завдань.

2.8 Визначення шляхів дослідження нейромережевого розпізнавання жестів у відеопотоці

Для дослідження методу нейромережевого розпізнавання зображень у відеопотоці було вирішено використати мову Javascript, бібліотеку для машинного навчання Tensorflow.js, бібліотеку React для побудови інтерфесу користувача, бібліотеку Bootstrap для верстки веб-застосунку, бібліотеку Three.js для створення об'єктів анімації в інтерфейсі користувача та бібліотеку Jest для розробки та проведення тестування складових частин веб-застосунку.

Метою дослідження є створення веб-застосунку для розпізнавання жестів (жести відповідають позначенням цифр 0, 1, 2, 3, 4, 5) у відеопотоці. З метою дослідження методу і розробки програмного продукту було створено дорожню мапу реалізації застосунку (Рисунок 2.11).

Функціонал застосунку складається з наступних функцій:

- запит на дозвіл браузера ввімкнути камеру на девайсі користувача;
- можливість створення додаткового навчального датасету користувачем;

- можливість вибору кількості вхідних даних для подальшого тренування моделі користувачем;
- можливість запускати навчання моделі у реальному часі;
- можливість переглядати статус навчання моделі у реальному часі;
- можливість почати процес розпізнавання образень;
- можливість призупинити процес розпізнавання зображень.

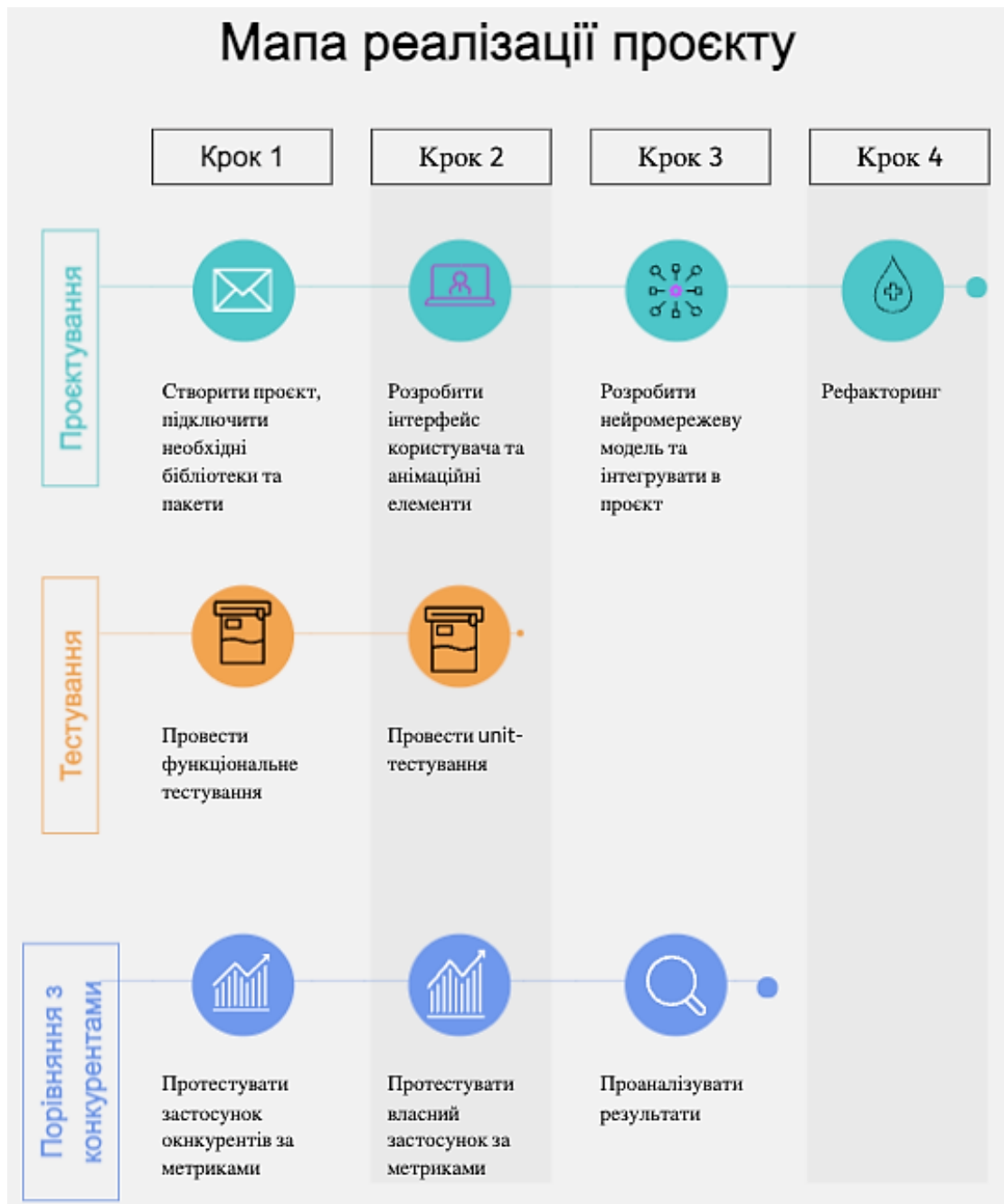


Рисунок 2.11 – Дорожня мапа реалізації застосунку

В якості підходів до тестування було обрано Functional Testing та Unit testing з перевіркою наступних функцій:

- запуск розпізнавання жестів користувача;
- збереження зображень з відеопотоку і відображення кількості збережених кадрів користувачу;
- запуск процесу тренування моделі.

Окремо для дослідження та тестування ефективності нейромережевої моделі для розпізнавання жестів у відеопотоці потрібно дослідити працездатність моделі на метриках Accuracy, Precision, Recall та матриці сплутувань.

2.9 Висновки до розділу 2

Під час виконання другого розділу було обрано архітектуру для створення нейромережевої моделі з метою розпізнавання жестів – CNN та нейромережеву модель – MobileNet, яка дозволяє якісно і точно проводити процес розпізнавання жестів і побудовано схему методу для розпізнавання жестів у відеопотоці.

Було розглянуто всі етапи роботи класичної моделі MobileNet, обрано варіант вхідних даних для моделі та створено діаграму послідовностей для розуміння взаємозв'язку компонент та підсистем інформаційної системи.

Також було проведено аналіз та автоматизацію обробки потоків даних і як наслідок побудовано діаграму варіантів використання системи користувачем.

Було досліджено складові архітектури CNN та побудовано програмну архітектуру системи, що складається з 4 підсистем:

- налаштування нейромережевої моделі MobileNet;
- формування персоналізованої частини датасету;
- навчання нейромережевої моделі для розпізнавання жестів;
- нейромережеве розпізнавання жестів у відеопотоці.

Кожна з підсистем включає в себе набір функцій, що забезпечують функціонування нейромережевої моделі для розпізнавання жестів у відеопотоці та взаємодії всіх її компонентів та інтерфейсу користувача для можливості створення веб-застосунку.

Враховуючи наявність можливості подання в модель різних вхідних даних, було проаналізовано можливості нейромережевої моделі та обрано вигляд вхідних даних для коректної роботи, в даному випадку це зображення розміром 244x244 пікселів з трьома кольоровими каналами в кожному пікселі вхідного зображення.

Було проведено аналіз програмних засобів, мов програмування, бібліотек та архітектур для здійснення поставлених мети і задач, і як результат було обрано мову Javascript, бібліотеку для машинного навчання Tensorflow.js, бібліотеку React для побудови інтерфесу користувача, бібліотеку Bootstrap для верстки веб-застосунку, бібліотеку Three.js для створення об'єктів анімації в інтерфейсі користувача та бібліотеку Jest для розробки та проведення тестування складових частин веб-застосунку.

Крім того, було створено дорожню мапу проєкту, кінцевим результатом якого є створення програмного веб-застосунку, який буде використаний як спосіб дослідження та тестування ефективності і працездатності обраного методу нейромережевого розпізнавання жестів у відеопотоці.

Розділ 3 Експериментальне дослідження методу нейромережевого розпізнавання жестів у відеопотоці

3.1 Структура та функціональне призначення програмних складових системи нейромережевого розпізнавання жестів

Програма для нейромережевого розпізнавання жестів складається з декількох частин, кожна з яких має свою зону відповідальності, а саме:

- інтерфейс користувача;
- програмна реалізація нейромережевого методу для розпізнавання зображень у відеопотоці.

Інтерфейс користувача включає в себе створення дизайну веб-застосунку, програмну реалізацію дизайну для відображення у браузері та оптимізацію анімаційних елементів для коректного відображення у веб-браузері. Основна задача проєктування інтерфейсу користувача – забезпечити якомога більш плавний та інтуїтивний користувацький досвід під час роботи із застосунком та забезпечити коректне відображення даних отриманих з моделі.

Проєктування нейромережевої моделі для розпізнавання зображень у відеопотоці включає в себе створення набору взаємопов'язаних функцій відповідно до моделі MobileNet архітектури CNN, що мають виконувати наступні основні функції:

- ініціалізація нейромережевої моделі та попереднє налаштування параметрів моделі;
- коригування блоку для відеозахоплення камери браузера;
- отримання і збереження кадрів з відеопотоку, їх нормалізація;
- формування датасету та навчання нейромережевої моделі на повному наборі даних;
- розпізнавання жесту з відеопотоку шляхом виокремлення кадру для розпізнавання;
- зупинка розпізнавання і повернення до моніторингу відеопотоку.

Відповідно до вище описаного функціоналу нейромережевої моделі було побудовано діаграму функцій (Рисунок 3.1).

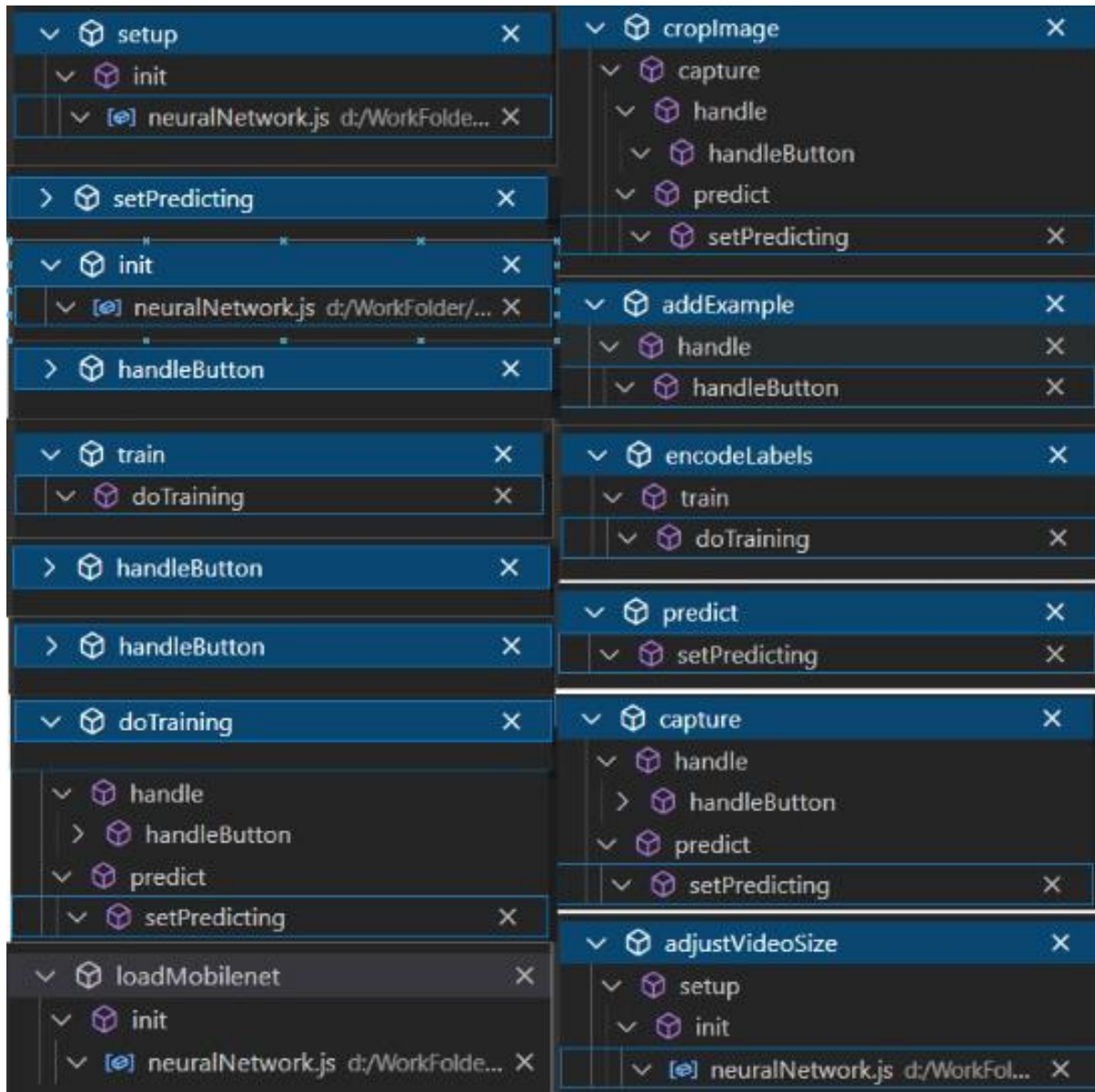


Рисунок 3.1 – Діаграма функцій системи нейромережевого розпізнавання жестів

Деякі з функцій, представлених вище, виконують логіку роботи моделі нейронної мережі, а деякі є допоміжними функціями, які використовуються, щоб забезпечити подальшу роботу застосунку.

На етапі ініціалізації нейромережевої моделі та попереднього налаштування параметрів моделі працює функція `loadMobilenet`, яка викликається за допомогою функції ініціалізації програми `init`. Функція

loadMobilenet завантажує модель Mobilenet для її інтеграції у проєкт і подальшого налаштування параметрів нейромережевої моделі.

На етапі коригування блоку для відеозахоплення камери браузера працюють функції adjustVideoSize та setup, яка в свою чергу викликається функцією ініціалізації програми init. Функція adjustVideoSize викликається для проведення коригування розмірів зображення у відеопотоці для подальшого створення відповідного розміру кадрів для нейромережевої моделі та коректного відображення блоку з відео в елементі відео webcamElement на екрані веб-застосунку.

На етапі отримання і збереження кадрів з відеопотоку, їх нормалізації та додавання до наявного набору зображень відповідного класу виконуються функції handleButton, capture, cropImage та encodeLabels. Функція handleButton є допоміжною функцією, що запускає процес захоплення і обробки зображень по кліку користувача на кнопку, що відповідає певному класу зображення для розпізнавання. Функція capture виконує захоплення зображень з відеокамери користувача. Функція cropImage приймає вхідне зображення та виконує обрізання кадру для подальшої ідентифікації. Функція encodeLabels дозволяє підготувати зображення в такий вид, щоб нейромережева модель могла провести навчання на тренувальному датасеті.

На етапі формування датасету та навчання нейромережевої моделі на повному наборі даних виконуються функції addExample, doTraining та train. Функція addExample отримує з відеокамери користувача зображення, яке обробляється за допомогою функцій capture, cropImage та encodeLabels, і додає його до тренувальної вибірки зображень для формування повного навчального датасету. Для запуску навчання нейромережевої моделі виконується функція doTraining, яка запускає функцію для тренування моделі – train, а також виводить користувачу повідомлення про те, що модель була навчання і можна починати процес розпізнавання образів або навпаки. Для навчання моделі використовується функція train, яка складається з декількох кроків та дозволяє

провести донавчання моделі на повному датасеті, що сформований на попередньому кроці роботи з моделлю.

На етапі розпізнавання жесту з відеопотоку шляхом виокремлення кадру для розпізнавання виконуються функції `capture`, `predict` та `setPredicting`. Функція `capture` захоплює зображення з веб-камери та підготовляє його для подачі як вхідні дані моделі. Функція `setPredicting` керує процесом розпізнавання жестів, якщо змінна `isPredicting` дорівнює `true`, то запускається функція `predict`. Функція `predict`, яка отримує передбачення моделі на обробленому вхідному масиві зображень і виводить результат у блоці для результатів передбачення моделі у інтерфейсі користувача.

На етапі зупинки розпізнавання і повернення до моніторингу відеопотоку виконується функція `setPredicting`, змінна `isPredicting` якої встановлюється у значення `false` і таким чином припиняється цикл розпізнавання жестів у відеопотоці.

3.2 Особливості реалізації програмних складових системи нейромережевого розпізнавання жестів у відеопотоці

На початку роботи застосунку важливо провести корекцію блоку для захоплення зображень камерою браузера, щоб переконатись, що подальше розпізнавання жестів є можливим. Для цього використовується функція `adjustVideoSize`. Дана функція отримує висоту і ширину відео користувача в якості вхідних параметрів і проводить коригування розмірів у випадках, коли ширина більша за висоту відео, або навпаки. Для точного відображення відеопотоку у відповідному вікні інтерфейсу користувача виконується функція `setup`, яка встановлює підключення до веб-камери, отримує відеопотік з веб-камери користувача (попередньо браузер очікує підтвердження дозволу на використання відеокамери у браузері), налаштовує розмір елемента відео `webcamElement` та дозволяє розпочати роботу із моделлю.

Важливим етапом роботи нейромережевої моделі для розпізнавання образів є захоплення відеокамерою зображення та його обробка для подальшого тренування моделі. Функція `capture` виконує захоплення зображень з відеокамери користувача, відображає кадр дзеркально (для отримання зображення з орієнтацією зліва направо), обрізає зображення для отримання квадратного кадру за допомогою функції `cropImage` та підготовляє передачу зображення до моделі шляхом приведення його до 4D тензору (бо модель очікує батчі – групи даних для передачі моделі для одночасної обробки, розміром `[batch_size, height, width, channels]`). Функція `cropImage` в свою чергу приймає зображення та вирізає квадратний кадр попередньо обчислюючи початкову точку обрізання в центрі оригінального зображення. Функція `encodeLabels` отримує масив даних – батчів, які мають відповідні мітки зображень та перетворює масив міток у вектор `one-hot` з подальшим об'єднанням векторів міток в один тензор, який використовується під час навчання моделі.

Після підготовки навчального датасету відбувається донавчання моделі, яке виконується за допомогою функції `train`. Спочатку функція очищає попередні значення параметрів моделі і втрат, які були збережені у змінній `loss`, далі створює модель використовуючи `TensorFlow.js Sequential API`, яка має 3 шари: згортковий шар, який перетворює вихідні дані вектора, прихований `Dense` шар з 100 нейронами та функцією активації `Relu` і останній `Dense` шар з 10 нейронами та функцією активації `Softmax`. Далі відбувається компіляція моделі з використанням алгоритму оптимізації `Adam` з коефіцієнтом навчання `0.0001` та функцією втрати `categoricalCrossentropy`. Наступним кроком є тренування моделі, для цього використовується метод `fit`, який виконує декілька епох навчання (в даному випадку 10) на вхідних даних (зображеннях) та вихідних даних (мітках зображень), оновлюючи параметри моделі після кожного батчу. Крім того, встановлюється функція зворотного виклику для виведення значень втрат після кожного батчу.

За успішного виконання процесу навчання нейромережевої моделі, користувач має змогу побачити інформаційне повідомлення – спливаючий поп-

ап зеленого кольору з текстом про те, що модель готова до розпізнавання образів (Рисунок 3.2). За умови неуспішного виконання тренування нейромережевої моделі, користувач має змогу побачити інформаційний поп-ап з повідомленням червоного кольору про те, що сталась помилка (Рисунок 3.3).

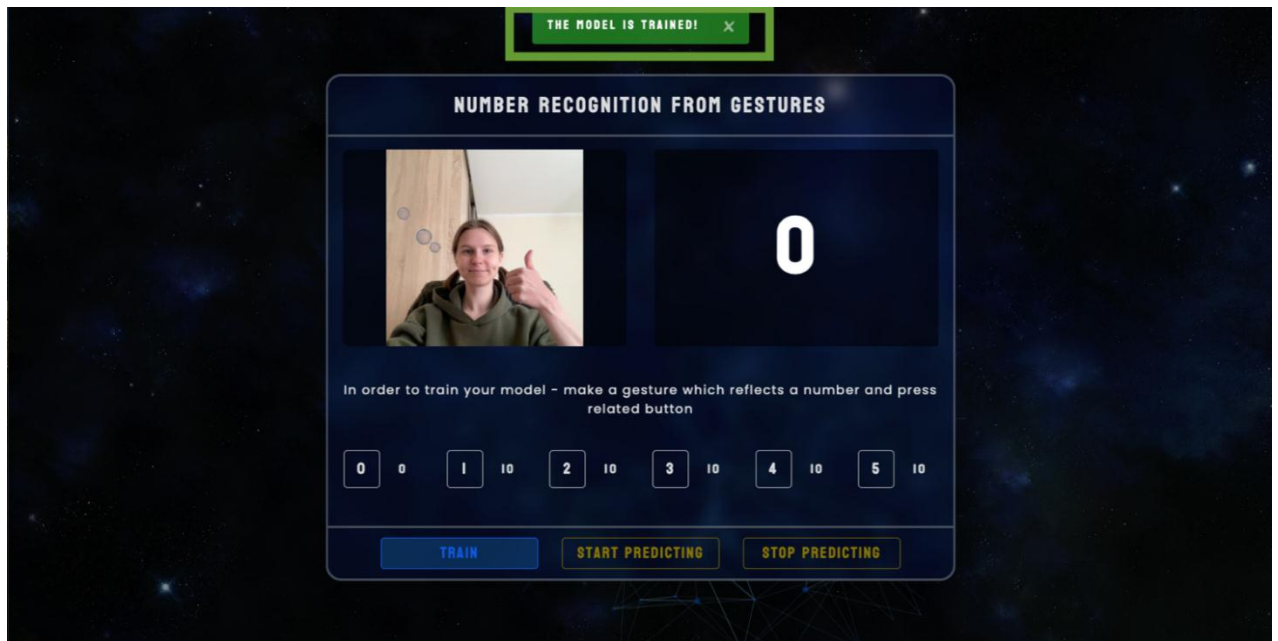


Рисунок 3.2 – Успішне виконання тренування модель. Модель готова до розпізнавання жестів

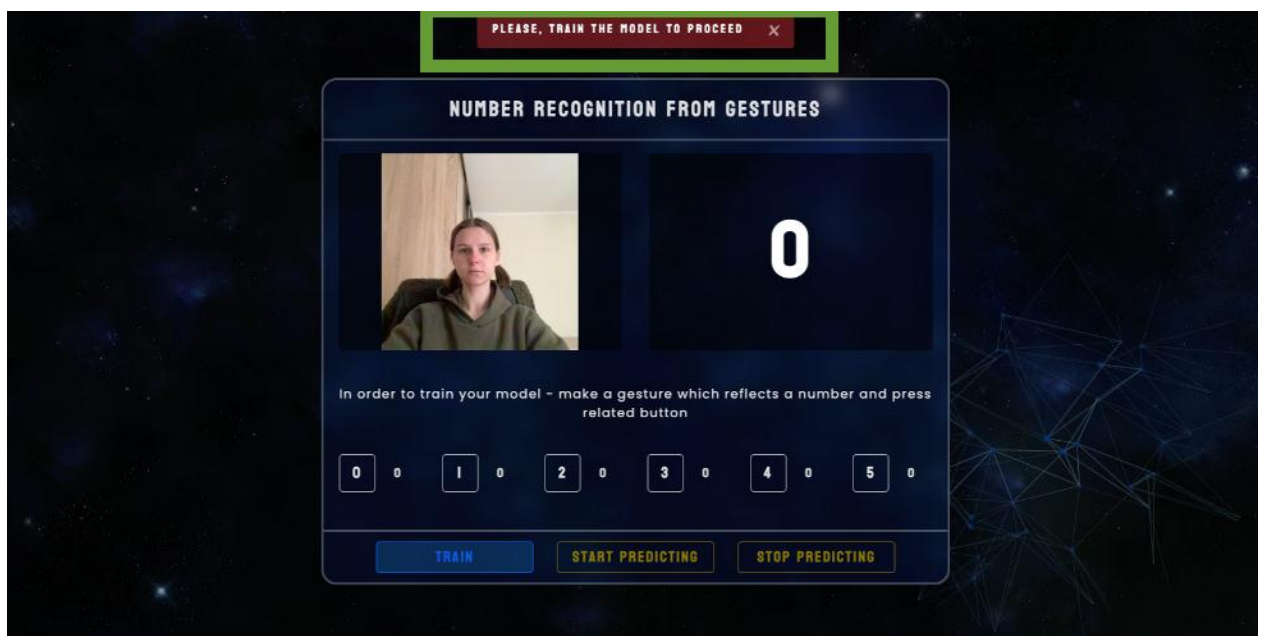


Рисунок 3.3 – Неуспішне виконання тренування модель. Модель не готова до розпізнавання жестів

За основний результат – розпізнавання жестів у відеопотоці, відповідає функція `predict`. На першому кроці функція отримує оброблене зображення за допомогою функції `capture`, та кодування його у вектор `one-hot` за допомогою функції `encodeLabels`. Після отримання вектору, він подається на вхід до натренованої моделі і модель генерує передбачення продемонстрованого жесту користувачем. Отримані передбачення відображаються у блоці для результатів у інтерфейсі користувача (Рисунок 3.4).

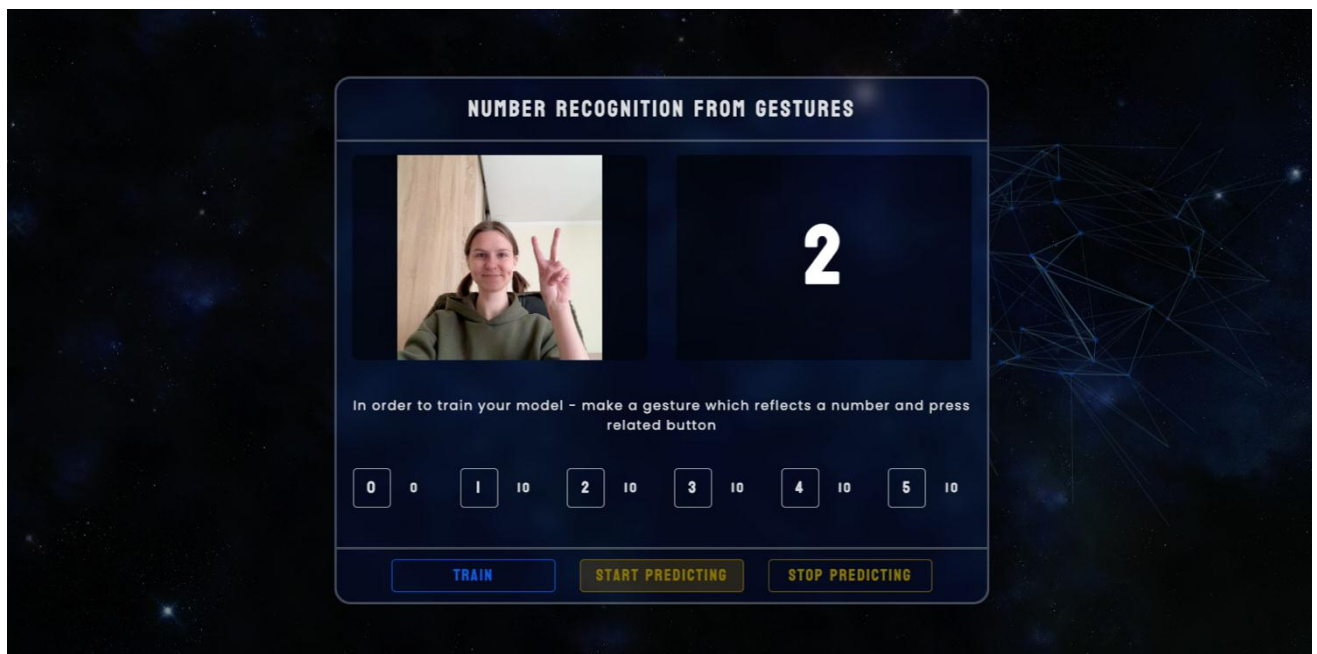


Рисунок 3.4 – Результат розпізнавання нейромережевою моделю класу зображення, що відповідає позначенню цифри «2»

За потреби, користувач має змогу припинити розпізнавання жестів та додатково провести донавчання нейромережевої моделі, користуючись інтерфейсом застосунку, з метою досягнення більшої точності розпізнавання жестів.

3.3 Тестування системи розпізнавання жестів у відеопотоці

Наступним етап після програмування веб-застосунку для нейромережевого розпізнавання жестів у відеопотоці є процес розробки автоматичних тестів та проведення мануального тестування.

В якості підходів до тестування веб-застосунку було обрано провести Functional Testing та Unit testing. Для тестування застосунку було створено декілька Unit-тестів із використанням бібліотеки Jest та кілька функціональних тестів для перевірки коректної роботи основних частин створеного веб-застосунку.

Функція, що запускає розпізнавання жестів користувача, є важливою для тестування, адже від цього буде залежати, чи почне нейромережева модель процес розпізнавання жестів до того, як користувач почне демонструвати жести у відеопотоці. Acceptance criteria, тобто умови вважати задачу успішно виконаною і коректно працюючою, для даного тесту наступні:

- При передачі у функцію параметру `predicting`, що дорівнює `false`, функція `setPredicting` поверне значення, що дорівнює `false`.

- При передачі у функцію параметру `predicting`, що дорівнює `true`, функція `setPredicting` поверне значення, що дорівнює `true`.

Для перевірки acceptance criteria 1 було створено Unit-тест, який при отримання параметру `predicting=falsy` має повертати `falsy` результат.

Результат виконання тесту підтвердив коректність роботи даного фрагменту коду (Рисунок 3.5).

Для перевірки acceptance criteria 2 було створено другий Unit-тест, який при отримання параметру `predicting=true` має повертати `truthy` результат.

Результат виконання тесту підтвердив коректність роботи даного фрагменту коду (Рисунок 3.6).

```

PASS ./neuralNetwork.test.js
  ✓ return (5 ms)

Test Suites: 1 passed, 1 total
Tests:      1 passed, 1 total
Snapshots:  0 total
Time:       1.535 s
Ran all test suites.

```

Рисунок 3.5 – Результат виконання Unit-тест 1. Acceptance criteria 1. Пройдено успішно.

```

PASS ./neuralNetwork.test.js
  ✓ return (7 ms)

Test Suites: 2 passed, 2 total
Tests:      2 passed, 2 total
Snapshots:  0 total
Time:       2.018 s
Ran all test suites.

```

Рисунок 3.6 – Unit-тест 2. Acceptance criteria 2. Пройдено успішно.

Для тестування функції збереження зображень з відеопотоку і відображення кількості збережених кадрів користувачу було проведено Functional-тест. Acceptance criteria для даного тесту наступні:

- При натисканні на кнопку, що відповідає класу зображення, відбувається захоплення зображень з відеокамери користувача і відображення кількості зібраних кадрів у інтерфейсі користувача навпроти відповідної кнопки. Кількість зібраних кадрів при одноразовому натисканні на кнопку становить 10 кадрів.

- При повторному натисканні кнопки, що відповідає певному класу зображення, кількість кліків збільшується на 10 і відображається в інтерфейсі користувача.

При виконанні тесту було перевірено коректність роботи функції збереження зображень з відеопотоку та відображення користувачу кількості кадрів. Так, по одному кліку на кнопку з класом зображення, що відповідає позначенню цифри «1» було зібрано 10 кадрів, тож acceptance criteria 1 було виконано (Рисунок 3.7).

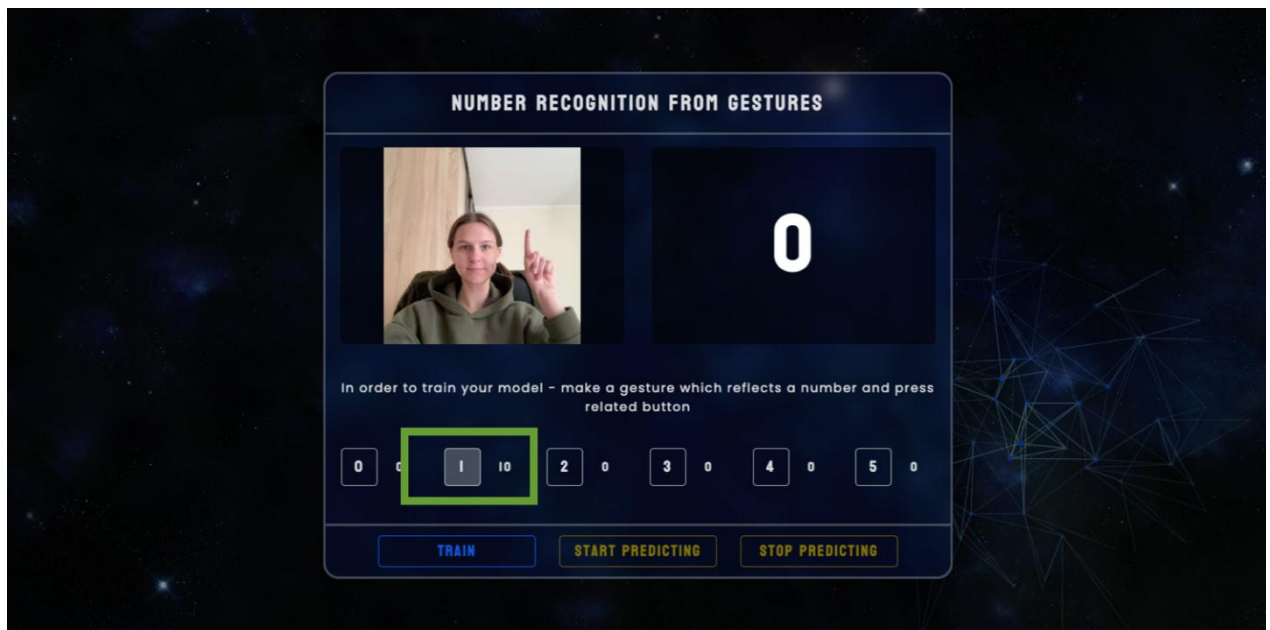


Рисунок 3.7 – Результат виконання Functional-test 1. Acceptance criteria 1.

Пройдено успішно.

Відповідно, натиснувши 4 рази на клас зображення, що відповідає позначенню цифри «2» було зібрано 40 зображень, тож acceptance criteria 2 також виконано (Рисунок 3.8).

Для тестування функції запуску процесу тренування моделі було використано Functional-тест. Acceptance criteria для даного тесту наступні:

– Якщо користувач не продемонстрував жодного жесту або не натиснув на кнопку відповідного класу зображення при демонструванні жесту, тренування моделі неможливе, потрібно відобразити користувачу інформаційне повідомлення з проханням додати зображення до навчального датасету і натренувати модель.

– Якщо користувач продемонстрував жести і модель пройшла процес навчання, потрібно відобразити інформаційне повідомлення з текстом про те, що модель готова до розпізнавання жестів у відеопотоці.



Рисунок 3.8 – Результат виконання Functional-тест 2. Acceptance criteria 2.
Пройдено успішно.

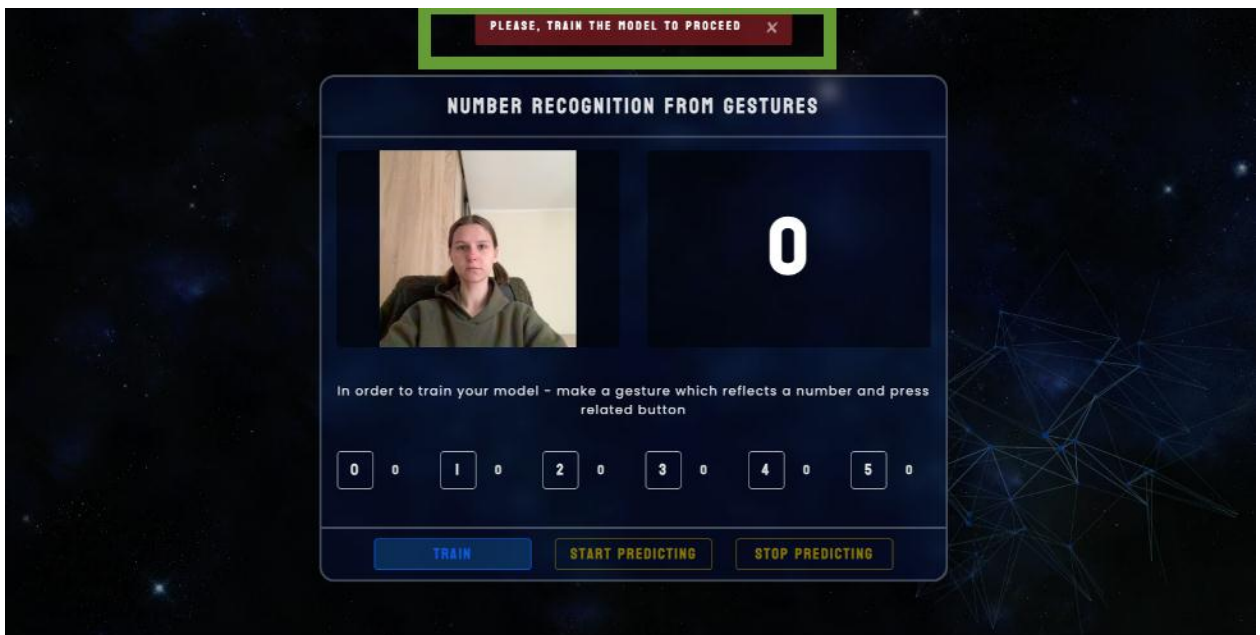


Рисунок 3.9 – Результат виконання Functional-тест 3. Acceptance criteria 1.
Пройдено успішно.

При виконанні тесту було перевірено роботу функції тренування моделі за acceptance criteria 1, де користувач не продемонстрував жодних жестів, застосунок відреагував очікувано, тож Acceptance criteria 1 виконано (Рисунок 3.9). При тестуванні функціоналу тренування моделі за acceptance criteria 2, де користувач продемонстрував жести для додавання нових кадрів, застосунок відреагував очікувано, тож acceptance criteria 2 також виконано (Рисунок 3.10).

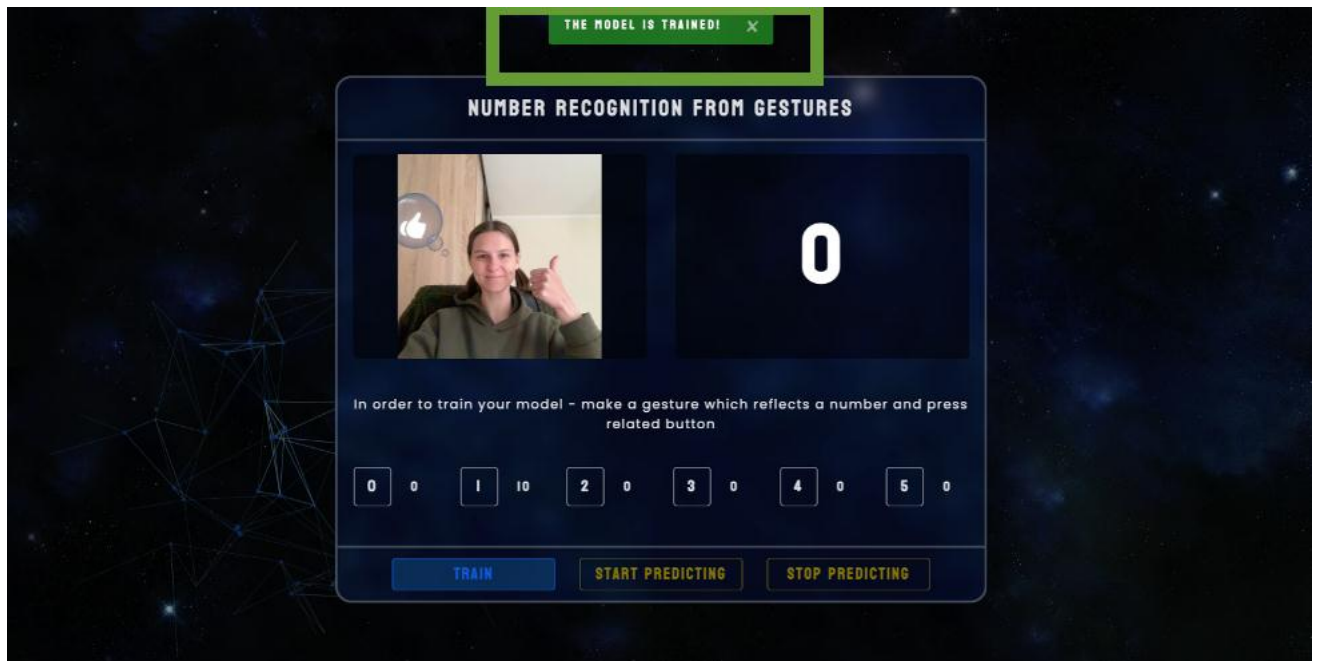


Рисунок 3.10 – Результат виконання Functional-тест 4. Acceptance criteria 2.

Пройдено успішно.

Для користування веб-застосунком для нейромережевого розпізнавання жестів у відеопотоці є рекомендовані параметри до апаратної та програмної частин інформаційної системи.

Рекомендовані апаратні вимоги для розгортання застосунку наступні:

- Процесор Intel Core i5 або еквівалентний. Застосунок вимагає обробки великої кількості даних, швидкі види процесорів забезпечують високий рівень швидкодії та оптимальної роботи застосунку.

- Оперативна пам'ять мінімум 8Гб. Чим більша за об'ємом оперативна пам'ять, тим швидше будуть виконуватись складні обчислення, уникаючи затримок.

– Графічна підсистема з підтримкою WebGL. Це необхідно для оптимального відображення інтерфейсу, графіки, а також анімаційних елементів застосунку.

– Веб-камера вбудована або зовнішня з підтримкою HD або Full HD.

Рекомендовані програмні вимоги для розгортання застосунку наступні:

– Операційна система Windows 10, macOS 10.14 (Mojave) або вище, Linux.

– Найновіші версії браузерів Google Chrome, Mozilla Firefox або Safari.

– Браузер повинен мати ввімкнений Javascript двигун для коректної роботи застосунку.

3.4 Аналіз функціональності системи нейромережевого розпізнавання жестів

Створений програмний застосунок дозволяє користувачу взаємодіяти з нейромережевою моделлю для розпізнавання жестів у відеопотоці за допомогою інтерфейсу та, за потреби, текстових підказок.

Для початку роботи потрібно запустити застосунок, далі користувач побачить перед собою програмний інтерфейс (Рисунок 3.11).

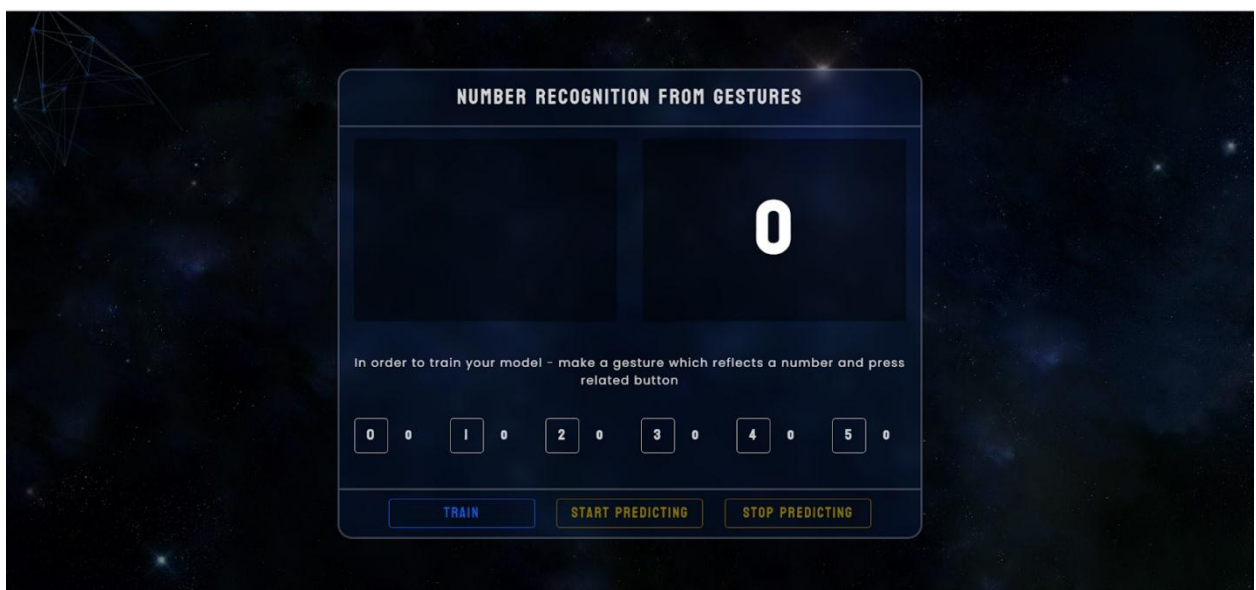


Рисунок 3.11 – Інтерфейс програми

Для того, щоб увімкнути камеру, потрібно дати дозвіл браузеру на використання камери, натиснувши на кнопку «Дозволити цього разу», що дозволить браузеру використати відеокамеру в поточній сесії користувача або «Дозволити під час кожного відвідування», що дозволить браузеру використовувати відеокамеру для даного застосунку незалежно від сесії на девайсі (Рисунок 3.12).

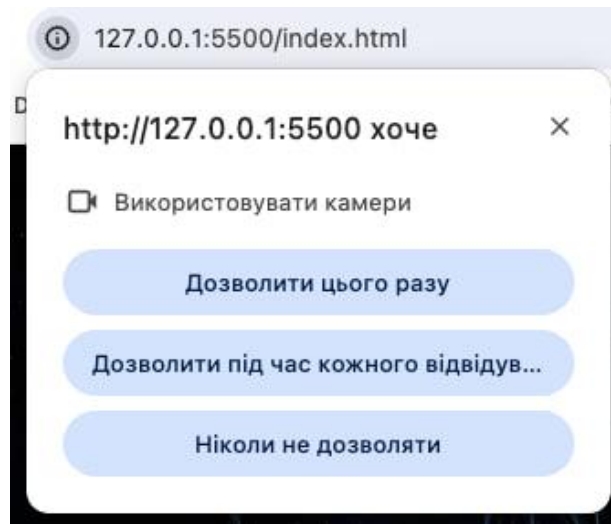


Рисунок 3.12 – Дозвіл браузеру на використання камери

Коли відео завантажиться, можна починати демонстрацію жестів для подальшого їх розпізнавання у відеопотоці. Для цього потрібно показати певний жест, в даному випадку це цифра від 0 до 5 і натиснути на кнопку, що відповідає продемонстрованому жесту і відповідному класу зображення. Цей крок необхідно здійснити щонайменше один раз для продовження роботи з моделлю. Кількість створених зображень буде відображена справа від кнопки з відповідною цифрою (Рисунок 3.13).

Після завершення демонстрації жестів, потрібно натиснути на кнопку «Train» для того, щоб модель почала навчання моделі з метою розпізнаванню жестів користувача під час етапу розпізнавання жестів і дочекатись поп-ап повідомлення про те, що модель готова до використання (Рисунок 3.14).

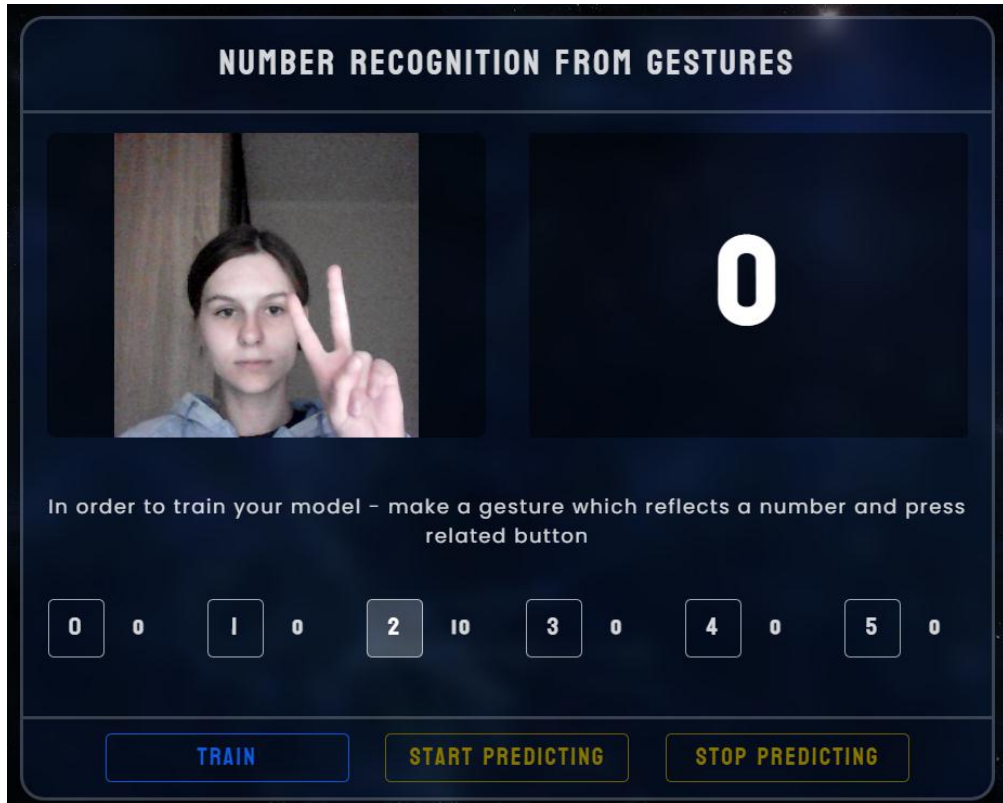


Рисунок 3.13 – Показ жесту для подальшого навчання моделі

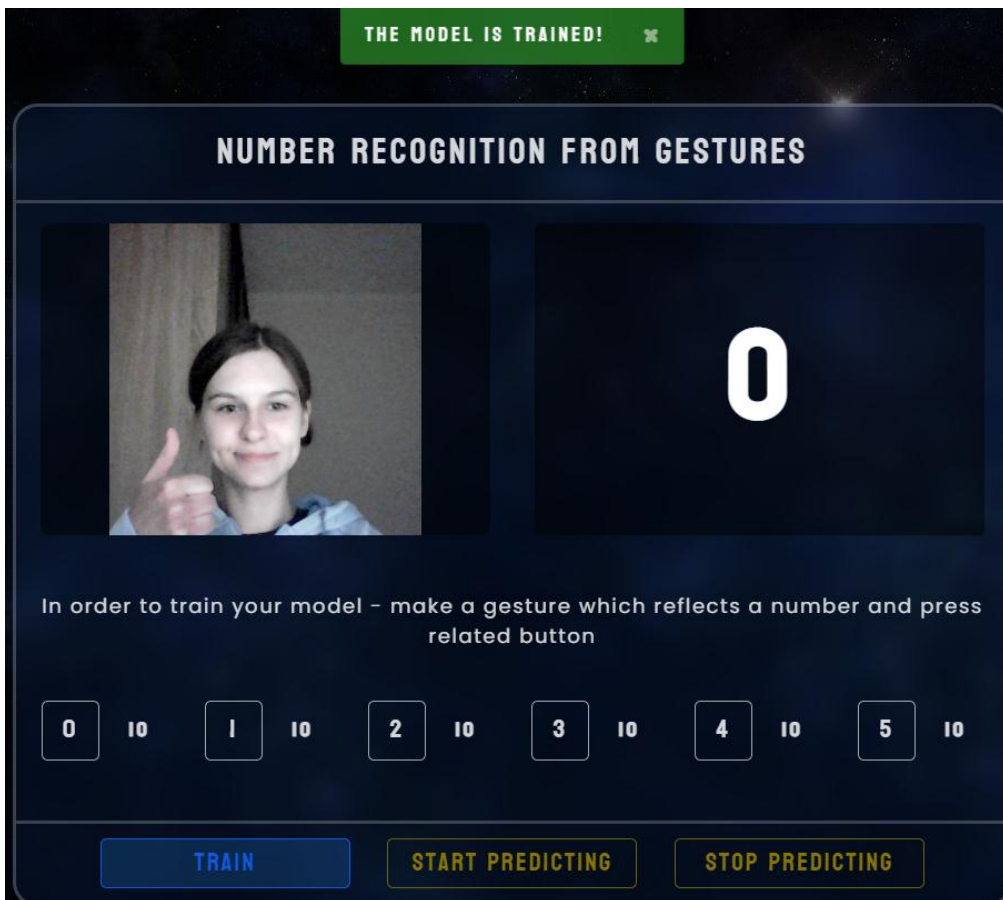


Рисунок 3.14 – Результат виконання навчання моделі

Далі настає процес використання нейромережевої моделі для розпізнавання жестів у відеопотоці. Для цього потрібно натиснути на кнопку «Start predicting», показати жест, і модель почне розпізнавання жесту користувача із «передбаченням» продемонстрованого жесту, що буде відображено справа від камери користувача у блоці результату розпізнавання жесту (Рисунок 3.15).



Рисунок 3.15 – Результат розпізнавання жестів

Для того, щоб завершити або призупинити роботу програми, слід натиснути на кнопку «Stop predicting».

3.5 Результати досліджень з розпізнавання жестів у відеопотоці

Для дослідження розпізнавання жестів у відеопотоці буде використовуватись метрика Accuracy, Precision, Recall та матриця сплутувань.

В якості валідаційного набору буде використано набір попередньо зібраних з відеокамери зображень жестів цифр 0, 1, 2, 3, 4, 5 по 100 зображень для кожного класу зображення.

Нижче наведено експеримент щодо впливу розміру набору зображень для навчання моделі для розпізнавання, освітлення (денне, вечірнє) та наявності обличчя користувача на фоні при демонстрації жестів (наявне, відсутнє) (Таблиця 3.1).

Таблиця 3.1 – Вплив параметрів на модель

<i>К-сть зображень</i>	<i>Параметри:</i>		<i>Метрики:</i>		
	<i>Освітлення</i>	<i>Обличчя користувача на фоні</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
10	денне	наявне	0.67	0.69	0.68
10	денне	відсутнє	0.71	0.74	0.74
10	вечірнє	наявне	0.65	0.64	0.68
10	вечірнє	відсутнє	0.68	0.69	0.69
50	денне	наявне	0.83	0.85	0.86
50	денне	відсутнє	0.85	0.87	0.89
50	вечірнє	наявне	0.77	0.76	0.78
50	вечірнє	відсутнє	0.79	0.79	0.8
100	денне	наявне	0.91	0.9	0.92
100	денне	відсутнє	0.93	0.94	0.91
100	вечірнє	наявне	0.87	0.88	0.88
100	вечірнє	відсутнє	0.89	0.88	0.89

Результати таблиці продемонстровані у вигляді графіка (Рисунок 3.16).



Рисунок 3.16 – Вплив параметрів на модель

За висновками експерименту та його результатів видно, що при виборі параметрів для найкращої роботи моделі варто вибирати параметри: кількість зображень = 100, освітлення – денне, обличчя користувача на фоні має бути відсутнім. За таких параметрів матриця сплутувань набуває наступного вигляду (Таблиця 3.2).

Таблиця 3.2 – Матриця сплутувань

	0	1	2	3	4	5
0	20	0	0	0	0	0
1	1	19	0	0	0	0
2	0	0	20	0	0	0
3	0	0	0	20	0	0
4	0	1	0	1	18	0
5	0	0	0	1	0	19

Як видно на матриці сплутувань, на головній діагоналі розташована кількість зображень відповідного класу, що пройшла розпізнавання успішно, якщо якийсь екземпляр зображення було помилково прийнято за інший клас зображення, він відображається у відповідній колонці.

На наборі у 100 зображень при денному освітленні та відсутності обличчя користувача на фоні було проведено дослідження впливу параметрів навчання методу: кількість епох навчання та розміру батчу (Таблиця 3.3).

Таблиця 3.3 – Вплив параметрів нейромережевої моделі

<i>Параметри:</i>		<i>Метрики:</i>		
<i>batch</i>	<i>epoch</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
16	10	0.89	0.93	0.92
32	10	0.83	0.91	0.9
64	10	0.98	0.99	1
16	20	0.83	0.95	0.94
32	20	0.85	0.89	0.88
64	20	0.91	1	0.98

Результати таблиці продемонстровані у вигляді графіка (Рисунок 3.17).

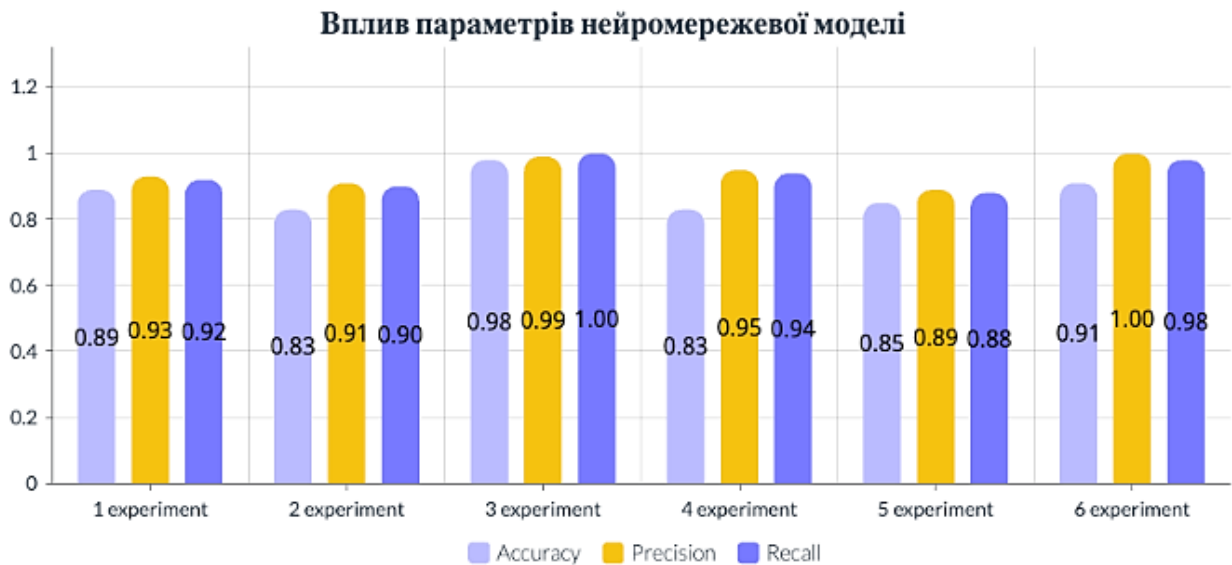


Рисунок 3.17 – Вплив параметрів нейромережевої моделі

Як видно з експерименту, при збільшенні кількості епох навчання та збільшенні батчу відбувається покращення роботи даної нейромережевої моделі. За результатами дослідження найкращі показники має модель з параметрами батч = 64 та кількістю епох – 10.

Дана розроблена система може використовуватись для точного і швидкого розпізнавання жестів у відеопотоці, що може бути корисним для людей, які мають проблеми із зором, слухом, опорно-руховим апаратом та іншими порушеннями, при спілкуванні з іншими людьми або взаємодії із комп'ютерними програмами.

При експериментальному дослідженні розробленої систему було досліджено вплив різних параметрів, таких як розмір батчу, кількість епох навчання, освітлення, характеристик фону зображення та кількості зображень в наборі для розпізнавання моделлю. При дослідженні моделі за допомогою метрик Accuracy, Precision, Recall та матриці сплутувань вдалось досягнути таких показників точності роботи нейромережевої моделі: Accuracy – 0.98, Precision – 0.99 та Recall – 1 при розмірі батчу – 64, кількості епох – 10, кількості

зображень для навчання моделі – 100 при денному освітленні та відсутності сторонніх об'єктів на фоні зображення.

3.6 Висновки до розділу 3

За результатами побудови діаграми класів та визначення структури і функціонального призначення програмних складових системи нейромережевого розпізнавання жестів було розроблено і програмно реалізовано веб-застосунок для нейромережевого розпізнавання жестів у відеопотоці за моделлю MobileNet архітектури CNN.

Застосунок було розроблено із використанням потужної бібліотеки Tensorflow.js та допоміжних інструментів для створення програми.

Створений веб-застосунок було протестовано за допомогою тестових підходів Functional testing та Unit testing і виявлено коректну роботу всіх складових застосунку, що містять наступний функціонал:

- запитування дозволу браузера ввімкнути камеру на девайсі користувача;
- створення додаткового навчального датасету користувачем;
- вибору кількості вхідних даних для подальшого тренування моделі користувачем;
- запуску навчання моделі у реальному часі;
- перегляду статусу навчання моделі у реальному часі;
- початку процесу розпізнавання зображень;
- зупинки розпізнавання зображень.

Було також проведено дослідження ефективності моделі за допомогою метрик Accuracy, Precision, Recall та матриці сплутувань. Дослідження показали, що модель має високу ефективність у розпізнаванні жестів у відеопотоці, зокрема було досягнуто показників точності роботи моделі Accuracy – 0.98, Precision – 0.99 та Recall – 1 при розмірі батчу – 64, кількості епох – 10, кількості зображень для навчання моделі – 100 при денному освітленні та відсутності

сторонніх об'єктів на фоні зображення. Подальші дослідження будуть спрямовані на розширення кількості класів зображень для розпізнавання та оптимізацію ресурсів у веб-середовищі для пришвидшення роботи програмного застосунку з метою розпізнавання жестів.

Загальні висновки

Метою кваліфікаційної роботи бакалавра було визначено покращення розпізнавання жестів у відеопотоці, для чого була виконана розробка та програмна реалізація нейромережевого методу розпізнавання жестів у відеопотоці у вигляді вебресурсу. В результаті виконання кваліфікаційної роботи бакалавра було досягнуто мети покращення розпізнавання жестів у відеопотоці, для чого було розв'язано наступні задачі:

- виконано дослідження предметної області розпізнавання жестів у відеопотоці;
- в рамках дослідження предметної області виконано огляд теоретичних підходів щодо розпізнавання жестів у відеопотоці;
- виконано аналіз наукових надбань й існуючих програмних рішень в області розпізнавання жестів у відеопотоці;
- розроблено метод нейромережевого розпізнавання жестів у відеопотоці;
- на основі розробленого методу виконано проектування структури системи нейромережевого розпізнавання жестів у відеопотоці;
- здійснено вибір засобів розробки для системи нейромережевого розпізнавання жестів у відеопотоці;
- здійснено програмну реалізацію системи розпізнавання жестів;
- проведено тестування розробленої системи нейромережевого розпізнавання жестів у відеопотоці;
- здійснено дослідження ефективності розробленого методу з використанням розробленої системи нейромережевого розпізнавання жестів.

В результаті було спроектовано та побудовано веб-застосунок, що виконує наступні основні функції:

- запитування дозволу браузера ввімкнути камеру на девайсі користувача;
- створення додаткового навчального датасету користувачем;

- вибору кількості вхідних даних для подальшого тренування моделі користувачем;
- запуску навчання моделі у реальному часі;
- перегляду статусу навчання моделі у реальному часі;
- початку процесу розпізнавання зображень;
- зупинки розпізнавання зображень;
- демонстрація результату розпізнавання жестів у відеопотоці.

Для розробки інформаційної системи було використано мову програмування Javascript, бібліотеку для машинного навчання Tensorflow.js, бібліотеку React для побудови інтерфесу користувача, бібліотеку Bootstrap для верстки веб-застосунку, бібліотеку Three.js для створення об'єктів анімації в інтерфейсі користувача та бібліотеку Jest для розробки та проведення тестування складових частин веб-застосунку.

Результат виконаної кваліфікаційної роботи бакалавра повністю відповідає поставленому завданню роботи. За результатами дослідження ефективності було виявлено значене покращення процесу нейромережевого розпізнавання жестів у відеопотоці за метриками Accuracy, Precision, Recall та матриці сплутувань, які показали ефективність роботи моделі з такими результатами, як Accuracy – 0.98, Precision – 0.99 та Recall – 1 при розмірі батчу – 64, кількості епох – 10, кількості зображень для навчання моделі – 100 при денному освітленні та відсутності сторонніх об'єктів на фоні зображення..

Актуальність досліджень і розробки програмних продуктів з метою розпізнавання жестів є досить високою, тому перспективи розробленого методу і програмної системи є широкими. Зокрема, розширення кількості класів зображень для розпізнавання та оптимізація ресурсів у веб-середовищі для пришвидшення роботи програмного застосунку з метою розпізнавання жестів.

Розроблена система призначена для розпізнавання жестів користувача у відеопотоці. Реалізована автоматизація розпізнавання жестів у відеопотоці може бути імплементована у сферах охорони здоров'я, сфері розваг, автомобільній галузі. Напрямами практичного використання розробленої інформаційної

системи визначено автоматизацію розпізнавання жестів користувача у відеопотоці для покращення комунікації користувачів без використання вербальних засобів комунікації.

Основні наукові та практичні результати доповідалися на Міжнародному конкурсі студентських наукових робіт «Black Sea Science 2023» (зайнято 2 місце в фіналі), за темою кваліфікаційної роботи бакалавра автором було виконано наукову публікацію «Gesture recognition using a neural network in real time» [31]. Також автором було виконано доповідь «An Approach to Using MobileNet CNN-model for Gesture Recognition» на XXIII Міжнародній науково-практичній конференції «Problems of Science and Technology: the Search for Innovative Solutions» (May 15-17, 2024, Munich, Germany) і відповідну публікацію [32].

Перелік посилань

1. Wikipedia. Вади слуху. URL: https://uk.wikipedia.org/wiki/Вади_слуху.
2. Wikipedia. Зображення. URL: <https://uk.wikipedia.org/wiki/Зображення>.
3. Imena.ua. Як технологія розпізнавання жестів завойовує різні галузі і яке майбутнє чекає цю технологію. URL: [Imena.ua. Як технологія розпізнавання жестів завойовує різні галузі і яке майбутнє чекає цю технологію. URL: https://www.imena.ua/blog/how-gesture-recognition-conquers-different-industries/](https://www.imena.ua/blog/how-gesture-recognition-conquers-different-industries/).
4. Unite.ai. Розпізнавання зображень проти комп'ютерного зору. URL: [Imena.ua. Як технологія розпізнавання жестів завойовує різні галузі і яке майбутнє чекає цю технологію. URL: https://www.unite.ai/uk/розпізнавання-зображень-проти-комп%27ютерного-зору/](https://www.unite.ai/uk/розпізнавання-зображень-проти-комп%27ютерного-зору/).
5. Library.nlu.edu.ua. Ідентифікація людини за ознаками зовнішності. URL: https://library.nlu.edu.ua/POLN_TEXT/KNIGI/1_DISKI/KRIM/html/%D0%93%D0%BB%D0%B0%D0%B2%D0%B0_12.htm.
6. Wikipedia. Жест. URL: <https://uk.wikipedia.org/wiki/Жест>.
7. С.С. Кондратюк. Моделювання та розпізнавання жестів української дактильної абетки за допомогою кросплатформених технологій. 2021. С. 35. URL: https://csc.knu.ua/media/filer_public/fe/64/fe643504-1482-48cb-b64e-db437738da63/dis_kondratiuk_ss.pdf.
8. Zaptest. Комп'ютерний зір – це майбутнє автоматизації тестування програмного забезпечення – історія минулого, сьогодення та майбутнього. URL: [Imena.ua. Як технологія розпізнавання жестів завойовує різні галузі і яке майбутнє чекає цю технологію. URL: https://www.zaptest.com/uk/компютерний-зір-це-майбутнє-автом](https://www.zaptest.com/uk/компютерний-зір-це-майбутнє-автом).
9. Wikipedia. Багатошаровий_перцептрон. URL: https://uk.wikipedia.org/wiki/Багатошаровий_перцептрон.
10. Wikipedia. Рекурсивна_нейронна_мережа. URL: https://uk.wikipedia.org/wiki/Рекурсивна_нейронна_мережа.

11. Wikipedia. Рекурентна_нейронна_мережа. URL: https://uk.wikipedia.org/wiki/Рекурентна_нейронна_мережа.
12. Analytics Vidhya. Introduction to Seq2Seq Models. URL: <https://www.analyticsvidhya.com/blog/2020/08/a-simple-introduction-to-sequence-to-sequence-models/>.
13. Wikipedia. Довга короткочасна пам'ять. URL: https://uk.wikipedia.org/wiki/Довга_короткочасна_пам'ять.
14. Evergreens. Згортова мережа - просте пояснення CNN та її використання. URL: <https://evergreens.com.ua/ru/articles/cnn.html>.
15. В.І. Дигодій, Т.Б. Мартинюк, М.А. Очкуров. Нейромережева технологія виявлення і розпізнавання людей у відеопотоці. 2020. С. 2. URL: <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/41884/19561.pdf?sequence=3&isAllowed=y>.
16. О.І. Проніна, Р.В. Остапенко, С.В. Альошин. Розпізнавання жестів за допомогою згортової нейронної мережі. Вісник приазовського державного технічного університету. 2021. С. 7–14. URL: <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/41884/19561.pdf?sequence=3&isAllowed=y>.
17. Finger-talk. Fingertalk. URL: <https://finger-talk.com/>.
18. Terabee. Custom designed gesture recognition applications. URL: <https://www.terabee.com/custom-designed-gesture-recognition-applications/>.
19. Google Play. Gesture Control. URL: <https://play.google.com/store/apps/details?id=com.conena.navigation.gesture.control&hl=uk&gl=US>.
20. Nauka.ua. Штучний інтелект забезпечив високу точність керування біонічною рукою. URL: <https://nauka.ua/news/shtuchnij-intelekt-pidvishchiv-tochnist-keruvannya-bionichnoyu-rukoyu>.
21. Robotdreams. Функції активації: ступінчаста, лінійна, сигмоїдна, Relu, Tanh. URL: <https://robotdreams.cc/uk/blog/327-funkciji-aktivaciji-stupinchasta-liniyna-sigmojida-relu-ta-tanh>.

22. Wikipedia. Softmax. URL: <https://uk.wikipedia.org/wiki/Softmax>
23. GeeksforGeeks. Intuition of Adam Optimizer. URL: <https://www.geeksforgeeks.org/intuition-of-adam-optimizer/#:~:text=Adam%20optimizer%20involves%20a%20combination,minima%20in%20a%20faster%20pace.>
24. Wikipedia. Convolutional neural network. URL: https://en.wikipedia.org/wiki/Convolutional_neural_network
25. Medium. Image Classification With MobileNet. URL: <https://medium.com/analytics-vidhya/image-classification-with-mobilenet-cc6fbb2cd470>
26. Medium. MobileNet Architectures. URL: <https://medium.com/@pandrii000/mobilenet-architectures-17fe7406d794>
27. Stackoverflow. How does the flatten layer work in keras. URL: <https://stackoverflow.com/questions/44176982/how-does-the-flatten-layer-work-in-keras>
28. Semantic scholar. MobileNets: efficient convolutional neural networks for mobile vision applications. URL: <https://www.semanticscholar.org/paper/MobileNets%3A-Efficient-Convolutional-Neural-Networks-Howard-Zhu/3647d6d0f151dc05626449ee09cc7bce55be497e>
29. W3schoolsUA. Tensorflow.js Підручник. URL: https://w3schoolsua.github.io/ai/ai_tensorflow_intro.html#gsc.tab=0
30. Morning Dough. TensorFlow. URL: <https://www.morningdough.com/uk/ai-tools/tensorflow/>
31. А.Г. Богданова, О.В. Мазурець, О.В. Собко. Gesture recognition using a neural network in real time. Black Sea Science: Proceedings of the International Competition of Student Scientific Works. 2023. С. 556–566. URL: <https://elar.khmnu.edu.ua/items/be3c7ee1-6434-41d1-87f7-587df03db830>
32. Mazurets O., Zalutska O., Tyschenko O., Bohdanova A. An Approach to Using MobileNet CNN-model for Gesture Recognition. Proceedings of XXIII International Scientific and Practical Conference «Problems of Science and

Technology: the Search for Innovative Solutions». May 15-17, 2024. Munich, Germany. 2024. Pp. 59-64. URL: https://isu-conference.com/wp-content/uploads/2024/05/Problems_of_science_and_technology_the_search_for_innovative_solutions_May_15_17_2024_Munich_Germany.pdf

ДОДАТКИ

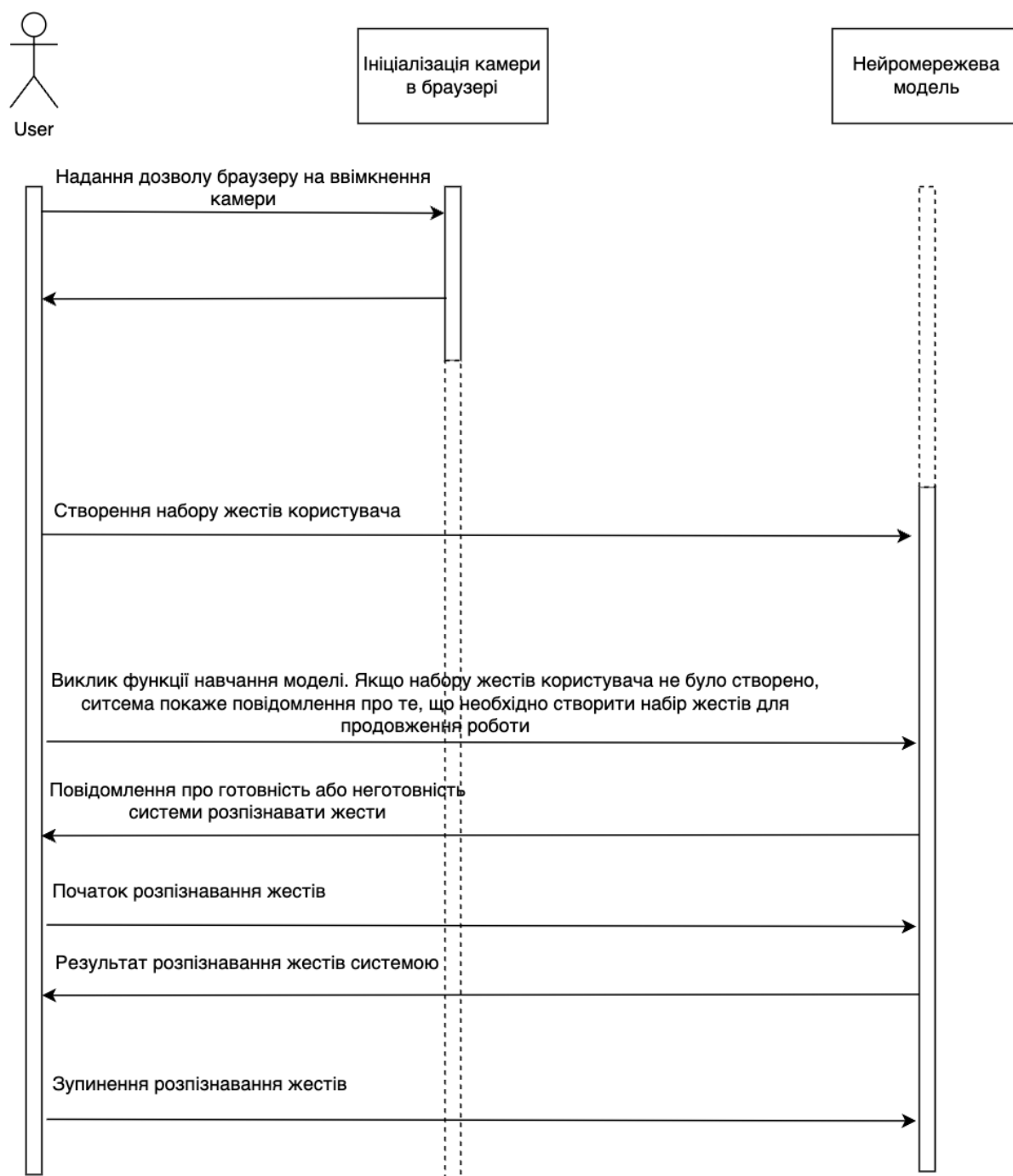
Додаток А

Схема методу нейромережевого розпізнавання жестів у відеопотоці



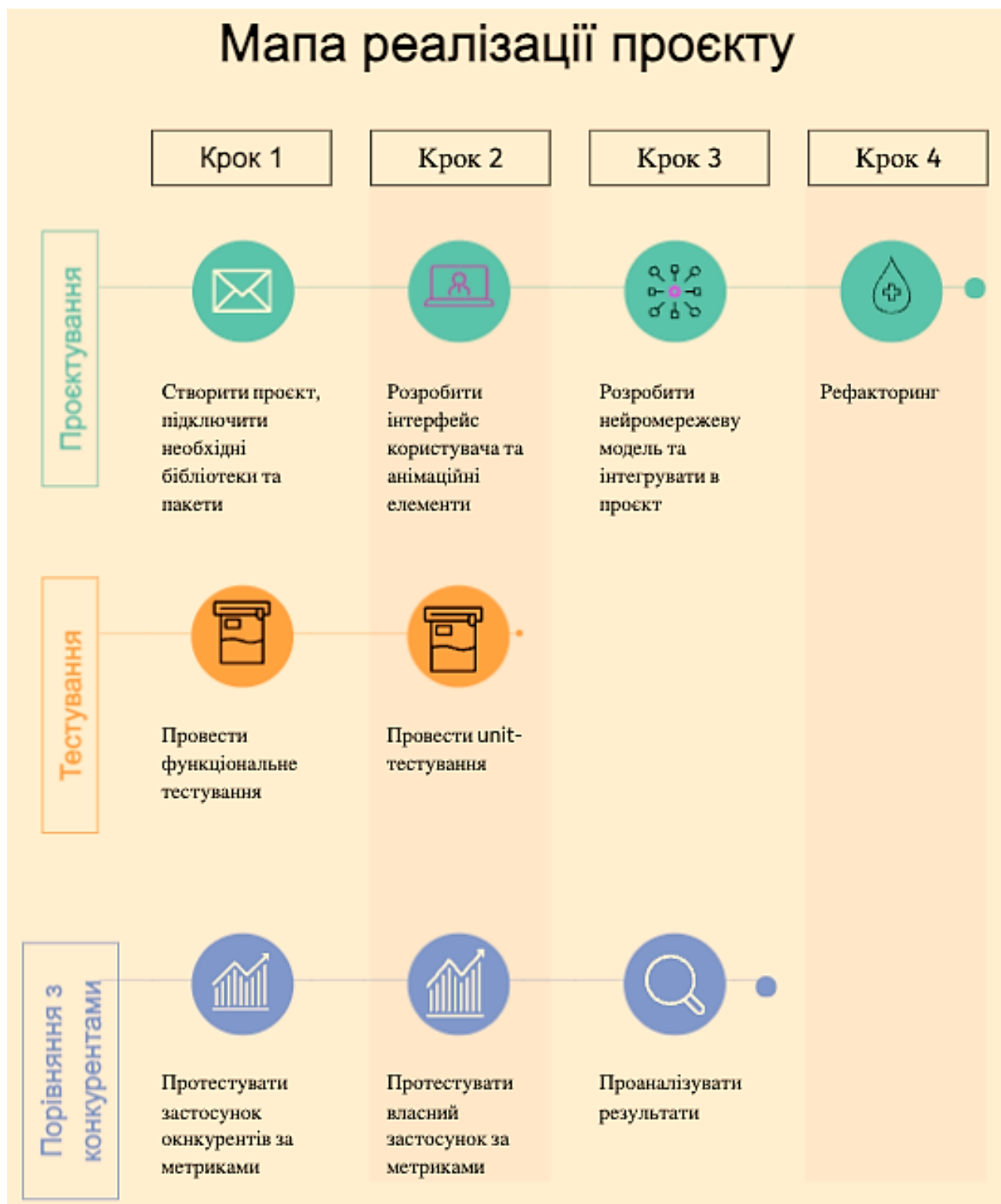
Додаток Б

Діаграма послідовностей системи нейромережевого розпізнавання жестів



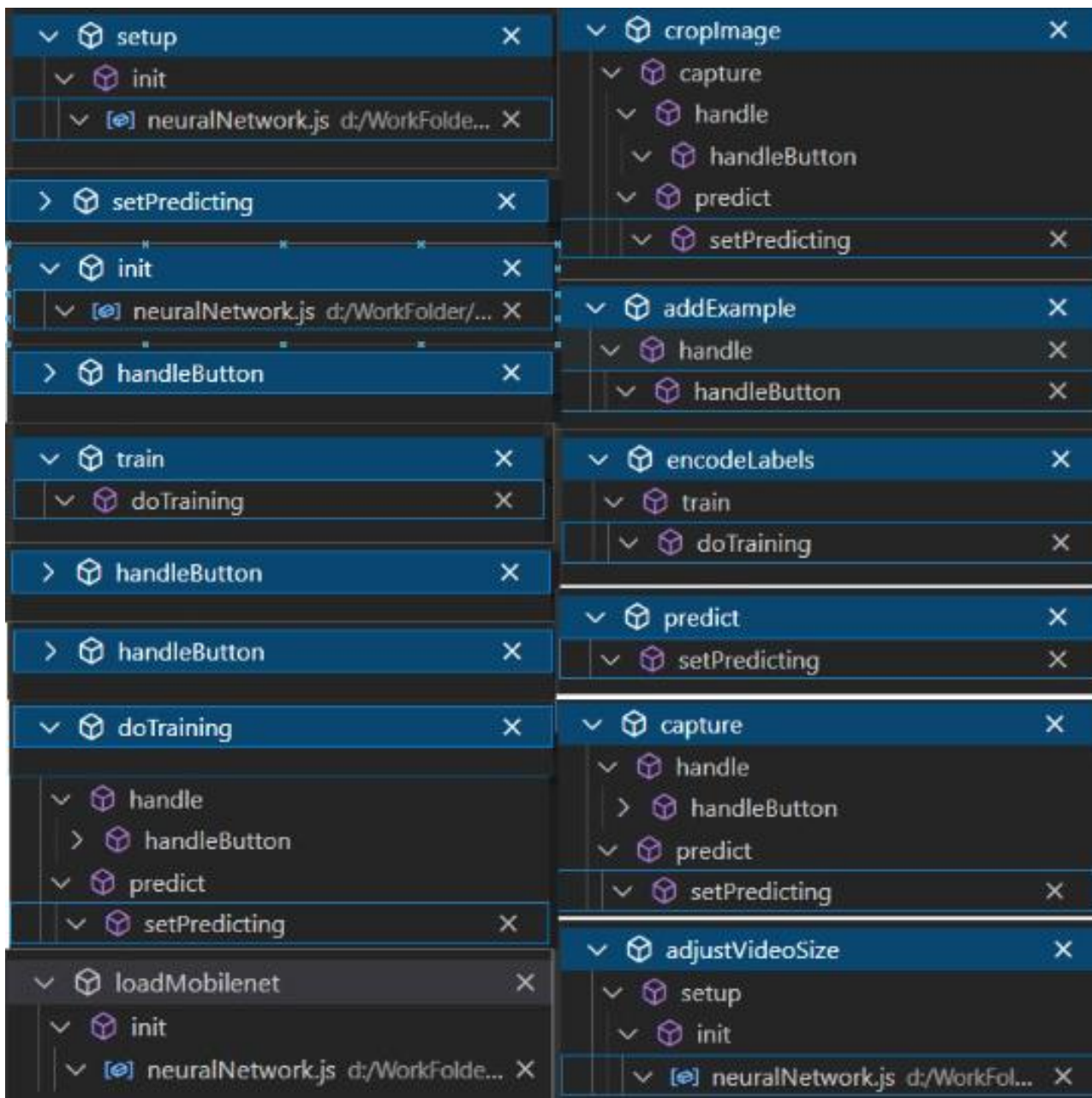
Додаток В

Дорожня карта реалізації системи нейромережевого розпізнавання жестів



Додаток Г

Діаграма класів системи нейромережевого розпізнавання жестів



Додаток Д

Презентаційний матеріал

КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА

МЕТОД НЕЙРОМЕРЕЖЕВОГО РОЗПІЗНАВАННЯ ЖЕСТІВ У ВІДЕОПОТОЦІ



Виконала:
Студентка 4 курсу, групи КН-20-1
Богданова Ангеліна Григорівна



Керівник:
к.т.н., доцент кафедри КН
Мазурець Олександр Вікторович

Актуальність

Дослідження мови жестів стали важливою сферою для розробки та імплементації програмного забезпечення в останні роки. Провідні науковці, розробники та аналітики створюють та презентують відповідні рішення на різноманітних виставках та конференціях з метою автоматизації розпізнавання жестів у відеопотоці.

Дані підходи були успішно імplementовані в різних галузях, зокрема таких, як кіберполіція, автомобільна промисловість, медицина, тощо. Технології орієнтовані на розпізнавання жестів користувачів допомагають зокрема людям з різноманітними хворобами, такими як хвороба Паркінсона, порушення мовлення або ж слуху, тунельний синдром, полегшенням взаємодії з комп'ютером, проведення ефективного навчання в школах та університетах і успішної соціалізації у суспільстві.

Окрім того, дослідження і розробки у даній сфері посилюють розвиток інформаційних технологій. Так, розробляються нові методи, архітектури, підходи та технології для вирішення поставлених завдань. Саме тому, дослідження даної проблеми є актуальним в наш час і необхідне для багатьох сфер життя і діяльності людини.

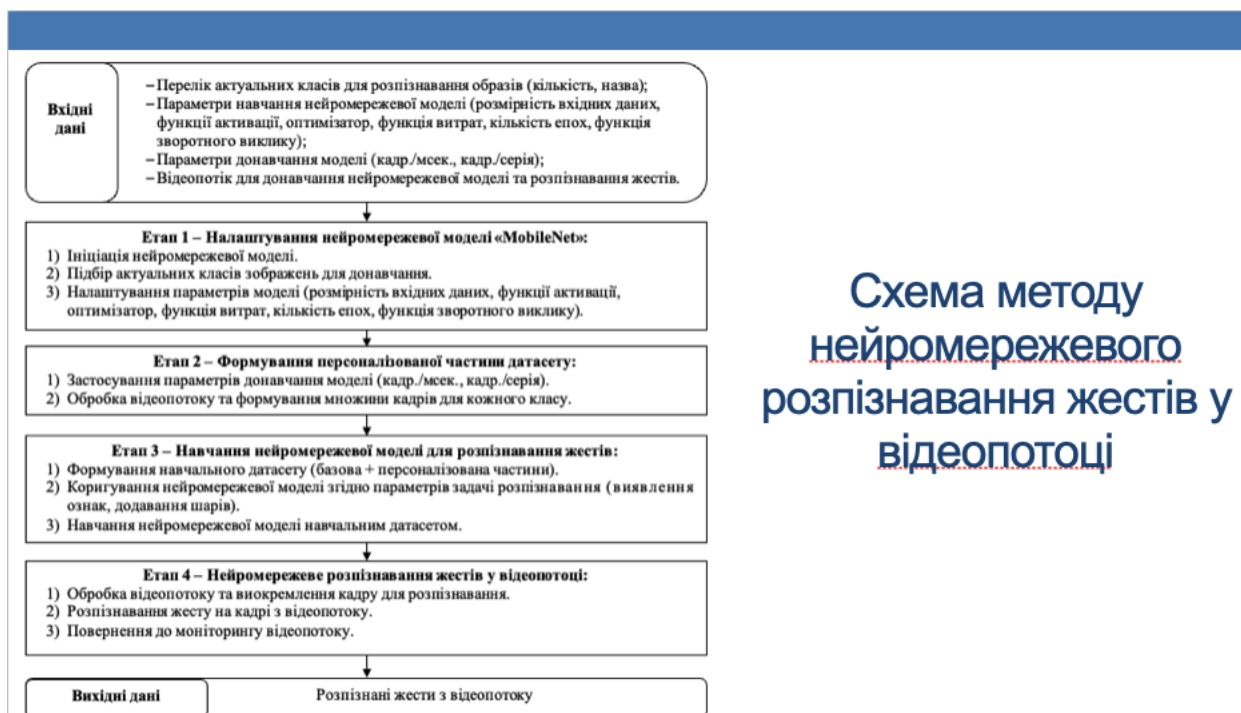
Мета і задачі роботи

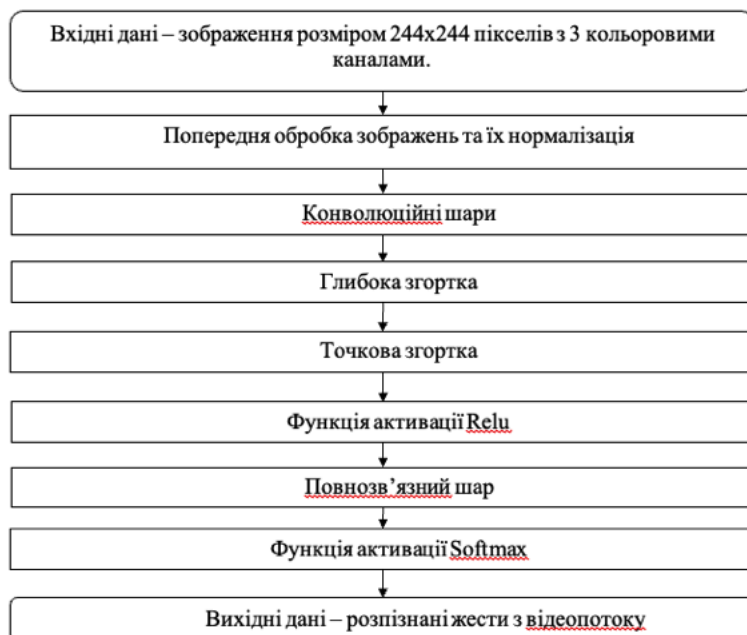
Метою кваліфікаційної роботи бакалавра є покращення розпізнавання жестів у відеопотоці, для чого потрібно розв'язати наступні задачі :

1. Виконати дослідження предметної області розпізнавання жестів у відеопотоці.
2. Виконати огляд теоретичних підходів щодо розпізнавання жестів у відеопотоці.
3. Виконати аналіз наукових надбань й існуючих програмних рішень в області розпізнавання жестів у відеопотоці.
4. Розробити метод нейромережевого розпізнавання жестів у відеопотоці.
5. Виконати проектування структури системи нейромережевого розпізнавання жестів у відеопотоці.
6. Здійснити вибір засобів розробки для системи нейромережевого розпізнавання жестів у відеопотоці.
7. Здійснити програмну реалізацію системи розпізнавання жестів
8. Провести тестування розробленої системи нейромережевого розпізнавання жестів у відеопотоці
9. Здійснити дослідження ефективності розробленого методу з використанням розробленої системи нейромережевого розпізнавання жестів

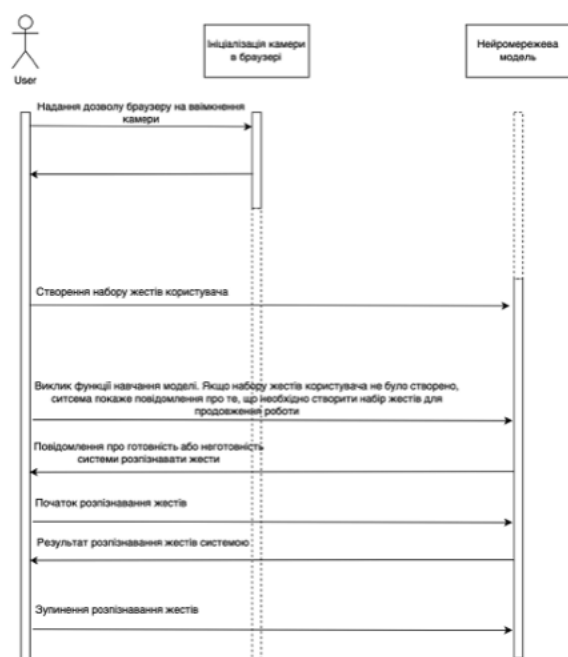
Для дослідження ефективності розробленого методу слід виконати його програмну реалізацію, що містить наступні основні функції:

- запитування дозволу браузера ввімкнути камеру на девайсі користувача;
- створення додаткового навчального датасету користувачем
- вибору кількості вхідних даних для подальшого тренування моделі користувачем
- запуску навчання моделі у реальному часі
- перегляду статусу навчання моделі у реальному часі
- початку процесу розпізнавання зображень
- зупинки розпізнавання зображень;
- демонстрація результату розпізнавання жестів у відеопотоці.



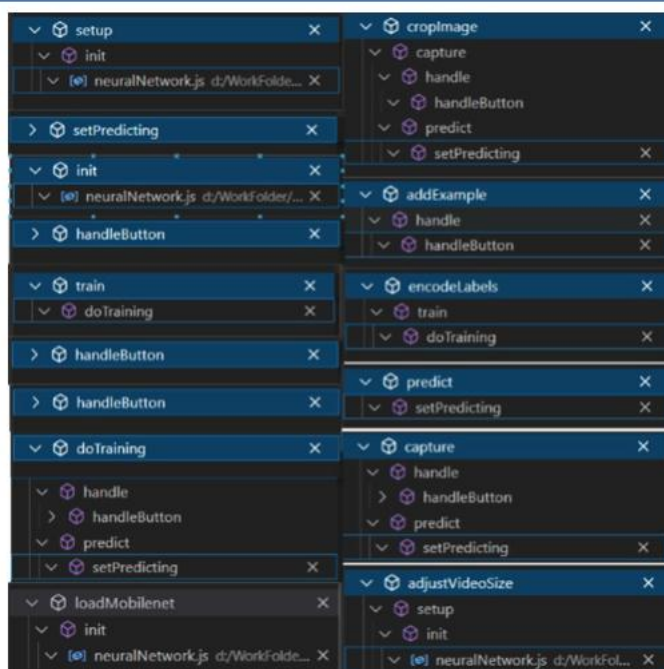
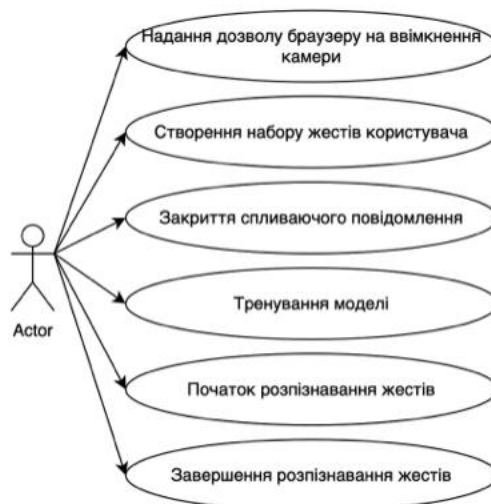


Блок-схема
основних етапів
моделі MobileNet



Діаграми послідовностей системи
та варіантів використання

Підхід до автоматизованого підбору відповіді на користувацькі запитання



Діаграма функцій системи нейромережевого розпізнавання жестів

Дослідження ефективності методу для розпізнавання жестів у відеопотоці

Матриця сплутувань, за параметрів кількість зображень = 100, освітлення - денне, відсутність сторонніх об'єктів на фоні.

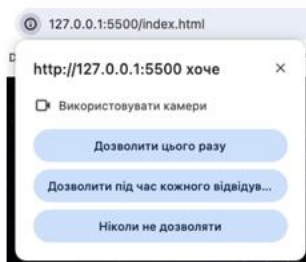
	0	1	2	3	4	5
0	20	0	0	0	0	0
1	1	19	0	0	0	0
2	0	0	20	0	0	0
3	0	0	0	20	0	0
4	0	1	0	1	18	0
5	0	0	0	1	0	19

Графік впливу параметрів нейромережевої моделі.

Найвищі показники моделі: Accuracy – 0.98, Precision – 0.99 та Recall – 1 при розмірі батчу – 64, кількості епох – 10, кількості зображень для навчання моделі – 100 при денному освітленні та відсутності сторонніх об'єктів на фоні зображення.



Програмна реалізація методу нейромережевого розпізнавання жестів у відеопотоці



Дозвіл браузеру на використання камери

Показ жесту для подальшого навчання моделі



Програмна реалізація методу нейромережевого розпізнавання жестів у відеопотоці



Результат розпізнавання жестів

Висновки

У рамках виконання кваліфікаційної роботи бакалавра було виконано **розробку методу** нейромережевого розпізнавання жестів у відеопотоці. Зокрема, було проведено аналіз предметної області й досліджено сучасні підходи до нейромережевого розпізнавання жестів у відеопотоці, розглянуто існуючі програмні реалізації за цим напрямком.

Розроблена **програмна реалізація** методу нейромережевого методу розпізнавання жестів у відеопотоці виконує наступні основні функції:

- запитування дозволу браузера ввімкнути камеру на девайсі користувача;
- створення додаткового навчального датасету користувачем
- вибору кількості вхідних даних для подальшого тренування моделі користувачем
- запуску навчання моделі у реальному часі
- перегляду статусу навчання моделі у реальному часі
- початку процесу розпізнавання зображень
- зупинки розпізнавання зображень;
- демонстрація результату розпізнавання жестів у відеопотоці.

У якості засобів розробки було обрано бібліотеку React, мову програмування Javascript, бібліотеку для машинного навчання Tensorflow.js, бібліотеку для створення анімації Three.js, бібліотеку Jest для тестування.

Ім'я користувача:
Кафедра КН

Дата перевірки:
17.06.2024 00:23:46 EEST

Дата звіту:
17.06.2024 00:26:49 EEST

ID перевірки:
1016366425

Тип перевірки:
Doc vs Internet + Library

ID користувача:
100005671

Назва документа: КН-20-1 Богданова_ЗАПИСКА

Кількість сторінок: 72 Кількість слів: 11384 Кількість символів: 92923 Розмір файлу: 2.59 MB ID файлу: 1016172697

Виявлено модифікації тексту (можуть впливати на відсоток схожості)

9.59%
Схожість

Найбільша схожість: 3.54% з джерелом з Бібліотеки (ID файлу: 1016172605)

7.03% Джерела з Інтернету

635

Сторінка 74

5.58% Джерела з Бібліотеки

177

Сторінка 79

0% Цитат

Вилучення цитат вимкнено

Вилучення списку бібліографічних посилань вимкнено

0%
Вилучень

Немає вилучених джерел

Модифікації

Виявлено модифікації тексту. Детальна інформація доступна в онлайн-звіті.

Замінені символи

7

Підозріле форматування

11
сторінок

Anti-Plagiarism v-15.257

Максимальне співпадіння з одним документом 7.0%

Словники перевірки: en_US, ru_RU, ua_UA. Помилки в документах: 15%

ID: 130867 Назва: КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА на тему Метод нейромережевого розпізнавання жестів у відеопотоці Додано в БД: 2024-06-16 Автора: Ангеліна БОГДАНОВА Керівники: Олександр МАЗУРЕЦЬ Консультанти: Опоненти:	Документ		Сумарний збіг по Базі Даних	
	Символи	Лексеми	Символи	Лексеми
	75437	1069	8254 (11%)	124 (12%)

Джерело плагіату

ID	Опис	Наявність плагіату в документі	
		Символи	Лексеми

**РІШЕННЯ ЕКСПЕРТНОЇ КОМІСІЇ КАФЕДРИ КОМП'ЮТЕРНИХ НАУК
ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ**

Підтверджуємо ознайомлення з результатом звіту подібності щодо роботи, генерованого системою виявлення текстових збігів/ідентичності/схожості:

Назва: Метод нейромережевого розпізнавання жестів у відеопотоці

Автор: студентка групи КН-20-1 Ангеліна Богданова

Спеціальність: 122 – Комп'ютерні науки

Освітня програма: освітньо-професійна

Науковий керівник: к.т.н., доц. каф. КН Олександр МАЗУРЕЦЬ

Після аналізу звіту подібності зроблено такий висновок:

№	Висновок	Позначка про відповідність
1	Запозичення, виявлені в роботі, є законними і не є плагіатом. Робота приймається до захисту.	<i>відповідає</i>
2	Виявлені запозичення не є плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота приймається до захисту, але має бути відкоригована. Відкоригований варіант має бути поданий на кафедру за 2 дні до захисту, разом із заявою щодо самостійності виконання письмової роботи та ідентичності друкованої та електронної версії роботи	
3	Виявлені запозичення не є плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. В зв'язку з цим мета роботи та поставлені завдання не були досягнені. Робота може бути допущена до захисту (наступного року) після того як буде відкоригована та допрацьована і успішно пройде повторну перевірку на академічний плагіат.	
4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	

Підтвердження:

Запозичення, виявлені в роботі Ангеліни Богданової, 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 програмного коду, що не мають авторства і містять поширені конструкції; серед запозичень знаходяться загальновідомі терміни, скорочення.

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

- за системою Anti-Plagiarism: 7%;

- за системою Unicheck: 9.59.

Керівник роботи

Олександр МАЗУРЕЦЬ

Гарант ОП

Олександр МАЗУРЕЦЬ

Завідувач кафедри КН

Олександр БАРМАК



ВІДГУК НАУКОВОГО КЕРІВНИКА на кваліфікаційну роботу бакалавра

студента гр. КН-20-1 Богданової Ангеліни Григорівни

за темою Метод нейромережевого розпізнавання жестів у відеопотоці

1. Актуальність теми

Розробка методу нейромережевого розпізнавання жестів у відеопотоці є актуальною завдяки своїм потенціалом у вдосконаленні систем комунікації, особливо для людей з обмеженими можливостями, що сприяє автоматизації та поліпшенню сприйняття жестів, що важливо для медицини, розважальної індустрії та інших галузей, де важлива точність і швидкість взаємодії з технологіями.

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

За стандартом, а саме описом предметної області, об'єктом дослідження є процес нейромережевого розпізнавання жестів у відеопотоці. Метою роботи є покращення розпізнавання жестів у відеопотоці. При вирішенні поставленої задачі використано методи та технології машинного навчання для розпізнавання об'єктів у відеопотоці. Отже, результати кваліфікаційної роботи бакалавра повністю відповідають стандарту бакалавра спеціальності 122 – Комп'ютерні науки.

3. Професійні та особистісні якості бакалавра

При виконанні кваліфікаційною роботою бакалавра Богданова Ангеліна Григорівна проявила себе як дисциплінована та самостійна студентка. Також показала достатні вміння та навички в розробці програмного застосунку заданої теми, що дозволило отримати відмінні результати.

4. Ступінь самостійності під час виконання кваліфікаційної роботи

Студентка самостійно виконувала усі завдання під час виконання кваліфікаційної роботи, тому результати роботи є її особистим надбанням.

5. Ступінь оволодіння методами дослідження

Під час роботи над кваліфікаційною роботою Богданова Ангеліна Григорівна показала високий рівень володіння методами дослідження області комп'ютерних наук. За темою кваліфікаційної роботи в здобувача є суттєві наукові результати, виконано публікації.

6. Повнота та якість розкриття теми роботи

Тема роботи є добре обґрунтованою і розкритою. У роботі проведено аналіз актуальності теми та сучасних досліджень та існуючих програмних рішень в цій області, а також підходів до розв'язку подібних задач. Усі поставлені завдання були успішно виконані, в тому числі розроблено інформаційну систему для оцінки ефективності запропонованого методу нейромережевого розпізнавання жестів у відеопотоці.

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

Кваліфікаційна робота бакалавра відзначається чіткою структурою, що сприяє послідовному та логічному викладенню матеріалу з повним аргументуванням. Також слід відзначити високий рівень літературної грамотності автора.

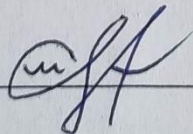
8. Можливість практичного застосування кваліфікаційної роботи бакалавра, окремих її частин

Розроблена система для розпізнавання жестів користувача у відеопотоці має потенціал для впровадження в різні сфери. Зокрема, вона може бути корисною у сферах охорони здоров'я, розважальній індустрії та автомобільній галузі. Головним напрямком практичного застосування є автоматизація розпізнавання жестів для поліпшення комунікації користувачів без використання словесних засобів.

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

Враховуючи високий рівень виконання та забезпечення усіх необхідних вимог, робота може бути допущена до захисту. Рекомендована оцінка «відмінно».

Керівник _____



к.т.н., доцент кафедри КН Олександр МАЗУРЕЦЬ



РЕЦЕНЗІЯ

на кваліфікаційну роботу бакалавра

студентки гр. КН-20-1 Богданової Ангеліни Григорівни

за темою: Метод нейромережевого розпізнавання жестів у відеопотоці

1. Актуальність обраної теми

Дослідження мови жестів набуває важливості у сфері розробки програмного забезпечення, зокрема для автоматизації розпізнавання жестів у відеопотоках. Розробка таких технологій сприяє спрощенню комунікації для людей з порушеннями слуху та знаходять застосування в різних галузях, включаючи медицину та кібербезпеку.

2. Повнота розкриття мети та завдань роботи

Авторка кваліфікаційної роботи бакалавра повністю розкрила мету та завдання. В роботі проаналізовано сучасний стан предметної області, виконано огляд існуючих рішень та підходів, що дало змогу розробити метод нейромережевого розпізнавання жестів у відеопотоці і детально його описати, а також провести відповідне дослідження ефективності.

3. Зміст кожного розділу роботи

Усі розділи роботи мають актуальну інформацію, що стосуються теми. У першому розділі надано характеристику предметної області розпізнавання жестів, а саме виконано аналіз моделей, методів та реалізацій, що дало змогу постави мету та задачі роботи. У другому розділі описано проектування системи нейромережевого розпізнавання жестів у відеопотоці, включаючи метод нейромережевого розпізнавання жестів у відеопотоці. У третьому розділі проведено експериментальне дослідження методу нейромережевого розпізнавання жестів у відеопотоці.

4. Оцінка розробленої інформаційної системи, її практична цінність

Розроблена інформаційна системи нейромережевого розпізнавання жестів у відеопотоці має високий потенціал застосування. Автоматизоване розпізнавання жестів у відеопотоках може знайти застосування у медицині, розважальній індустрії та автомобільному секторі.

5. Якість оформлення кваліфікаційної роботи бакалавра

Авторка якісно оформила кваліфікаційну роботу бакалавра. Робота містить необхідні розділи, ілюстративні матеріали, такі як таблиці та графіки, що допомагають краще розуміти представлену інформацію. Використана література підтверджує обґрунтованість висновків, а послідовне викладення матеріалу забезпечує легке сприйняття тексту.

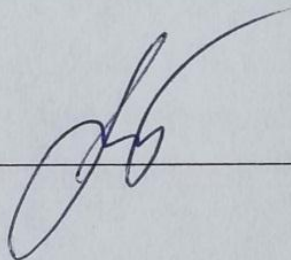
6. Недоліки кваліфікаційної роботи бакалавра

Суттєвих недоліків немає. Було б доречним вивести на інтерфейсну частину розробленої інформаційної системи у відсотковому відображенні оцінку приналежності образу до розпізнаного класу. По тексту пояснювальної записки виявлено кілька пунктуаційних помилок, переважно у 3 розділі. Втім наведене не впливає на загальну якість роботи та отримані результати.

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

Враховуючи рівень виконання та забезпечення усіх необхідних вимог, робота може бути допущена до захисту. Рекомендована оцінка «відмінно».

Рецензент _____



Говорушченко Т.О.