

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Факультет інформаційних технологій

Кафедра технологій управління

Спеціальність 122 – Комп’ютерні науки,
освітня програма «Інформаційна аналітика та впливи»

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

на тему:

“Прогнозування фільмографічних уподобань користувача по
аналізу емоцій з селфі фото”

**Студента 2-го курсу групи
ІАВ-21**

Олександр ШЕЛЕЯ

(прізвище, ім’я, по батькові)

(підпис студента)

Науковий керівник:

к.т.н., асистент

(науковий ступінь, вчене
звання)

Андрій ХЛЕВНИЙ

(прізвище, ім’я, по батькові)

(дата)

(підпис)

Попередній захист:

(Висновок: «До захисту в Екзаменаційній комісії»)

Завідувач
кафедри технологій
управління

Віктор МОРОЗОВ

(підпис)

(прізвище, ініціали)

(дата)

Київ – 2022

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
Факультет інформаційних технологій**

Кафедра технологій управління
Освітньо-кваліфікаційний рівень Магістр
Спеціальність 122 - Комп'ютерні науки
Освітня програма Інформаційна аналітика та впливи

ЗАТВЕРДЖУЮ
Завідувач кафедри
професор Морозов В.В.

«___» _____ 20__ року

**ЗАВДАННЯ
НА ВИКОНАННЯ КВАЛІФІКАЦІЙНОЇ РОБОТИ**

Студент *Олександр ШЕЛЕЯ*

Група ІАВ-21

1. Тема кваліфікаційної роботи “Прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото”

Затверджена наказом по від «17» листопада 2021 р. № 4.

2. Строк подання студентом готової роботи – “18” травня 2022 р.

3. Цільова установка та вихідні дані до роботи

Дослідження особливостей розпізнавання емоцій з фото та застосування методів штучного інтелекту в сфері прогнозування фільмографічних уподобань.

4. Зміст роботи

Аналіз основних методів визначення емоцій, дослідження застосування методів штучного інтелекту, побудова архітектури та програмного рішення для прогнозування фільмографічних уподобань, та отримання з них інформації для подальшого використання для мобільного додатку.

5. Перелік графічного матеріалу (слайдів)

45 рисунків, 2 додатки, 18 слайдів презентації доповіді.

6. Календарний план виконання роботи:

№ з/п	Назва частин роботи	%	Виконання роботи	
			За планом	Фактично

1	Вибір теми дипломної роботи	3	01.10.21	01.10.21
2	Протокол кафедри ТУ про затвердження тем дипломних робіт та призначення наукових керівників	2	24.12.21	24.12.21
3	Формування переліку нормативних матеріалів, літератури з проблематики дипломної роботи	10	07.01.22	07.01.22
4	Складання розгорнутого плану кваліфікаційної роботи	5	18.01.22	18.01.22
5	Ознайомлення наукового керівника з розгорнутим планом кваліфікаційної роботи. Внесення змін.	5	19.01.22 - 20.01.22	19.01.22 - 20.01.22
6	Написання розділу 1 дипломної роботи	15	14.02.22	14.02.22
7	Написання розділу 2 дипломної роботи	17	08.03.22	08.03.22
8	Написання розділу 3 дипломної роботи	19	01.04.22	01.04.22
9	Оформлення кваліфікаційної роботи. Підготовка висновків і пропозицій	15	03.05.22	03.05.22
10	Передача кваліфікаційної роботи науковому керівникові	2	04.05.22	04.05.22
11	Передача кваліфікаційної роботи рецензенту для рецензування	2	11.05.22	11.05.22
12	Попередній захист кваліфікаційної роботи	5	17.05.22	17.05.22

Дата видачі завдання «17» листопада 2022 р.

Керівник роботи к.т.н., ас. Андрій ХЛЕВНИЙ

(підпис)

Завдання прийняв до виконання:

Здобувач освіти групи ІАВ-21 Олександр ШЕЛЕЯ



(підпис)

ЗМІСТ

АНОТАЦІЯ	6
ВСТУП	7
РОЗДІЛ 1. ПОСТАНОВКА ЗАДАЧІ ТА АНАЛІЗ РІШЕННЯ	10
1.1 Постановка задачі магістерської роботи	10
1.2 Аналіз існуючих у світі способів вирішення задачі	19
1.3 Визначення вхідних даних для вирішення поставленої задачі	31
1.4 Висновки до першого розділу	33
РОЗДІЛ 2. УПРАВЛІННЯ ПРОЕКТОМ ОБРОБКИ ІНФОРМАЦІЇ	34
2.1 Методи обробки інформації	34
2.2 Розробка проекту проведення аналітичної діяльності	44
2.2.1 Вибір програмного середовища	44
2.2.2 Система управління базами даних MySQL	52
2.2.3 Створення користувацьких інтерфейсів	54
2.3 Засоби обробки інформації	58
2.4 Висновки до другого розділу	71
РОЗДІЛ 3. РЕАЛІЗАЦІЯ МЕТОДІВ ОБРОБКИ ІНФОРМАЦІЇ НА ОСНОВІ РЕАЛЬНИХ ДАНИХ.	72
3.1 Візуалізація вхідних даних проекту обробки інформації	72
3.1.1 Ідентифікація архетипу ІС	72
3.1.2 Інтерфейс користувача (UI View)	72
3.1.3 Логічне уявлення ІС (Logical View)	74
3.1.4 Архітектура БД	75
3.1.5 Опис стеку технологій	76
3.2 Представлення поетапної обробки даних засобами обробки інформації	78
3.3 Візуалізація результату проекту	85
3.4 Висновки до третього розділу	95

ВИСНОВКИ	96
ПЕРЕЛІК ВИКОРИСТАНИХ ІНФОРМАЦІЙНИХ ДЖЕРЕЛ	97
Додаток А	103
Додаток Б	106

АНОТАЦІЯ

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет інформаційних технологій

Кафедра технологій управління

Спеціальність 122 - Комп'ютерні науки,
освітня програма "Інформаційна аналітика та впливи"

Дипломна робота магістра Шелеї Олександра Івановича.

Тема роботи – «Прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото».

Мета дипломної роботи магістра – розробити додаток для прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото.

Об'єкт дослідження – прогнозування фільмографічних уподобань на основі емоцій.

Предмет дослідження – методики застосування аналізу емоцій у задачах прогнозування фільмографічних уподобань.

Наукова новизна роботи – розроблено перший у світі додаток, що дозволяє прогнозувати фільмографічних уподобання на основі емоцій з селфі фото.

У роботі досліджуються існуючі підходи використання штучного інтелекту. Розробляється нова методика його використання, а також проводиться обґрунтування доцільності впровадження рішення.

Дипломна робота складається зі вступу, основної частини, яка включає три розділи, висновків, 2 додатки та списку використаних джерел. Всього налічує 107 сторінок та перелік посилань з 51 джерела на 6 сторінках.

Ключові слова: мобільний додаток, проект, інтелектуальний аналіз даних, емоції, фільми.

ВСТУП

Науково-технічний розвиток суспільства призвело до широкого використання людьми різних аудіовізуальних засобів зв'язку (засоби масової інформації, телефони, супутниковий космічний зв'язок). Різке підвищення їхньої ролі в соціально-політичному житті зумовило інтерес вчених до досліджень особливостей опосередкованого та обмеженого спілкування, вивчення впливу різних факторів на емоції.

Програми для смартфонів, або так звані мобільні додатки, стали невід'ємною частиною сучасного життя. З розвитком технологій програмування збільшуються і можливості в розробці мобільних додатків.

Мобільний додаток являє собою програмний код (програму) для смартфона або планшета. Програма встановлюється на ту чи іншу платформу і має певний функціонал для виконання дій, спрямованих на вирішення поставленого кола питань – програми для бізнесу, розваг, навчання, органайзери з опціями нагадування, розважальний контент, різні сервісні служби з замовлення та доставки товарів, авіаквитків, замовлення таксі тощо.

В світі програмістами щодня створюються програмні засоби, які роблять взаємодію з мобільним пристроєм більш комфортною, або спрощують процеси спілкування, допомагають організувати роботу або відпустку, дізнатися прогноз погоди або новини в світі, та багато іншого. Врешті-решт, логічно, що сучасний мобільний пристрій має мати хоча б мінімальний набір встановлених додатків.

Ретельно продумавши функціонал програми та стратегію монетизації, можна перетворити мобільний додаток в потужний інструмент для бізнесу.

Розпізнавання емоцій на обличчі перетворюється на надзвичайно важливе значення у різних методах комунікації для зміни емоцій у суспільстві, яке можна оцінити як інтерпретацію дискурсу. Хоча про це говорили різні вчені у галузі комп'ютерного зору. Більшість із цих прикладів дослідження з ентузіазмом ставляться до візуального чи віртуально розпізнавання емоцій. За

допомогою різних центрованих абстрактних експертиз вченими виявлено, що на вигляд люди розрізняють гнів, щастя, подив і ворожість.

У роботі розглядається питання можливості використання підходу розпізнавання емоцій обличчя для розробки програмного продукту.

Отже курсова робота присвячено актуальній на сьогоднішній день задачі аналізу емоцій з фото та використання результатів цього аналізу у програмному продукті.

Метою дослідження є програмна реалізація алгоритму прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото.

Для досягнення мети були визначені наступні задачі:

1. проаналізувати існуючі аналогічні системи та додатки, виділити основні риси, переваги та недоліки;
2. описати особливості функціоналу для майбутнього додатку;
3. вибрати інструменти для розробки і проектування;
4. спроектувати дизайн, інтерфейс та макет додатку згідно визначених особливостей;
5. реалізувати функціонал додатку з використанням сучасних методів та засобів розробки;
6. протестувати функціонал додатку.

Об'єкт дослідження – прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото.

Предмет дослідження – це алгоритм розпізнавання емоцій обличчя.

Для вирішення поставлених завдань використовувалися такі методи дослідження:

- вивчення літератури та результатів існуючих рішень;
- порівняння;
- аналіз;
- тестування.

Наукова новизна одержаних результатів – вперше одержано програмний продукт, здатний на основі аналізу емоцій з фото, сформувати перелік фільмографічних вподобань.

Практичне значення одержаних результатів полягає у вивченні впливу різних факторів на успішність розпізнавання емоцій за виразом обличчя. На підставі даного дослідження планується подальша розробка програми для підвищення точності розпізнавання емоції обличчя.

РОЗДІЛ 1. ПОСТАНОВКА ЗАДАЧІ ТА АНАЛІЗ РІШЕННЯ

1.1 Постановка задачі магістерської роботи

Буквально 15-20 років тому на питання, що таке мобільний додаток, власник мобільного телефону не знайшов би відповіді. Можливості пристроїв обмежувалися функціями дзвінків, відправлення СМС і найпростішими іграми. Втім, невігадливу гру на мобільному можна було вже тоді називати додатком.

Сьогодні мобільний додаток – це спеціально розроблене під функціональні можливості гаджетів програмне забезпечення. Призначення ПО може бути найрізноманітнішим: сервіси, магазини, розваги, онлайн-помічники і інше. Ці додатки завантажуються і встановлюються самим користувачем через мобільні маркетплейси. Найбільші майданчики – AppStore, Google Play. Технічно всі програми створюються під конкретну платформу мобільного гаджета. Найбільш популярні операційні системи - iOS, Android, Windows Phone.

Мобільний додаток – це програмний пакет, функціонал і дизайн якого «заточений» під можливості мобільних платформ. Перерахуємо кілька основних плюсів програми [1]:

- інтерфейс програми створений саме під роботу на мобільному пристрої через сенсорний екран або кнопки;
- зручна і зрозуміла для користувачів гаджетів навігація, мобільне меню;
- краща взаємодія з користувачем через повідомлення, пуш-повідомлення, нагадування. Додаток може виконувати функції навіть у фоновому режимі, чого не можна сказати про сайт. Для роботи з програмою не потрібно відкривати браузер, а багато додатків підтримують ряд функцій і при відключеному інтернеті;
- зберігання персональних даних користувача. Ця функція розширює можливості персоналізації додатків. Наприклад, викликає таксі додому (прописка), записує на прийом до лікаря з медичного полісу та інші переваги;

- більш гнучкий зворотний зв'язок з компанією, сервісом;
- можна задіяти більше ресурсів. Наприклад, підключити геолокацію і викликати машину в будь-яку точку міста;
- програми можуть враховувати біологічні ритми людини і сповіщати його про необхідність дотримуватися режиму.

Насправді функціонал мобільних додатків вже давно перевершив адаптовані сайти. Сьогодні можна скачати і встановити на смартфон програми для бізнесу, навчання, органайзери з опціями нагадування, розважальний контент, різні сервісні служби.

Програми для смартфонів стали частиною життя сучасних людей. Різні типи мобільних додатків допомагають організувати роботу і відпочинок, дізнатися останні новини, спілкуватися з друзями. Щодня з'являються нові програми для iOS, Android, Java, Symbian, Windows, інших операційних систем.

Під час карантину попит на нові розважальні послуги різко зріс. Із закриттям кінотеатрів по всьому світу єдиним варіантом стали екрани телевізорів і комп'ютерів, а на допомогу прийшли всілякі стрімінгові платформи, які стали транслювати довгоочікувані прем'єри.

Розважальні сервіси дозволяють захопити широку аудиторію завдяки широкому виборі контенту і мультиплатформенності. Ретельно продумавши функціонал програми та стратегію монетизації, можна перетворити мобільний додаток в потужний інструмент для бізнесу.

7 травня 2020 р. Sandvine Incorporated – компанія, що займається розробкою додатків та мережевим аналізом, а також продуктів для контролю мережевих політик, включаючи цензуру в Інтернеті, управління навантаженнями та безпеку – оголосила про випуск спеціального звіту про глобальні явища в Інтернеті під час пандемії COVID-19 «Special COVID-19 Global Internet Phenomena Report» [2]. Цей звіт аналізує, як трафік на споживчих широкосмугових мережах зазнав значних змін, оскільки споживачі в основному використовують свої стаціонарні домашні мережі для доступу до Інтернету. Це видання Звіту про явища надає унікальний погляд на Інтернет-трафік.

Всесвітній лозунг «залишатися вдома» сконцентрував переважну більшість інтернет-трафіку на одній мережі, даючи відповіді на питання, які програми насправді сприяють задоволенню споживачів і де операторам мереж слід зосередити свої зусилля.

Коментуючи звіт, Лін Кантор, президент і генеральний директор Sandvine, сказала: «Світова пандемія COVID-19 різко вплинула на використання мережі в усьому світі, і багато операторів відчують 100% Зростання обсягу трафіку в міру переходу використання від підприємств та освітніх мереж до переважно споживчих широкосмугових мереж. Мета Sandvine – забезпечити, щоб оператори мережі могли надалі надавати високоякісні Інтернет-послуги, від яких залежать споживачі в ці часи, і цей звіт є нашим внеском у мережеву галузь у ці періоди змін».

«Special COVID-19 Global Internet Phenomena Report» – це авторитетний погляд на те, як програми споживають світову пропускну здатність у світі. Завдяки встановленій базі понад 2,5 млрд. абонентів у всьому світі в більш ніж 500 операторах фіксованого, мобільного, Wi-Fi та супутникового зв'язку, бачення Sandvine тенденцій в Інтернеті не має собі рівних в галузі. Це видання звіту повідомляє про загальний обсяг трафіку для тисяч різних додатків, постачальників контенту та інтернет-протоколів, які є основною інтелектуальною власністю мережевих розвідницьких рішень Sandvine.

У звіті висвітлено зміни у змішаному мережевому трафіку під час пандемії COVID-19, що допомагає мережевим операторам зрозуміти, чому змінюється структура трафіку та як вони можуть адаптувати свої послуги до нової норми. Мережеві оператори, які намагаються впоратись із цими змінами, яким недостатньо чітко видно, як ці зміни вплинуть на якість своїх абонентів, отримують величезне розуміння того, що може бути новим нормальним явищем у їхніх мережах в майбутньому.

Деякі основні моменти цього видання включають:

- трафік у глобальному масштабі з 1 лютого по 19 квітня зріс майже на 40%;

- YouTube – безперечний король інтернет-трафіку, понад 15% всього трафіку в споживчих широкосмугових мережах;
- Netflix складає понад 11% від загального обсягу трафіку, на другому місці після категорії YouTube;
- Відео, ігри та соціальний доступ становлять понад 80% всього інтернет-трафіку, підкреслюючи їх значення для споживачів як рятувальний круг для розваг під час відключення.

Звіт висвітлює лідерів частки трафіку для відео, соціальних мереж, обміну повідомленнями та ігрових програм під час пандемії COVID-19 (рисунок 1.1).



Рисунок 1.1 Глобальний звіт Sandvine Incorporated про зміни в Інтернеті під час Пандемії COVID-19 [2]

Ще раніше, звіт Sandvine Incorporated «The Global Internet Phenomena Report» [3] розкрив веб-сервіси, що відповідають за найбільший у світі трафік (рисунок 1.2).

Згідно з дослідженнями Sandvine Incorporated у [3] передача відео займає 60% всього інтернет-трафіку (рисунок 1.2), велика частина якого – онлайн-стрімінг новин, телешоу, спорту, серіалів, фільмів і відеоігор на Netflix, YouTube, Twitch і інших майданчиках.



Рисунок 1.2. Частка трафіку сервісів у глобальній мережі [3]

Створення якісних сервісів вигідно для всіх учасників ринку. Користувачі отримують інструменти для вирішення повсякденних завдань. Розробники отримують прибуток за рахунок [4]:

- продажі рекламних блоків;
- платних додаткових функцій (в безкоштовній програмі);
- платних поширення програм;
- створення контенту на замовлення.

Від затребуваності контенту напряму залежить сума прибутку. Для того, щоб мобільний додаток став популярним, він як мінімум, повинен бути зручним для користувача і неодмінно корисним. Не останнє місце при цьому займає стиль та дизайн оформлення.

Якщо додаток розробляється для бізнесу, то запуск сервісу в такому випадку надає наступні переваги: розширення цільової аудиторії, зворотний зв'язок від клієнтів, збільшення прямих продажів, зростання впізнаваності бренду.

В залежності від призначення класифікація мобільних додатків для всіх смартфонів відбувається за наступною схемою: всіх їх можна розділити на 4 категорії [4]:

- Ігри для смартфонів;
- Промо-додатки;
- Контентні сервіси;
- Соціальні мережі.

Безумовно, інші типи додатків також існують, наприклад, сервіси, створені для управління компаніями або для ведення бізнесу. Але такі додатки менш затребувані, ніж чотири вищеперераховані категорії. Як показано у [4], за статистикою – 70% смартфонів мають серед своїх додатків хоча б одну програму з перерахованих категорій.

Коротко оглянемо кожен з категорій.

У групу ігрових додатків об'єднані мобільні ігри всіх жанрів: від стрілянини до логічних завдань, включаючи, гонки, аркади, стратегії, квести тощо. Ця група включає сімейні, дорослі та дитячі ігри. Якщо говорити про цільову аудиторію ігрових додатків – це молоді люди до 27 років. Однак популярність мобільних ігор у людей старшого віку збільшується з кожним роком. Найчастіше мобільні ігри надають безкоштовний доступ користувачеві, проте є можливість покупки додаткового контенту, що є найпоширенішим способом заробітку для розробників.

Промо-додатки – це маркетинговий інструмент просування бренду, тому створюються на замовлення конкретного бізнесу чи торгівельної марки. Через такі мобільні сервіси користувачі можуть оформляти доставку товарів чи замовляти послуги, приймати участь у програмі лояльності, накопичувати бонуси, бути в курсі акцій, новинок та знижок або розпродажів, а також залишати відгуки. Саме тому головною метою розробників промо-додатків – це максимальне охоплення клієнтів, незалежно від того, якими моделями смартфонів вони користуються. Найпопулярніші сервіси цієї категорії – це доставка їжі, виклик таксі та купівля квитків в кінотеатри.

Контентні додатки – це програми, розроблені для можливості швидкого доступу до певного контенту. Кілька прикладів інформації, яку вони можуть містити:

- Публікації новин;
- Мотиваційні цитати;
- Актуальні курси валют;
- Погода;
- Курси валют;
- Гороскопи;
- Системи схуднення або тренувань;
- Навчальні мовні курси;
- Обробка фото/відео;
- Книги;
- Рецепти або кулінарні гайди;
- Правильне харчування;
- Карти і створення маршрутів.

Як видно, це велика група програм, яка включає багато видів додатків на різну тематику.

Для заробітку на тематичних сервісах часто використовуються рекламні блоки. Відключити рекламу користувач може за додаткову плату, яка поповнює бюджет проекту.

Соціальні мережі – сервіси, для спілкування в соціальних мережах за допомогою смартфона або планшета. Більшість популярних соцмереж, наприклад Facebook, Instagram, YouTube, Google має свій мобільний додаток, який може бути вбудований в базову комплектацію операційної системи смартфона.

За бізнес-спрямованістю мобільні додатки можуть бути двох видів: перші оптимізують внутрішні процеси організації, компанії, сервісу; інші використовуються в маркетинговому плані, тобто для комунікацій з клієнтами, продаж і просування бренду.

Серед програм для клієнтів в мобільному середовищі реалізуються різноманітні онлайн-сервіси. До цієї групи відносяться програми для бронювання квитків та номерів в готелях, відстеження посилок, інтернет-банкінгу, різні маркетплейси, онлайн-вітрини з товарами і послугами;

Програми лояльності для клієнтів. Додатки на зразок знижкових і бонусних карт для постійних клієнтів, дисконтні програми, кешбек і інші.

Окремий тип додатків – розважальні. До них відносяться тематичні/контентні програми та ігри.

Існує ще одна класифікація мобільних додатків. Це класифікація за особливостями роботи: нативні, гібридні та мобільні додатки для сайтів [5].

1. Мобільні веб-додатки та сайти (наприклад, last.fm чи google.com.ua/maps). У таких рішень є кілька плюсів – це кросплатформеність, простота створення та оновлення. Мінус в низькій функціональності. Це непоганий варіант для старту, щоб проаналізувати мобільний трафік в бізнес-ніші. Однак з такими програмами практично нічого не заробиш через їхній низький функціонал. Веб-додаток – це адаптована під характеристики екрану смартфона версія сайту. Веб-додатки створюються, щоб користувачі Інтернет могли відвідати сайт за допомогою смартфона, минаючи необхідність доступу до персонального комп'ютера або ноутбука. Деякі веб-сервіси потребують завантаження та встановлення, решта – запускається автоматично при заході на сайт через мобільний браузер.

2. Гібридні додатки – це вже більш сучасний варіант, який працює на API (наприклад, HeartCamera для iOS, що дозволяє прикрасити фото мальованими серцями, чи TripCase – органайзер для планування подорожей). У програмах вже є push-повідомлення, додаток може розміщуватися в плейсмаркетах для вільного або платного скачування. Такі програмні рішення мають можливість незалежного оновлення, що знімає необхідність випуску нових версій. Мають обмежений доступ до апаратної частини смартфона (камера, мікрофон, геолокація, адресна книга). Вимагають підключення до інтернету, оскільки

завантажують контент із зовнішнього джерела, розміщеного на сервері. Більшість промо-сервісів відноситься до цієї категорії.

Проте у гібридних програм є свої недоліки. Ці сервіси оперують малим обсягом інформації. Дизайн не адаптується до розміру і розширення екрану, що може викликати незручності.

3. Нативні додатки – самі «накручені» фичи, які дають максимальну функціональність і швидкість взаємодії (прикл. Shazam, Instagram). Однак для їх стабільної роботи потрібні серйозні ресурси системи. Створюються для конкретної операційної системи (iOS, Android, Windows). Для охоплення аудиторії потрібно розробити кілька окремих додатків для різних операційних систем. Вони можуть виконувати одні функції, мати однаковий дизайн, але будуть різними програмами. Ця необхідність збільшує термін роботи над проектом і бюджет розробки.

Нативні сервіси можуть працювати незалежно від підключення до інтернету, хоча частина з них вимагає наявності підключення. Вони займають менше пам'яті, працюють швидко, витрачають менше заряду батареї. Можуть отримати доступ до апаратної частини телефону з дозволу власника.

Отже, що стосується об'єкту проектування – це нативний інтелектуальний онлайн-сервіс з прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото, яка розроблена для мобільного додатку під операційну систему iOS/Android.

Функції об'єкту проектування: обробка, пошук та надання користувачам інформації за запитом.

Основне завдання – зчитування емоції з обличчя людини і на основі отриманих результатів надати список/перелік фільмів, які найкраще підійдуть під настрій людини.

Хоча в основі успіху будь-якого мобільного додатку лежить в першу чергу затребуваний контент, важливо відповісти на наступні питання: яка інфраструктура необхідна і як організувати якісний і масштабований сервіс?

Для відповіді на поставлені питання були визначені наступні задачі:

1. проаналізувати існуючі аналогічні системи та додатки, виділити основні риси, переваги та недоліки;
2. описати особливості функціоналу для майбутнього додатку, враховуючи побажання українських споживачів;
3. розробити технічне завдання, дизайн, інтерфейс та макет додатку згідно визначених особливостей;
4. реалізувати функціонал додатку з використанням сучасних методів та засобів розробки.

1.2 Аналіз існуючих у світі способів вирішення задачі

Теорія розпізнавання образу – розділ інформатики та суміжних дисциплін, що розвиває основи та методи класифікації та ідентифікації предметів, явищ, процесів, сигналів, ситуацій тощо об'єктів, які характеризуються кінцевим набором деяких властивостей та ознак.

Проблема розпізнавання образу набула визначного значення в умовах інформаційних перевантажень, коли людина не справляється з лінійно-послідовним розумінням повідомлень, що надходять до неї, внаслідок чого її мозок переключається на режим одночасності сприйняття і мислення, якому властиве таке розпізнавання.

Визначення людей на фото – практичний додаток теорії розпізнавання образів, завданням якого є автоматична локалізація людської особи на фотографії або зображенні і, у разі потреби, ідентифікація персони на основі наявних баз даних [6].

Розпізнавання облич існує вже десятки років, але стрибкоподібно ця технологія розвивається лише останніми роками завдяки досягненням у галузях комп'ютерного зору та штучного інтелекту.

Глибоке навчання (від англ. Deep learning) – це сукупність методів машинного навчання, що дозволяє навчати модель і передбачати результат набору вхідних даних. Технологія базується на штучних нейронних мережах.

Вони отримують алгоритми навчання та обсяги даних, які постійно зростають для підвищення ефективності навчання. Чим більше даних – тим ефективніший процес (рисунок 1.3)

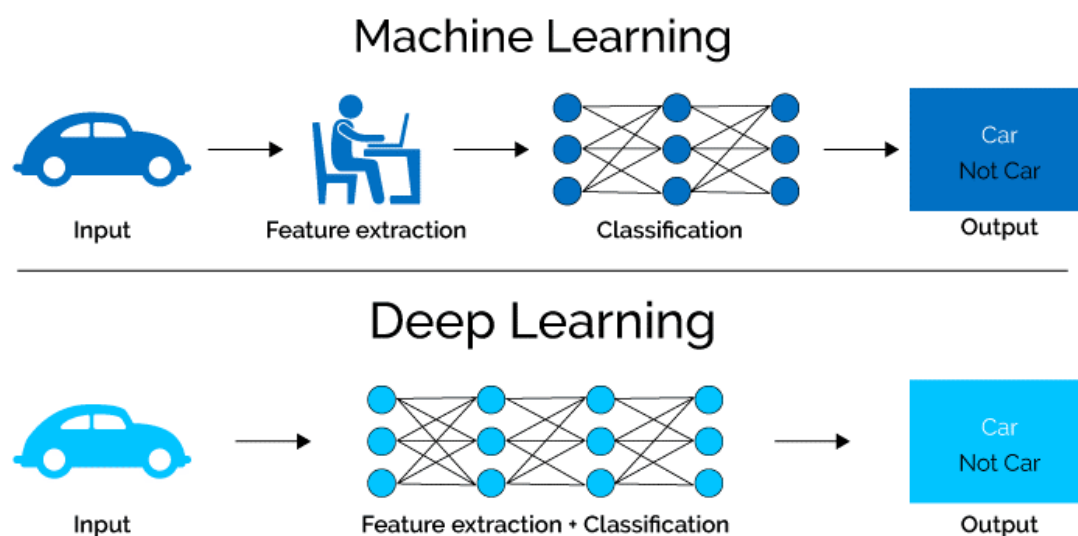


Рисунок 1.3 – Відмінність між машинним навчанням та глибоким навчанням

Навчання називається глибоким, тому що згодом нейронна мережа покриває дедалі більшу кількість рівнів. Подібний процес складається з двох фаз – навчання та формування висновків. Перший розглядається як спосіб маркування великих обсягів даних та визначення їх характеристик, які в подальшому порівнюються та запам'ятовуються системою. Це необхідно для того, щоб зробити правильні висновки, коли система стикається з подібними даними.

Етапи глибокого навчання:

1. Штучні нейронні мережі задають набір запитань з варіантами відповідей «Так» або «Ні»;
2. Відбувається вилучення числових значень із блоків даних;
3. Дані класифікуються відповідно до отриманих відповідей;
4. Дані маркуються.

Коли відбувається формування висновків, система робить деякі висновки, після чого маркує невивчені дані виходячи з отриманої раніше інформації.

Поява глибокого навчання призвело до значних досягнень у сфері комп'ютерного зору. Завдяки цьому ми можемо класифікувати зображення, знаходити в них об'єкти та відзначати їх заголовками. Для цього глибокі нейронні мережі з багатьма прихованими шарами можуть послідовно вивчати більш складні функції вихідного зображення:

- перші приховані шари можуть вивчати лише локальні контури;
- потім кожен наступний шар або фільтр вивчає складніші уявлення;
- останній шар може класифікувати зображення.

Ці типи глибоких нейронних мереж називаються згортковими нейронними мережами. Вони є багат шарові нейронні мережі, які припускають, що вхідні дані є зображеннями.

Інтерес до цих систем дуже великий у зв'язку з широким колом завдань, що вони вирішують:

- забезпечення безпеки у місцях великого скупчення людей;
- системи охорони, уникнення незаконного проникнення на територію об'єкта; пошук зловмисників;
- фейс-контроль у сегменті громадського харчування та розваг, пошук підозрілих та потенційно небезпечних відвідувачів;
- верифікація банківських карток;
- онлайн-платежі;
- контекстна реклама, цифровий маркетинг, Intelligent Signage та Digital Signage;
- фототехніка;
- криміналістика;
- телеконференції;
- мобільні додатки;
- пошук фото у великих базах фотографій;
- відмітка людей на фото в соціальних мережах та багато інших.

Масові скупчення людей створюють проблеми в різних областях (рітейл, держслужби, банки, забудовники). Замовникам необхідно об'єднувати та моніторити інформацію про кількість людей у багатьох місцях: в офісах обслуговування, адміністративних приміщеннях, на будівельних майданчиках тощо.

Завдання підрахунку людей мають готові рішення, наприклад, застосування камер з вбудованою аналітикою. Однак у багатьох випадках важливо використовувати велику кількість камер, які раніше вже встановлені в різних відділеннях [7].

Штучний інтелект (ШІ) – це система або машина, яка здатна імітувати людську поведінку для виконання певних завдань і може поступово навчатися, використовуючи отриману інформацію. На сьогоднішній день системи ШІ мають безліч втілень, наприклад [8]:

- чат-боти використовують ШІ, щоб швидше аналізувати звернення замовників та давати вірні відповіді;
- «розумні помічники» використовують ШІ, щоб витягувати інформацію з великих наборів даних у довільній формі та оптимізувати планування;
- механізми рекомендацій автоматично підбирають користувачам схожі телепрограми на основі переглянутих раніше тощо.

ШІ – це не формат і не функція, це процес та вміння думати та аналізувати дані. Незважаючи на те, що при згадуванні «штучного інтелекту» багато хто представляє розумних людиноподібних роботів, що захоплюють світ, насправді ШІ не призначений для заміни людей. Його мета – розширити межі людських здібностей та можливостей. Тому ця технологія є цінним бізнес-ресурсом.

Сьогодні термін «ШІ» широко використовується для позначення додатків для складних завдань, які раніше могли виконувати лише люди, наприклад, обслуговування замовників або гра в шахи. Нерідко його використовують як синонім машинного навчання та глибокого вивчення, які насправді є підрозділами науки про штучний інтелект і мають свою специфіку. Наприклад,

машинне навчання фокусується на створенні систем, які навчаються та розвиваються шляхом обробки та аналізу даних. Важливо розуміти, що якщо машинне навчання завжди має на увазі використання ШІ, то ШІ далеко не завжди має на увазі машинне навчання.

Одним із важливих напрямів штучного інтелекту є комп'ютерний зір. Комп'ютерний зір – це наука про комп'ютери та системи програмного забезпечення, які можуть розпізнавати та розуміти зображення. Комп'ютерний зір також складається з різних аспектів, таких як: розпізнавання зображень, виявлення об'єктів, генерація зображень, супер-роздільна здатність зображень тощо. Виявлення об'єктів є найбільш глибоким і складним аспектом комп'ютерного зору через величезну кількість практичних випадків.

Виявлення об'єкта посиляється на здатність комп'ютера та систем програмного забезпечення знаходити об'єкти на зображенні та ідентифікувати кожен об'єкт. Виявлення об'єкта широко використовується для виявлення осіб, виявлення транспортних засобів, підрахунку пішоходів, веб-зображень, систем безпеки та машин без водія.

Існує багато способів для використання технології виявлення об'єктів, так само як і багато областей для вивчення. Як і з будь-якою іншою комп'ютерною технологією, широкий спектр унікальних та креативних застосувань технології виявлення об'єктів, безумовно, виходитиме від програмістів та розробників програмного забезпечення.

На відміну від розпізнавання людей технологія розпізнавання емоцій людини має зовсім інші цілі. Якщо технологія розпізнавання людини чи обличчя застосовується для ідентифікації особистості людини, то розпізнавання емоцій спрямовано на їх виявлення та класифікацію шляхом аналізу будь-якого обличчя. Ці системи вже впливають на поведінку людей та функціонування соціальних інститутів, незважаючи на відсутність серйозних наукових доказів того, що вони справді ефективні [9].

Розпізнавання емоцій (зі статичних та динамічних зображень) — процес створення «зліпків» виразів обличчя для виявлення різних емоцій (огидність,

радість, гнів, здивування, страх чи смуток) на обличчі людини за допомогою програмного забезпечення для обробки відеозображень [10].

Вираз обличчя може бути представлений у вигляді геометричних моделей особливостей зовнішності або набору параметрів, вилучених із перетворених зображень. Ці характеристики надалі представляються як динамічних і 3D-моделей або обробляються за допомогою певних алгоритмів.

Популярність цієї технології пояснюється великою кількістю галузей, у яких потрібне подібне рішення. Розглянемо декілька прикладів.

В 2017 Apple випустила iPhone X, який можна розблокувати за допомогою сканера Face ID. Нова функція позиціонувалася як ще один спосіб посилити захист телефону від зловмисників (рис. 1.4).

Сьогодні ця технологія використовується не тільки для розблокування телефонів, а й для розпізнавання людей на кордоні, упіймання злочинців, підтвердження банківських операцій та здійснення функції оплати за допомогою сканування особи покупця на касі. Але деякі технологічні компанії доводять, що алгоритми розпізнавання осіб можуть аналізувати емоційний стан людини [11].



Рисунок 1.4 – Розблокування iPhone за допомогою сканера Face ID

Ще 1970 року психологи здатні розпізнавати приховані емоції, вивчаючи «мікровирази» на обличчях на фотографіях і відео. Проте сучасні алгоритми та камери високої чіткості впораються з цим навіть краще та швидше, кажуть у технологічних компаніях.

"Ми вже бачимо, як це використовується з комерційною метою", - пояснює Олів'є Філіппу, експерт із систем відеоспостереження в IHS Markit. "Супермаркети можуть використовувати цю технологію в проходах, не для встановлення особистості, а для аналізу публіки - у плані віку, статі, а також її настрою. Це може допомогти в цільовому маркетингу та розміщенні товарів", - розповідає Філіппу.

Агентство з досліджень ринку Kantar Millward Brown використовує технологію, розроблену американською фірмою Affectiva для вимірювання реакції глядачів на ТВ-рекламу. Affectiva записує відео з особами людей - з їхнього дозволу - і потім "кодує" їх вирази обличчя по кадрах, щоб зрозуміти їхній настрій. "Ми опитуємо людей, але набагато більше нюансів отримуємо, якщо дивимося і на вирази обличчя. Можна побачити, яка точно частина рекламного ролика добре працює і викликає емоційний відгук", - говорить Грехем Пейдж, керуючий директор з пропозицій та інновацій у Kantar Millward Brown.

Це може здатися сумнівним, але деякі стартапи пропонують "виявлення емоцій" як захід безпеки. Наприклад, британська компанія WeSee, що працює з правоохоронною організацією, аналізуючи людей під час інтерв'ю, стверджує, що штучний інтелект виявляє підозрілу поведінку, "читаючи" непомітні неозброєному оку підказки (рисунки 1.5). Такі емоції, як сумнів і агресивність, можуть ховатися під маскою і контрастувати з тим, що говорить людина.



Рисунок 1.5 – Використання технології WeSee для доступу до емоційного стану людей під час інтерв'ю [11]

"Використовуючи лише відео низької якості, наша технологія може встановити настрій людини або її наміри за допомогою її виразів, пози, жестів та рухів", - заявив Бі-бі-сі виконавчий директор Девід Фултон. "У майбутньому відеокамери на платформах станцій метро зможуть використовувати нашу технологію для розпізнавання підозрілої поведінки та сповіщати владу про потенційну терористичну загрозу". «Те саме можна використовувати в натовпах на заходах на кшталт футболу або політичних мітингах», - каже Фултон.

Експерт із систем відеоспостереження Олів'є Філіппу, однак, скептично відгукується про точність емоцій по камерах. "Коли потрібно просто розпізнати обличчя, все ще нерідко виникають помилки - навіть найкращі компанії кажуть, що можуть розпізнати людей із точністю 90-92 відсотки. Коли ви намагаєтеся оцінити ще й емоції, ймовірність помилки набагато більша", - наголошує він.

Поєднання психології та кіно стало предметом окремих наукових досліджень та цілої наукової галузі – нейросинематики. Вона вивчає вплив кіно на мозок та свідомість, настрої та почуття людей. Група вчених з Португалії [12]

навіть розробила програму, в якій глядачі можуть обирати фільми з набору емоцій, які хочуть випробувати.

На підставі низки досліджень [13-16], фільмам у кожному жанрі притаманні певні емоції. Звичайно, кожен фільм у різних людей може викликати різний набір емоцій та їхню велику кількість, але існує «маст-хев» жанру – центральний настрій, який прагнуть створити під час зйомок фільму. Вчені визначають такий настрій різними способами: проводять опитування, сканують мозок людей (МРТ), записують рухи очей, збирають враження у рецензіях та есе до фільмів.

Отримуємо 8 великих груп:

1. Смуток, занепокоєння, емпатія - драма;
2. Лють, хвилювання, агресія, сміливість – «фільми-дії»: пригоди, екшн/бойовик, військовий, вестерн, спорт;
3. Страх, «саспенс» (тривожне очікування чогось невідомого), задоволення від провини – кримінал, детектив, фільм-нуар;
4. Інтерес, здивування – фільми, що показують щось несподіване, незвичне: наукова фантастика, фентезі, анімація, історія;
5. Веселощі, задоволення – комедії;
6. Страх, тривога – трилери та фільми жахів, хоррор;
7. Любов, турбота, щастя – романтичні та сімейні фільми, мюзикли;
8. Цікавість, захоплення – фільми-біографії.

Нейробіологи з'ясували, що на сьогоднішній день немає переконливих доказів того, що вираз обличчя розкриває справжні емоції людини. Проте великі ІТ-гіганти протягом останніх кілька років випустили на ринок рішення, засновані на розпізнаванні емоцій людини. Причому вони використовуються в аеропортах, освітніх установах, програмному забезпеченні для виявлення психічних захворювань та різних інформаційних системах поліції [17].

Нині широко використовуються автоматизовані системи виявлення емоцій, наприклад, при прийому на роботу. Компанія з найму HireVue, яка підбирає персонал у тому числі для Goldman Sachs, Intel та Unilever, застосовує

машинне навчання для визначення компетенцій претендентів на роботу. У 2014 році компанія запустила свою систему штучного інтелекту для вивчення міміки особи, тону голосу та інших даних із відео співбесід на роботу, які вона використала для порівняння кандидатів на роботу з найкращими діючими співробітниками компанії. Після критики з боку вчених та правозахисників компанія відмовилася від аналізу особи у 2021 році. Однак кандидати досі проходять перевірку за допомогою голосового тону, який використовується як один із критеріїв оцінки при найманні на роботу.

У січні 2016 року Apple придбала стартап Emotient, який розробив програмне забезпечення, здатне розпізнавати емоції щодо зображень облич. Такі IT-гіганти, як Amazon, Microsoft і IBM, розробили системи виявлення емоцій. Microsoft пропонує розпізнавання сприйманих емоцій у своєму API Face, ідентифікуючи гнів, презирство, огиду, страх, щастя, смуток і здивування, тоді як інструмент розпізнавання Amazon аналогічним чином заявляє, що він може ідентифікувати те, що він характеризує як «усі сім емоцій» [9].

Аналіз та класифікація відеоконтенту пов'язані з методами вилучення інформації з відео або користувачів для пошуку значимих сегментів. Деякі методи використовують автоматичну обробку, інші – ручну. Вилучення емоційної інформації, вираженою в відео, ґрунтується на низькорівневих характеристиках, таких як колір, текстура, освітлення, рухи, звуки, ритм, лексика і т.ін., натхненних інструментами теоретиків кіно, які є результатом наміру режисерів викликати емоції [12].

Класифікація фільмів з їх фактичного ефективного впливу нещодавно була в центрі уваги деяких досліджень, що використовують: біометричні методи, засновані на фізіологічних сигналах, таких як дихання, частота серцевих скорочень, кров'яний тиск, електроміограми і гальванічна реакція шкіри; або розпізнавання емоційних виразів особи. Нещодавно були розроблено роботи, які доводять, що фільми можуть бути емоційними індукторами, які допомагають психологам у специфічному лікуванні, або для автоматичного

індексування, пошуку чи резюмування відео, відповідно до емоційного впливу, яке, як правило, значно відрізняється від глядачів [12].

В роботі [18] вивчалися емоції користувачів та залученість до інтернет-відео, використовуючи фізіологічні показники для оцінки валентності та збудження, емоційні самозвіти і особисті відгуки, і було зроблено висновок, що емоції впливають на переваги у відео.

Розглянемо системи, які знаходяться на перетині Інтернету та телебачення, так зване Інтернет-телебачення. Основний продукт та джерело доходу – послуга підписки, яка дозволяє користувачам дивитися будь-які відеозаписи з колекції фільмів та телепередач у будь-який час на широкому спектрі підключених до Інтернету пристроїв. Простір інтернет-телебачення молодий, конкуренція висока, тому інновації мають вирішальне значення.

YouTube, ймовірно, найвідоміший сайт для публікації та перегляду відео, пошуку, коментування, обміну та отримання рекомендацій. IMDb(.com) надає інформацію про акторів, режисерів, жанри, рейтинги фільмів і т.ін. Інші сайти, наприклад Netflix, також дозволяють отримати доступ і дивитися фільми.

Але жодна з цих систем не підтримує емоційну інформацію, і вони не досліджують візуалізацію відеопростору далі списків.

Ключовим елементом продукту Netflix є система рекомендацій, яка допомагає користувачам знаходити відео для перегляду кожного сеансу. Система рекомендацій Netflix – це не один алгоритм, а скоріше набір різних алгоритмів, що обслуговують різні сценарії використання, які поєднуються для створення повного досвіду Netflix (на даний момент у Netflix понад 65 мільйонів користувачів, які щодня переглядають понад 100 мільйонів годин фільмів та телепередач) [19].

Система Film Finder [20] підтримує пошук та перегляд фільмів за тривалістю, жанрами, назвою, акторським та режисерським складом, використовуючи графіки зоряного поля, засновані на даті та популярності. Однак більшість інструментів та програм для візуалізації не розглядають відео. Серед винятків, найбільш пов'язаних з даною роботою можна назвати:

- 1) двомірне представлення YouTube, що представляє відео як розрізнені кругові нерухомі зображення, що дозволяє візуально орієнтуватися в околицях на основі подібності;
- 2) Video Sphere (Bestiario.org/research/videosphere), що представляє відео TED у вигляді тривимірної сфери, пов'язаною семантичними зв'язками.

У роботі [21] автори надали інтерактивну 3D візуалізацію та навігацію по відео, щоб досліджувати культурні та естетичні властивості відео та відеопростору; та 2D інтерактивну систему, засновану на системі фізичних частинок [22] для візуалізації та вивчення відео на основі переважання кольору, ритму та руху. У цих системах розглядається візуалізація відео, але не емоцій.

Ще одна система – Система iFelt – це інтерактивне веб-відео додаток, що дозволяє каталогізувати, отримувати доступ, досліджувати і візуалізувати емоційну інформацію про фільми. Воно створено для вивчення афективних аспектів фільмів з точки зору їх властивостей і відповідно до емоційними профілями, вибором та станами користувачів [12].

Програма Emovi із підбору фільмів на основі емодзі, створена у червні 2019 року групою студентів технічних університетів, за чотири дні потрапила у топи App Store та Google Play. Сервіс допомагає вибирати фільми за настроєм або їх комбінаціями, а не лише за стандартною шкалою рейтингу від одного до десяти [23]. Для зручності емоції представлені як емодзі. Система використовує нейронні мережі, щоб надавати фільмам певні настрої. Користувачі також можуть «емоційно» оцінювати фільми після перегляду, що покращує рекомендації та полегшує вибір.

Програма створювалась як російський аналог існуючим рекомендаційним системам американських онлайн-кінотеатрів. Тому що сьогоденні аналоги для стран минулого СНД не можуть похвалитися високою точністю підбору релевантного контенту. Ретельно вивчивши ринок, розробники визначили мету – нова рекомендаційна система плюс агрегатом (рисунок 1.6).

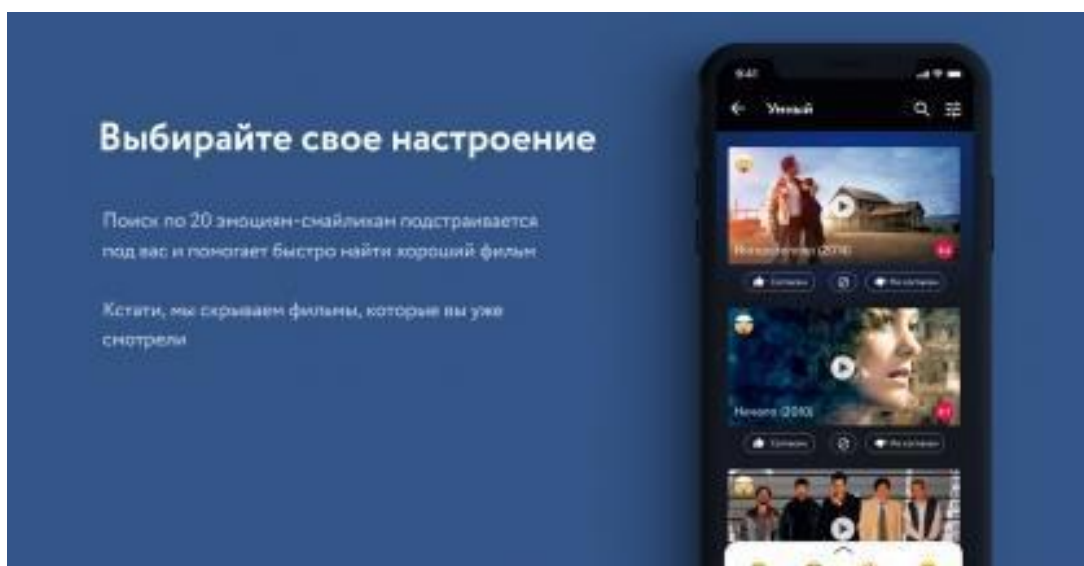


Рисунок 1.6 – Зовнішній вигляд програми eMovi [23]

Надалі планується розвивати eMovi, а також створити єдиний сервіс для вибору та перегляду фільмів, об'єднавши всі онлайн-кінотеатри в одному місці.

Отже, можна стверджувати, що ідея автоматичного розпізнавання емоцій дуже приваблива та прибуткова. ІТ-компанії згенерували величезний обсяг фото людських емоцій серед селфі в Instagram, фото в Pinterest, відео TikTok та фотографій Flickr. Як і розпізнавання осіб, розпізнавання емоцій стало частиною базової інфраструктури багатьох платформ, зокрема стартапів.

1.3 Визначення вхідних даних для вирішення поставленої задачі

Головна перевага розробки мобільних додатків як сучасної сфери бізнесу – великий попит та популярність. Завантаженість додатків перевищує показник у 200 млрд. у 2021 році (Statista). Отже в найближчому майбутньому перспективність даного напрямку беззаперечна. Але необхідно пам'ятати, що в умовах високої конкуренції, коли пропозиція перевищує попит, а для вирішення одних й тих самих задач вже існують сотні схожих рішень, щоб обійти конкурентів, необхідно запропонувати користувачеві не просто програмний продукт, а щось інноваційне – або сам контент, або якісно новий рівень реалізації вже існуючої ідеї.

Другий головний аспект веб-розробки – це необхідність постійних оновлень. Додаток має працювати однаково добре на будь-яких пристроях, незалежно від моделі смартфона чи версії операційної системи. В протилежному випадку – додаток буде видалено.

Також програма повинна відповідати запитам та очікуванням користувачів. Якщо не запропонувати вчасно клієнтам нові покращені функції, то це встигнуть зробити конкуренти. Тому для утримання аудиторії, охоплення нових користувачів та збільшення рівня їхньої лояльності є необхідність постійного оновлення контенту та переліку функцій, що використовуються у додатку.

У цій роботі показано, як можна використовувати підхід розпізнавання емоцій із селфі фото для підбору фільмів під настрій та формування переліку рекомендацій до перегляду фільмів з цього контенту.

Перший експеримент полягає у розпізнаванні лицьових емоцій людини, таких як сум, щастя, подив (100% емоційний зміст).

Другий експеримент полягає в тому, щоб зробити прогнозування фільмографічних уподобань користувача. Після цього на просунутому етапі майбутнього дослідження об'єднуються обидва наступні етапи.

Наведемо функціональні та нефункціональні вимоги – вимоги, які визначають критерії якості роботи системи, у якій проектуватиметься інтелектуальна система з прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото:

Цільовою платформою є операційна система Android 9 PPR1.180610.0

Технічні вимоги: CPU: 4 ядра ARM Cortex-A53 2 ГГц; RAM: 2 Гб; внутрішня пам'ять 16 Гб; дисплей IPS 5,45", 1440×720 (18:9); GPU: PowerVR GE8320 до 650 МГц.

Ізольованість – система має працювати незалежно від представлення, тобто передбачається можливість замінити стандартні асети на інші, не втрачаючи функціоналу.

Гнучкість – система має можливість розширяти функціонал.

Доступність – наскільки безперебійною планується робота ІС. В даному випадку 24/7. Максимально сукупний допустимий час простою системи (downtime) в рік 99% (3-4 дні).

Оптимізація – функціонал та ігрова графіка мають оптимально використовувати ресурси пристрою.

1.4 Висновки до першого розділу

Таким чином технологія розпізнавання облич дедалі більше ускладнюється, і алгоритми вже здатні читати емоції. Поєднання нейронних мереж з поведінковим аналізом є найновішим напрямком розвитку технології, який продовжує швидко зростати та доопрацьовуватися завдяки стрімкому розвитку систем штучного інтелекту.

Існує безліч сфер, у яких детектор емоцій знайде собі застосування: запобігання масовим заворушенням, безпека на дорозі, статистика та, звичайно ж, роздрібний маркетинг. Що стосується масових заворушень, велика кількість осіб у кадрі з високим показником подразнення чи страху – явна ознака метушні чи бійки. У машині: штучний інтелект, який постійно спостерігає за емоційним станом водія, може заспокоювати його у разі сильного стресу, а також не давати йому заснути за кермом. Статистика і маркетинг чудово працюють разом: знаючи, як часто відвідувачі висловлюють позитивні емоції, дивлячись на той чи інший товар, маркетолог зможе легко визначити, чи вдало обрано місце для товару (впадає у вічі), чи достатньо він релевантний для обраної аудиторії. Це особливо актуально для нових видів товару.

Аналіз об'єкта дослідження, визначення проблеми та огляд існуючих реалізацій – їх переваг та недоліків – дозволив сформулювати чітке уявлення про обраний напрям та очікуваний кінцевий результат.

Сформовано вимоги до програмного продукту та користувацького інтерфейсу, а також задачі для його розробки.

РОЗДІЛ 2. УПРАВЛІННЯ ПРОЕКТОМ ОБРОБКИ ІНФОРМАЦІЇ

2.1 Методи обробки інформації

Проектований програмний продукт призначено для користувачів смартфонів та вирішує задачу прогнозування фільмографічних уподобань користувача по аналізі емоцій з селфі фото. Основний алгоритм складається з наступних етапів:

1. Реєстрація селфі-фото;
2. Первинна обробка фото;
3. Виділення обличчя на фото;
4. Виділення елементів обличчя;
5. Виділення ключових точок на обличчі;
6. Класифікація емоції;
7. Створення добірки фільмів (список фільмів) на основі аналізу емоції з селфі фото;
8. Інформація про фільм;
9. Перегляд трейлеру обраного фільму з YouTube.

Перші п'ять етапів відносяться до процесу автоматичного розпізнавання емоцій.

Після того, як користувач завантажує свою селфі фото, відбувається реєстрація фото у додатку.

Первинна або попередня обробка зображення включає видалення шумів, геометричні і кольорові перетворення (рис. 2.1).

Джерелами шуму можуть бути конструктивні недоліки засобів реєстрації фотографії, погане освітлення, механічні пошкодження на обладнанні, розташування об'єктива смартфона тощо.

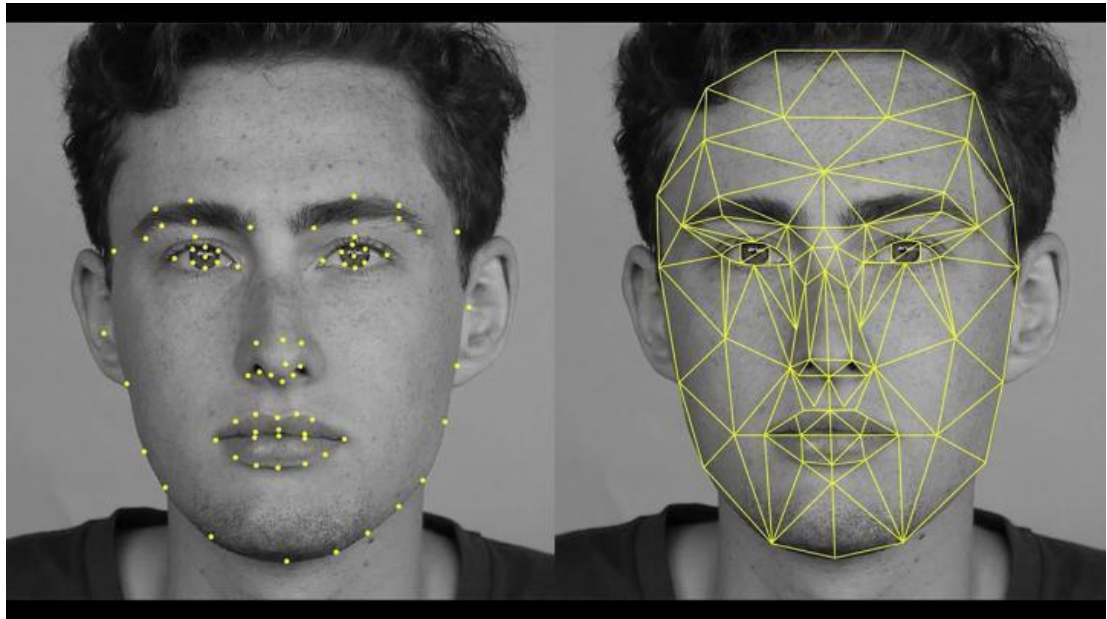


Рисунок 2.1 – Попередня обробка фото

Для видалення шумів можна застосувати середні фільтри (наприклад, фільтри, засновані на обчисленні середнього арифметичного, середнього геометричного, середнього гармонійного, середнього контр гармонійного) або фільтри, що базуються на порядкових статистиках (наприклад, медіанний фільтр, фільтр максимуму, фільтр мінімуму) [24].

Для виділення обличчя на зображенні можуть використовуватись декілька методів. Розглянемо деякі з них.

1. Метод основних компонент, в якому вхідні вектори є відцентрованими і приведені до єдиного масштабу зображення облич. Власні вектори, обчислені для набору зображень облич, називаються власними обличчями. За допомогою обчислених раніше матриць вхідне зображення розкладається на набір лінійних коефіцієнтів, званих основними компонентами. Сума N перших основних компонент, помножених на відповідні власні вектори є апроксимацією зображення порядку N . Для кожного зображення обличчя обчислюються його основні компоненти. Зазвичай береться від 55 до 200 основних компонент. Інші компоненти кодують дрібні відмінності між обличчями та шум.

Процес розпізнавання полягає у порівнянні основних компонент невідомого зображення з компонентами решти зображень.

2. Метод Віоли-Джонса, який був запропонований Полом Віолою і Майклом Джонсом в 2001 році. У цьому методі використовується принцип скануючого вікна. На вхід надходить початкове зображення розмірністю $N \times M$ пікселів, в якій кожен піксель має значення від 0 до 255 для кожного колірному каналу (одного колірному каналу для монохромного зображення і трьох колірних каналів для кольорового зображення). Вихідне зображення попередньо обробляється (масштабування, видалення шумів, застосування різних фільтрів тощо) і сканується за допомогою скануючого вікна. Скануюче вікно проходить по кожному пікселю зображення і до кожного положення застосовується класифікатор на основі методу бустингу – посилення слабких класифікаторів [24].

Процес виявлення обличчя методом Віоли-Джонса має такі особливості:

- для швидкого виконання необхідних розрахунків зображення представляються в інтегральному вигляді;
- пошук потрібних об'єктів на зображеннях здійснюється за результатами аналізу ознак Хаара;
- для вибору найбільш підходящих ознак при пошуку необхідного об'єкта на визначеній частині зображення застосовується метод посилення слабких класифікаторів (метод бустингу);
- для прийняття рішень використовуються прості бінарні класифікатори, які видають два значення - «Істина» і «Брехня»;
- для швидкого викидання фото, де не знайдено обличчя, використовуються каскади ознак.

Інтегральне представлення дозволяє виконати швидке обчислення сумарної яскравості довільного прямокутника на вихідному зображенні з постійним часом, незалежно від розмірів цього прямокутника. Таке представлення зображень – це матриця, що збігається за розмірами з вихідним

зображенням, в кожному елементі якої зберігається сума інтенсивностей всіх пікселів, що знаходяться лівіше і вище даного елемента.

Ще одна особливість, яку використовували Віола і Джонс у своєму методі – це вейвлети Хаара. Вейвлети Хаара – це прямокутні хвилі однакової довжини [25]. Однак прямокутні комбінації, які використовуються для виявлення об'єкта на зображенні, не є справжніми вейвлетами Хаара. Замість цього вони містять прямокутні комбінації, які краще підходять для завдань розпізнавання. Через цю різницю, ці функції називають не вейвлетами, а функціями або примітивами Хаара (рис.2.2).

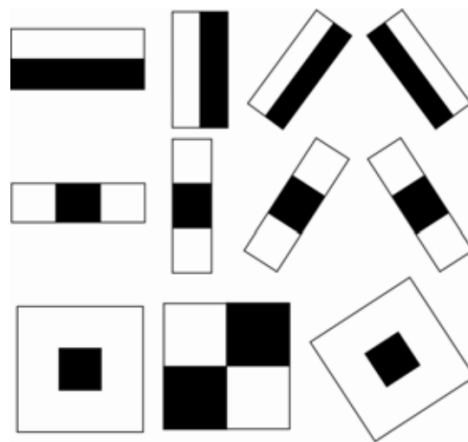


Рисунок 2.2 – Примітиви Хаара

Наявність функції Хаара визначається за допомогою віднімання середнього значення області темних пікселів із середнього значення області світлих пікселів. Якщо різниця перевищує поріг, який визначається в процесі навчання, то кажуть, що функція є існуючою.

3. Локальний бінарний шаблон є описом околиці пікселя зображення в двійковому уявленні. Початкове зображення розбивається на блоки 3×3 пікселі. Для обчислення ЛБШ в деякій точці зображення використовується вісім пікселів її околиці, а значення інтенсивності центрального пікселя приймається в якості порогу. Пікселі, які мають значення більше, ніж центральний піксель

(чи рівне йому), набувають значення = 1. Пікселі, які менше за центральний, набувають значень = 0. Так отримується восьмирозрядний бінарний код, який описує околицю цього пікселя, наприклад, 11100001, як показано на рис. 2.3.

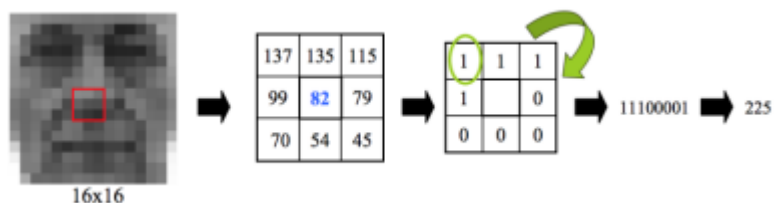


Рисунок 2.3 – Отримання локального бінарного шаблону пікселя у методі ЛБШ

ЛБШ обчислюється для кожного блоку зображення, після цього обчислюються гістограми кожного блоку і конкатенуються в загальну гістограму особливостей зображення людського обличчя. Приклад розбиття зображення на блоки і розрахунок гістограм показано на рисунку 2.4.

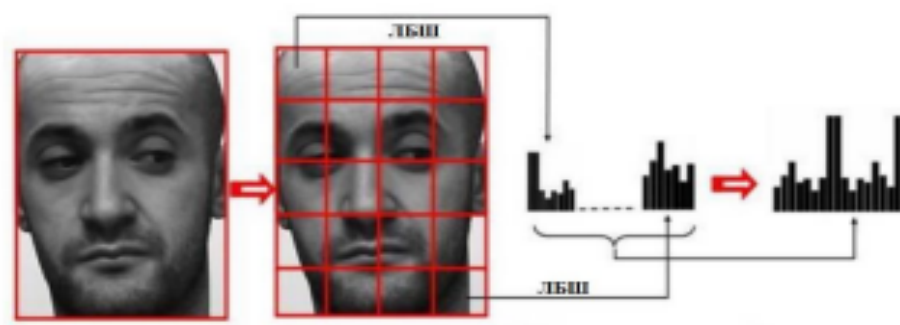


Рисунок 2.4 – Розбиття зображення на блоки і розрахунок гістограм

Розпізнавання відбувається за рахунок порівняння гістограм облич з тренувальної бази даних і гістограми обличчя, що розпізнається.

4. Метод порівняння еластичних графів, в якому обличчя представляються у вигляді графів із зваженими вершинами і ребрами. На етапі розпізнавання

один з графів – еталонний – залишається незмінним, тоді як інший деформується з метою найкращої підгонки до першого.

У вершинах графа обчислюються значення ознак, найчастіше використовують комплексні значення фільтрів Габора або їх впорядкованих наборів – Габоровських вейвлетів, які обчислюються в деякій локальній області вершини графа локально шляхом згортки значень яскравості пікселів з фільтрами Габора.

Ребра графа зважуються відстанями між суміжними вершинами. Відмінність між двома графами обчислюється за допомогою деякої цінової функції деформації, що враховує як відмінність між значеннями ознак, обчисленими у вершинах, так і міру деформації ребер графа. Результат розпізнавання системи – еталон з найкращим значенням цінової функції деформації.

5. Згорткова нейронна мережа (ЗНМ), відмінними особливостями якої є локальні рецепторні поля (забезпечують локальну двовимірну зв'язність нейронів), загальні ваги (забезпечують детектування деяких рис у будь-якому місці зображення) і ієрархічна організація з просторовими семплінгом (spatial subsampling).

Завдяки цим нововведенням ЗНМ забезпечує часткову стійкість до змін масштабу, зміщень, поворотів, зміни ракурсу і інших спотворень.

На рисунку 2.5 представлена архітектура згорткової нейронної мережі, яка складається з входу, згорткових і підвибіркових шарів (які чергуються між собою), шару звичайних нейронів і виходу.

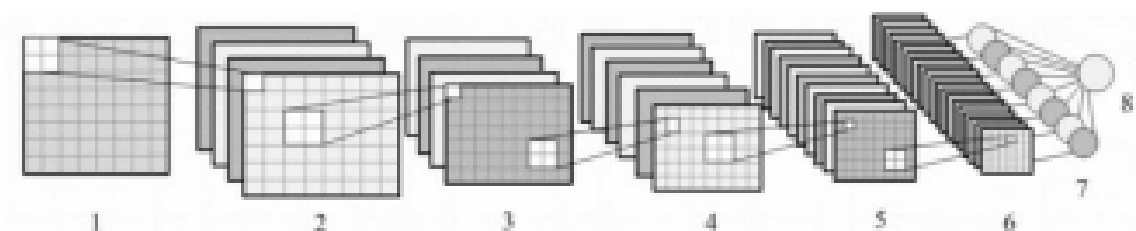


Рисунок 2.5 – Архітектура згорткової нейронної мережі

Порівняльна характеристика вищеописаних алгоритмів представлена в таблиці 2.1.

Таблиця 2.1 – Порівняльна характеристика алгоритмів [6]

	Метод головних компонент	Метод Віоли-Джонса спільно з використанням ЛБШ	Метод порівняння еластичних графів	Згорткова нейронна мережа
Швидке розпізнавання	+	+	-	+
Стійкість до змін освітлення	-	+	+	+
Стійкість до змін емоційного виразу обличчя	-	+	+	+
Стійкість до поворотів обличчя	-	+	+	+
Висока ефективність	-	+	+	+

Критеріями при виборі алгоритму є швидке розпізнавання та стійкість до змін. Метод Віоли-Джонса виявився найбільш ефективним, бо забезпечує достатньо високу швидкість і точність виявлення заданих об'єктів на зображеннях [24]. Проте хоча результати детектування обличчя дуже швидкі, але навчання класифікаторів йде дуже повільно, тому було вирішено поєднати метод Віоли-Джонса з локальними бінарними шаблонами (ЛБШ), які швидко справляються із завданням розпізнавання обличчя [6].

Після того як обличчя було виділено, необхідно виділити його основні елементи. Людина виявляє емоції за допомогою брів, очей і губ. Для виділення цих елементів на зображенні обличчя може використовуватися метод Віоли-Джонса. У якості вихідного зображення подається зображення обличчя, що було виділено на попередньому етапі. Щоб прискорити розпізнавання і зменшити помилкові виявлення, на обличчі задаються певні зони (рот завжди знаходиться у нижній половині обличчя, а брови і очі – у верхній).

Наступним кроком є знаходження ключових точок вібраних елементів обличчя. Визначити емоції можна на основі аналізу декількох ключових точок. Наприклад, на рисунку 2.6 показано комбінації точок брів та рота рис. 2.6 а) – зображення брів та рота; рис. 2.6 б) – комбінація ключових точок брів та рота, що відповідає їх зображенням; рис. 2.6 в) – комбінація ключових точок, що відповідає іншим положенням елементів, що розглядаються на зображенні [20].

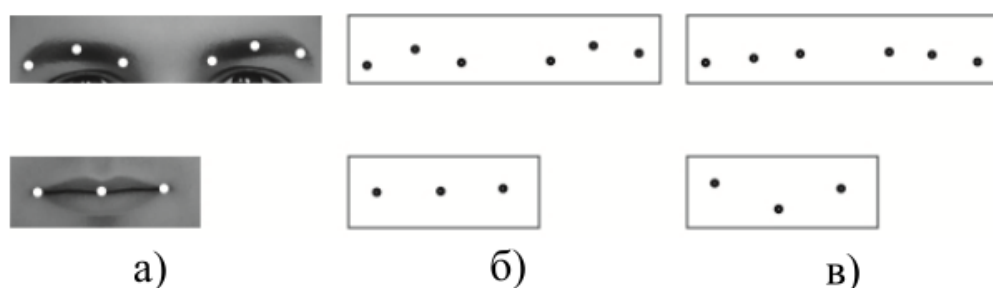


Рисунок 2.6 – Ключові точки елементів обличчя

Передостанній етап – це виділення ключових точок на обличчі, яке здійснюється таким чином:

1. Перетворення кольорового зображення до напівтонового вигляду;
2. Перетворення з півтонової форми до бінарного вигляду;
3. Застосування до бінарного зображення градієнтної маски;
4. Локалізація ключових точок.

Перехід від кольорового зображення до напівтонового здійснюється на етапі виділення обличчя. Для отримання бінарного зображення використовується адаптивний поріг. Застосування градієнтної маски до бінарного зображення дозволяє отримати контурне представлення елемента, що аналізується. Локалізація ключових точок полягає у визначенні певної кількості точок, що лежать на контурі елемента. Процес виділення ключових точок проілюстровано на рис. 2.7 [26].

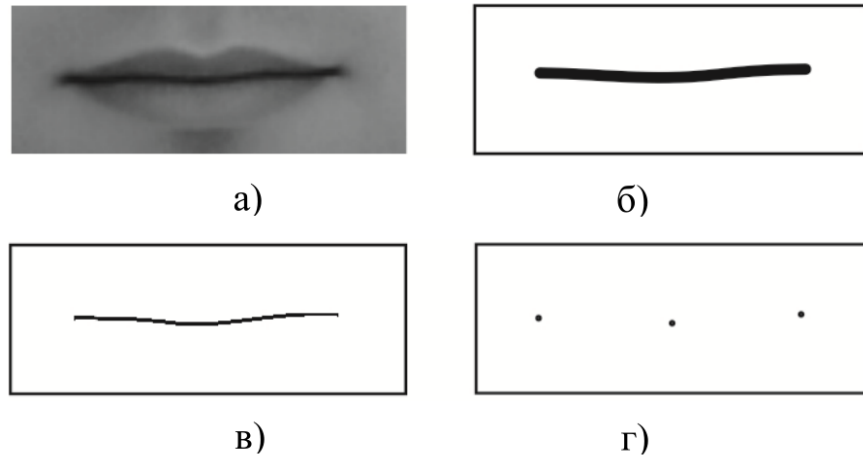


Рисунок 2.7 – Вибір ключових точок: а) напівтонове зображення, б) бінарне зображення, в) застосування градієнтної маски, г) локалізація ключових точок

Після виділення ключових точок можна виконати класифікацію емоцій. Наприклад, у таблиці 2.2 представлені характеристики емоцій з поєднання брови та рота [26].

Таблиця 2.2 – Характеристики емоцій за поєднанням брови та рота

Емоції	Брова	Рот
Здивування	Піднімається	Відкривається
Страх	Піднімається та зморщується	Відкривається та розтягується
Відраза	Знижується	Піднімається і кінці знижуються
Гнів	Знижується та зморщується	Відкривається або кінці знижуються
Щастя	Піднімається	Кінці піднімаються
Смуток	Кінці знижуються	Кінці знижуються

Це завершальний етап автоматичного розпізнавання емоцій у проєктованому додатку з прогнозування фільмографічних уподобань користувача по аналізі емоцій з селфі фото.

Наступним етапом алгоритму роботи проектованого додатку є формування переліку фільмів у відповідності до класифікованої емоції на обличчі з селфі фото, так звані фільмографічні вподобання користувача.

Парсинг інформації відбувається з ресурсу IMDb (Internet Movie Database – Інтернет-база фільмів).

IMDb.com – це веб-сайт з найбільшою в світі базою даних про кінематограф, яка вільно редагується. Станом на січень 2022 року в базі зібрано інформацію про більш ніж 6,5 млн кінофільмів, телесеріалів і окремих їхніх серій, а також близько 10,4 млн персоналій, пов'язаних з кіно і телебаченням, – акторів, режисерів, сценаристів та інших [27].

Майже вся інформація IMDb знаходиться у вільному доступі, сайт IMDb.com функціонує на основі вільного програмного забезпечення (Apache, Perl та ін.).

Парсинг – це збір та систематизація інформації, розміщену на певних сайтах, за допомогою спеціальних програм, що автоматизують процес [28]. В нашому випадку сайт IMDb.com.

Переваги парсингу у тому, що є можливість:

- збирати дані швидше і в будь-якому режимі, наприклад цілодобово;
- дотримуватися всіх вказаних параметрів, навіть дуже тонких;
- використання програмних засобів виключає помилки від неувважності або втоми;
- виконувати регулярну перевірку за заданим інтервалом (щодня, щотижня тощо);
- представити зібрані дані в будь-якому необхідному форматі без зайвих зусиль;
- рівномірно розподіляти навантаження на сайт, де проходить парсинг (зазвичай одна сторінка за 1-2 секунди), щоб не створювати ефект DDOS-атаки.

Алгоритм роботи парсингу залежить від цілей. Але схематично він виглядає так [28]:

1. Пошук на зазначених сайтах Інтернет-мережі дані, що відповідають параметрам.

2. Інформація збирається і проводиться початкова систематизація (її глибина також визначається при налаштуванні);

3. З даних формується звіт у форматі, що відповідає необхідним критеріям. Більшість сучасних парсерів мультиформатні і можуть успішно працювати хоч з PDF, хоч з архівами RAR, хоч з TXT.

В результаті парсингу в проєктованому додатку Користувач може обрати будь-який фільм з переліку, що був сформований на попередньому етапі та подивитись інформацію про нього.

Також Користувач може переглянути трейлер обраного фільму, який транслюється з youtube.com.

Тепер перейдемо до розробки програмного продукту.

2.2 Розробка проєкту проведення аналітичної діяльності

В рамках даної дипломної роботи запропоновано варіант вирішення задачі зчитування емоцій людини на фото. Але спочатку визначимося з видом програмного середовища для написання коду.

2.2.1 Вибір програмного середовища

Мобільний додаток можна розділити на два великі блоки – це Front та Back -end. Відповідно до частини Front-end входять компоненти і опції програми, з якою взаємодіє користувач. Наприклад, панель вибору, дашборд, налаштування опцій та інше. Back-end – це прихована частина. З цими компонентами взаємодіє розробник за допомогою серверного софту.

Іншими словами мобільний додаток нагадує спліт-систему, в якій одна частина знаходиться на стороні користувача – це Front-end, а інша на стороні розробника – це Back-end.

Використання сучасних методів виявлення об'єктів у додатках та системах, а також створення нових додатків на основі цих методів не є прямим завданням. Ранні реалізації технології виявлення об'єктів включали використання класичних алгоритмів, наприклад з тез, які підтримуються в OpenCV – популярної бібліотеці комп'ютерного зору. Проте ці класичні алгоритми не змогли забезпечити достатню продуктивність для роботи в різних умовах.

Швидке впровадження глибокого вивчення призвели до появи таких сучасних та високоточних алгоритмів та методів виявлення об'єктів, як R-CNN, Fast-RCNN, Faster-RCNN, RetinaNet та поки що швидких та високоточних SSD та YOLO. Використання цих методів та алгоритмів, заснованих на глибокому вивченні, що також базується на машинному навчанні, вимагає хорошого розуміння фреймворків математичного та глибокого навчання [29].

Існують мільйони експертів-програмістів та розробників програмного забезпечення, які хочуть інтегрувати та створити нові продукти, які використовують технологію виявлення об'єктів. Але ця технологія залишається поза досяжністю через додатковий складний шлях по її розумінню та практичному використанню.

Для вирішення цієї проблеми провідними розробниками ПЗ були створені бібліотеки Python, які дозволяють програмістам та розробникам програмного забезпечення легко інтегрувати новітні технології комп'ютерного зору у свої вже існуючі або нові програми, використовуючи лише кілька рядків коду.

Python – це динамічно типізована, універсальна мова програмування, призначена для швидкого (вивчення, використання та розуміння) застосування чистого та одноманітного синтаксису.

Як стверджується авторами в [30-32] Python вважається найкращою мовою програмування для роботи в галузях машинного навчання та великих даних із наступними можливостями:

1. Об'єктно-орієнтоване програмування (ООП);
2. Узагальнене та функціональне програмування;

3. Модулі та пакети. Програмне забезпечення (ПЗ) на Пітоні оформляється у вигляді модулів, які можуть бути зібрані в пакети;
4. Інтроспекція – можливість запросити тип та структуру об'єкта під час виконання програми. У мові програмування Python для будь-якого об'єкта можна дізнатися всю інформацію про його внутрішню структуру;
5. Обробка винятків, ітератори та генератори;
6. Управління контекстом виконання;
7. Декоратори, регулярні вирази.

Це визначає сфери використання: програмні продукти, мобільні додатки, ігри (наприклад, World of Tanks), вбудовані системи(наприклад, для керування банкоматами), скрипти і плагіни для створення нових модулів та автоматизації процесів, алгоритми для Machine Learning програм, а також багато аналітичних програм, тестування (автоматизація тестування).

До переваг, зазначених у [30], відносять:

1. Мова характеризується логічним синтаксисом, внаслідок чого вихідний код програм, написаних на Python, легко читається та сприймається.
2. Умовна легкість. Він вважається найбільш підходящим для фахівців-початківців. Крім того, якщо розробник стикається з питаннями та труднощами, він завжди може запитати поради у колег у велика інтернет-спільноті що значно прискорює вирішення проблем.
3. Гнучкість та масштабованість. Пайтон дозволяє розробникам адаптувати високорівневу логіку програми, що дозволяє легко розширювати складні програми.
4. Python є інтерпретованою мовою програмування. Це означає, що до запуску він є звичайним текстовим файлом. Відповідно програмувати можна майже на всіх платформах.

До недоліків можна віднести [31]:

1. У порівнянні з Ruby та деякими іншими мовами, у Python відсутня можливість модифікувати вбудовані класи, такі як int, str, float, list та інші.

Отже Ruby у веб-розробці зазвичай ефективніше та економічно вигідніше (з ідентичною якістю результату).

2. Мови компілювання (Go/Elixir/Java) – значно ефективніші за продуктивністю.
3. Розробка проекту на PHP буде помітно дешевшою.

Проте в сферах аналізу даних і машинного навчання Python зараз поза конкуренцією. У рейтингу GitHub Octoverse (Рейтинг популярності мови серед користувачів GitHub) за 2020 рік [33] Python посідає друге місце, поступаючись лише JavaScript.

У рейтингу RedMonk [34] Python також посідає друге місце. Співзасновник RedMonk Джеймс Гавернер зазначає, що Python вже став основною мовою Data Science.

Проте згідно з індексом ТЮВЕ [35], у листопаді 2021 року Python посів перше місце у списку найпопулярніших мов програмування, у листопаді минулого року поступав лише С. Він випереджає JavaScript, PHP, Swift та інші поширені мови.

Український профільний ресурс DOU.UA провів власне чергове щорічне опитування [36] про мови програмування, в рамках якого було зібрано 7211 анкет (92% респондентів перебувають в Україні). І виявилось, що вперше з 2014 року у Python негативна динаміка, спостерігається тенденція переходу з JavaScript на TypeScript, найбільш задоволені користувачі Elixir, і нарешті на графіках з'явився Rust (рис. 2.8).

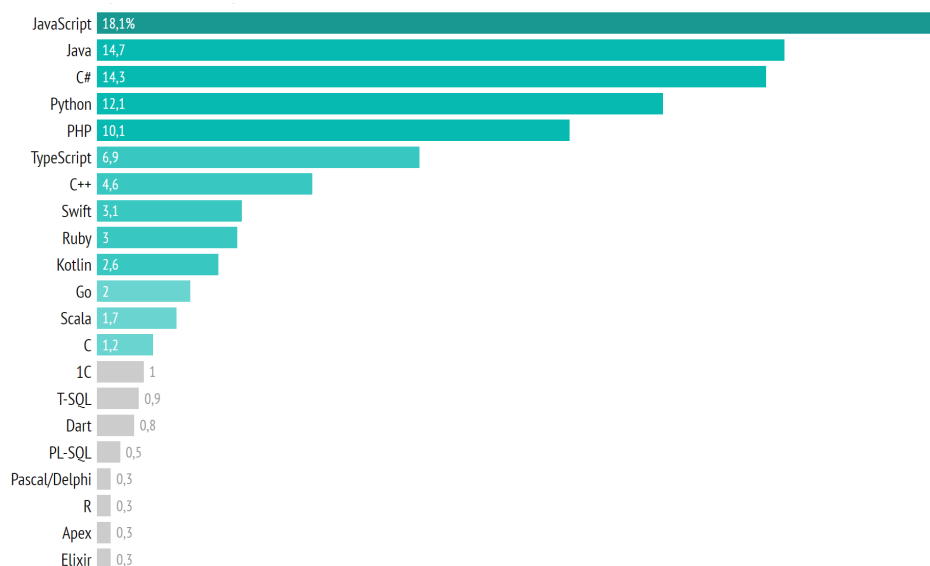


Рисунок 2.8 – Використання мов програмування [36]

Проте, якщо врахувати сферу застосування (бекенд, сегменти фронтенд і мобільної розробки, обробка масивів даних (туди відносяться і Big Data і машинне навчання), Desktop-програмування, системне програмування (зараз невелика частка) і Full-Stack розробка), то серед лідерів знаходиться ще одна дуже популярна мова PHP (Hypertext Preprocessor) – поширена мова програмування спільного призначення з відкритим вихідним кодом (рис. 2.9).

PHP була опублікована у 1995 р. Расмусом Лердорфом, з націлюванням на підтримку маленьких, простих динамічних веб-додатків, на кшталт гостьових книг та лічильників відвідувачів, популярних на зорі Інтернету.

З моменту релізу PHP, вона була використана у набагато складніших проектах, ніж спочатку очікували її автори. Вона зазнала кількох мажорних змін, кожна з яких принесла нові механізми для приборкання цих складних програм. Сьогодні, в 2022 році, вона є багатою фічами членом сім'ї Змішаної Парадигми Продуктивних Мов (MPDPL) [36], яка включає JavaScript, Python, Ruby та Lua.

PHP – це широко використовувана мова сценаріїв спільного призначення з відкритим вихідним кодом. PHP спеціально сконструйований для веб-розробок і його код може вбудовуватися безпосередньо в HTML.

Простіше кажучи, PHP – це мова програмування, спеціально розроблена для написання веб-програм (сценаріїв), що виконуються на веб-сервері.

Перевагою PHP є надання web-розробникам можливості швидкого створення динамічно генерованих web-сторінок.

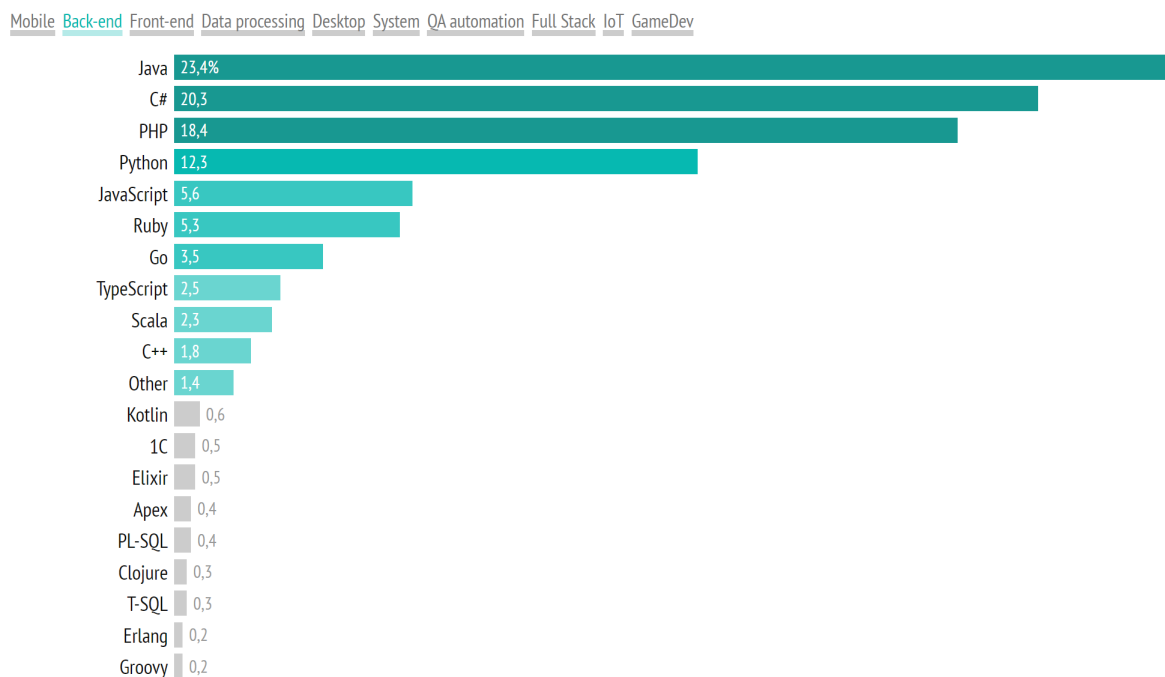


Рисунок 2.9 – Використання мов програмування у сегменті Back-end-розробки [36]

Важливою перевагою мови PHP перед такими мовами, як мови Perl і C полягає в можливості створення HTML документів з впровадженими командами PHP.

Значною відмінністю PHP від якогось коду, що виконується на боці клієнта, наприклад, JavaScript, є те, що PHP-скрипти виконуються на боці сервера. Є можливість настроїти сервер таким чином, щоб HTML-файли оброблялися процесором PHP, тож клієнти навіть не зможуть дізнатися, чи отримують вони звичайний HTML-файл або результат виконання скрипту.

PHP дозволяє створювати якісні Web-програми за дуже короткі терміни, отримуючи продукти, які легко модифікуються і підтримуються в майбутньому.

PHP простий для освоєння, і разом з тим здатний задовольнити запити професійних програмістів, про що доводить його позиція у ТОП-10 провідних мов програмування 2021 року [36-39].

Крім того, мова інтегрована з низкою популярних баз даних, включаючи MySQL, PostgreSQL, Oracle, Sybase, Informix і Microsoft SQL Server.

Мова PHP постійно вдосконалюється, і їй забезпечено довге майбутнє в області мов web-програмування, принаймні, найближчим часом [37].

У даний час PHP – одна з лідерів серед серверних мов програмування, що застосовуються для створення динамічних веб-сайтів та веб-програм. Велика частина систем управління сайтами написана саме на PHP, мова підтримується переважною більшістю хостинг-провайдерів, мова набула широкого поширення завдяки своїй простоті, швидкості, багатій функціональності та кросплатформеності.

Більш того, PHP стала популярною завдяки відкритому вихідному коду та універсальній функціональності. Завдяки невисокому порогу входження її можуть використовувати як розробники-початківці, так і професійні програмісти для більш складних завдань.

Вивчення і розуміння цієї мови, безсумнівно, відкриває нові можливості: створення будь-яких типів сайтів, інструмент для розробників WordPress або розробка веб-додатків.

Синтаксис PHP схожий на синтаксис мови C/C+. Деякі елементи, наприклад, асоціативні масиви, запозичені з Perl. Робота програми може починатися з оператора PHP, описувати змінні не потрібно.

Мова виконує код всередині обмежувачів `<? php? >`. Дані поза тегами з'являються у стандартній розмітці HTML документа. Назва змінної починається з символу `$` і чутливо до регістру, як і назви класів, функцій і константи. Змінні обробляються в укладених у подвійні лапки або апострофи рядках, а також створених за допомогою оператора `<<<`.

Скрипти обробляються інтерпретатором в порядку, що забезпечує кросплатформеність програми [38].

Як видно з вищесказаного PHP – дуже популярна мова програмування, але при цьому має досить спірну репутацію. Основні недоліки PHP зводяться до таких проблем [39]:

1. Неузгодженість навколо посилань та семантичних значень. PHP 3 мала чітку семантику передачі аргументів, повернення всього за значенням, створюючи логічну копію даних у запиті. Програміст може вибрати семантику посилання разом зі знаком `&`. Це виникло разом із запровадженням об'єктно-орієнтованих засобів програмування в PHP 4 і 5. Більшість PHP об'єктно-орієнтованих анотацій були запозичені з Java, і Java має семантику, у якій об'єкт передається за посиланням, тоді як примітивні типи передаються за значенням.

У підсумку, поточний стан семантики PHP полягає в тому, що об'єкти передаються за посиланням (вибираємо Java, замість, скажімо, C++), примітивні типи передаються за значенням (тут Java, C++, і PHP), але стара семантика і знак `&` залишилися, іноді взаємодіючи з новим світом неоднозначними методами.

2. Філософія обчислень, що ігнорують відмови. PHP намагається зберігати запит запущеним, навіть якщо вже процес йде неналежним чином. Так, наприклад, ділення на нуль не залишає винятки, не повертає INF, і не завершує фатально запит. За замовчуванням відбувається просто попередження і присвоюється значення як `false`. Оскільки `false` за замовчуванням розглядається як 0 у числових контекстах, безліч додатків розгортаються та запускаються з недіагностованими поділами на нуль. Саме ця проблема була вирішена в PHP 7, але імпульс у дизайні до обробки неоднозначностей, навіть коли вони можуть мати сенс, має місце навіть у бібліотеках.

3. Суперечності у стандартній бібліотеці, пов'язані з тим фактом, що на початку використання мови PHP, його аудиторія була найбільш знайома з C, і безліч API використовували дизайн стандартної бібліотеки мови C: шести-символьні імена в нижньому регістрі, відповіді про успішне/неуспішне виконання, що повертають реальне значення в параметр «out», і так далі. З

розвитком PHP, цей стиль поділу на простори імен через префікси з `_` став більш поширеним: `mysql_...`, `json_...`, тощо. А нещодавно, `camelCase` стиль іменування методів з Java на класах `CamelCase` став найпоширенішим способом введення нових функціональних можливостей.

Тому в результаті, іноді можна зустріти приклади коду з перемішаними виразами на кшталт `DirectoryIterator($path)` разом з `if (!($f = fopen($p, 'w+')))`... в логіці, що збиває з пантелику.

4. Неточності під час перетворення. Майже всі мови в наші дні дозволяють порівняти, наприклад, `integer` та `float` з оператором `>=`. Цілком зрозуміло, що тут мається на увазі. Менш очевидним є порівняння рядка і числа за допомогою `==`, і різні мови робили різний вибір. Вибір PHP у цій ситуації такий, що призводить до неприємних помилок. Наприклад, `123 == '123foo'` оцінюється як істина, але `0123 == '0123foo'` є брехнею.

Отже, хоча мова PHP має безліч недоліків, які, безсумнівно, уповільнили її розвиток, але середовище PHP має такі переваги, які більш ніж компенсують дані недоліки. І вона також має способи вирішення своїх власних мовних проблем, які цілком вражають. За підсумковими результатами, PHP надає найкращий фундамент для створення, зміни та експлуатації успішних веб-проектів порівняно з конкуруючими середовищами.

Таким чином для вирішення задачі проектування в даній роботі для реалізація алгоритму прогнозування фільмографічних уподобань користувача по аналізі емоцій з селфі фото обрано мову програмування PHP.

2.2.2 Система управління базами даних MySQL

PHP підтримує взаємодію з великою кількістю різних систем управління базами даних (СУБД). Вданій роботи використовується MySQL. MYSQL була розроблена шведською компанією 1984 року. У 2008-му систему управління викупила американська компанія Sun Microsystems, а в 2010-му – Oracle [40].

Велика кількість веб-додатків, таких як Facebook, Twitter, YouTube, Google і Yahoo! використовують MySQL для зберігання даних. Хоча спочатку він створювався для обмеженого використання, тепер він сумісний з багатьма важливими обчислювальними платформами, такими як Linux, macOS, Microsoft Windows і Ubuntu [41].

База MySQL – це реляційна база. Так називаються бази, організовані у вигляді пов’язаних між собою таблиць. Можна змінювати відкритий початковий код або самостійно встановити програму. Код можна налаштувати під конкретні задачі. Тільки у випадку необхідності розширеної підтримки, потрібно придбати ліцензійну версію.

Комп’ютери, що запускають програму СУБД, називаються клієнтами. Для отримання даних вони підключаються до сервера. Така система і називається «клієнт-сервер» (рис. 2.10).

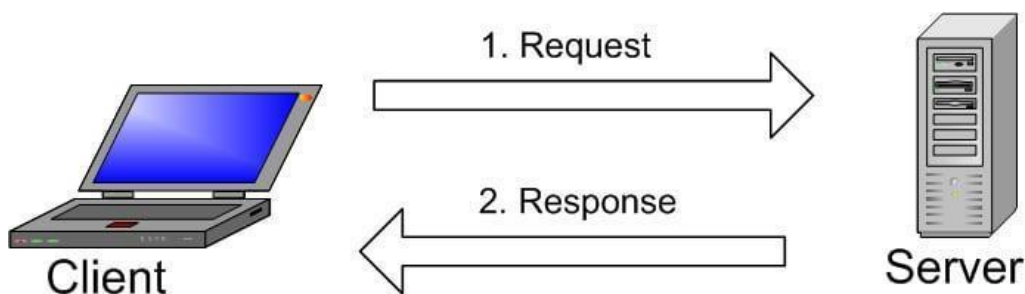


Рисунок 2.10 – Базова структура «клієнт-сервер»

Рисунок 2.10 пояснює базову структуру «клієнт-сервер». Один або декілька пристроїв (клієнтів) з’єднуються з сервером через певну мережу. Кожен клієнт може зробити запит з графічного інтерфейсу користувача (GUI) на своїх екранах, і сервер видасть бажаний результат, якщо обидва кінці розуміють інструкцію.

Клієнт і сервер взаємодіють в середовищі СУБД з використанням специфічної для домену мови – мову структурованих запитів (SQL), яка

використовується як основна мова для взаємодії з базою даних. Сам MySQL написаний на C і C++.

Основні процеси, що відбуваються в середовищі MySQL, однакові [41]:

1. MySQL створює базу даних для зберігання і керування даними, що визначають відносини кожної таблиці.
2. Клієнти можуть робити запити, вводячи певні команди SQL на MySQL.
3. Програма сервера відповість запитаною інформацією і з'явиться на боці клієнта.

MySQL не єдина СУРБД на ринку, але вона є однією з найпопулярніших і поступається тільки Oracle Database, коли оцінюється з використанням таких важливих параметрів, як кількість згадок у результатах пошуку, професійних профілів в LinkedIn і частоти технічних дискусій на інтернет-форумах. Той факт, що багато великих технологічних гігантів покладаються на нього, ще більше зміцнює заслужену позицію. Чому так? Причини наступні [40]:

1. Гнучкість і простота у використанні – є можливість змінити початковий код, процес установки відносно простий і не повинен займати більше 30 хвилин.
2. Висока продуктивність – широкий спектр кластерних серверів підтримує MySQL.
3. Промисловий стандарт – галузі використовують MySQL протягом багатьох років, а це означає, що для досвідчених розробників є рясні ресурси.
4. Безпека даних – за допомогою системи доступу і управління обліковими записами MySQL встановлює високий рівень безпеки. Доступна перевірка на основі вузла і шифрування пароля.

2.2.3 Створення користувацьких інтерфейсів

Інструмент для створення користувацьких інтерфейсів – це React. Його основне завдання забезпечити відображення на екрані того, що можна побачити

на веб-сторінках. React дозволяє легко створювати інтерфейси, розділяючи кожен сторінку на невеликі фрагменти і компоненти. Він дуже зручний для створення веб-додатків, крім цього не вимагає великого порогу входження [42].

Тобто, React Native – це JS-фреймворк для створення нативних iOS- і Android-додатків. В його основі лежить розроблена в Facebook JS-бібліотека React, призначена для створення користувацьких інтерфейсів. Але замість браузерів вона орієнтована на мобільні платформи. Іншими словами, якщо ви веб-розробник, то можете використовувати React Native для написання чистих, швидких мобільних додатків, не залишаючи комфорту звичного фреймворка і єдиної кодової бази JavaScript [43].

Інакше кажучи, React – це JS-бібліотека для створення користувацьких інтерфейсів, зазвичай для веб-додатків. Вона поширюється під ліцензією open source з 2013 року. React широко поширена, і на відміну від більш великих MVC-фреймворків вирішує відносно вузьку задачу: рендеринг інтерфейсу.

React популярний з кількох причин. Він компактний і відрізняється високою продуктивністю, особливо при роботі з швидкозмінними даними. За рахунок своєї компонентної структури, React заохочує писати модульний і багаторазово використовуваний код.

Перше, що дивує в React Native – він «дійсно» нативний. Інші рішення «JavaScript-для-мобільних-платформ» просто «обертають» JS-код у веб-уявлення. Вони можуть ще раз реалізувати яку-небудь нативну поведінку інтерфейсу, наприклад, анімацію, але все ж це залишається веб-додатком. В React компонент описує власне відображення, а потім бібліотека обробляє рендеринг. Ці дві функції розділені прозорим рівнем абстракції. Якщо потрібно відобразити компоненти для веб, то React використовує стандартні HTML-теги. Завдяки тому ж рівню абстракції – «мосту» – для рендеринга в iOS і Android React Native викликає відповідні API. В iOS компоненти відображаються в справжні UI-види, а в Android – в нативні (рис. 2.11).

Якщо вихідний код дуже схожий на стандартний JavaScript, CSS і HTML, то замість компіляції в нативний код React Native запускає додаток за

допомогою JS-движка хост-платформи, без блокування основного UI-потoku. Таким чином, отримуються переваги нативних продуктивності, анімації і поведінки без необхідності писати на Objective-C або Java. Інші методи розробки кросс-платформних додатків, на кшталт Cordova або Titanium, ніколи не досягнуть такого рівня нативної продуктивності або відображення.

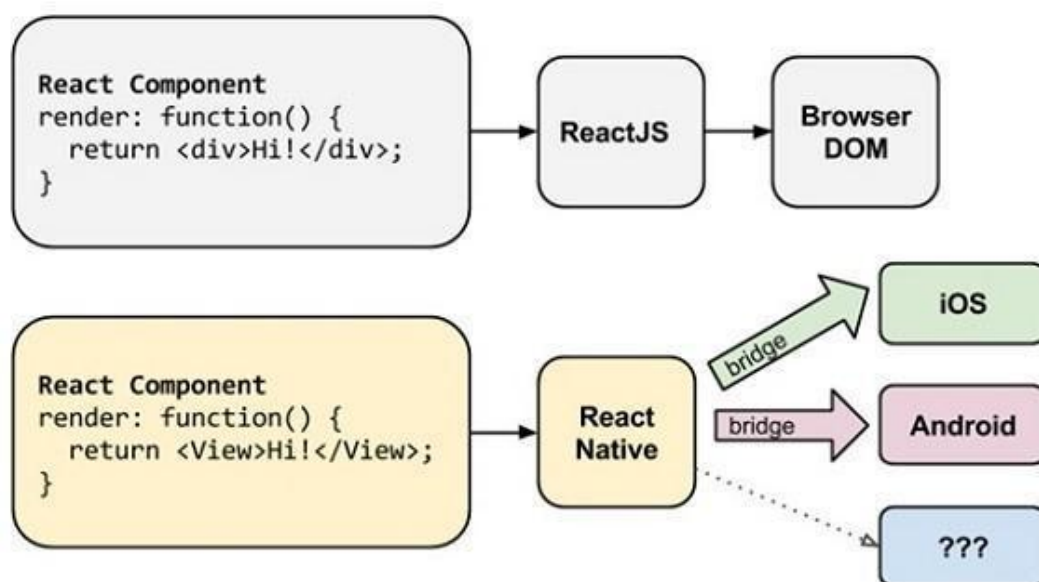


Рисунок 2.11 Рендеринг в iOS і Android [43]

Як показано у [43], у порівнянні зі стандартною розробкою під iOS і Android, React Native має набагато більше переваг. Оскільки проєктований додаток здебільшого складається з JavaScript, то є можливість користуватися численними перевагами веб-розробки. Наприклад, для того, щоб побачити внесені в код зміни, можна миттєво «оновити» додаток замість тривалого очікування завершення традиційного rebuild.

Крім того, React Native надає «розумну» систему повідомлень про помилки і стандартні інструменти налагодження JavaScript, що сильно полегшує процес мобільної розробки.

Ще одна істотна перевага React Native – витончена обробка різних платформ. Переважна більшість API у фреймворку – кросплатформені, так що досить просто написати компонент React Native, і він буде без проблем

працювати в iOS і Android. У Facebook заявляють, що в їх додатку Ad Manager 87% коду багаторазово використовується на обох платформах [43].

Якщо навіть потрібно написати залежний від платформи код – у зв'язку з різними правилами взаємодії в iOS і Android, або внаслідок переваг платформозалежного API – то з цим не буде труднощів. React Native дозволяє призначати платформозалежні версії кожного компонента, які можна потім інтегрувати в проєктований додаток.

У світі кросс-платформенної мобільної розробки, як показано у [42], вже були свої рішення, наприклад Apache Cordova – технологія, яка дозволяє використовувати HTML + CSS + JavaScript + нативний функціонал платформи, на якій додаток було запущено, для його роботи. Однак, технологія має велику проблему – швидкодія. Так само на даний момент існують і інші, такі як Xamarin, Flutter, QT тощо. У React Native код пишеться на JavaScript, і здійснюється за допомогою JavaScriptCore – движка, який використовує Safari. Так само можна використовувати нативні модулі платформи, наприклад камеру або Bluetooth. Для цього пишеться код, який реалізує функціональність на мові, яка призначається для розробки під конкретну платформу (Java / Swift / Objective C) і взаємодіє із середовищем JavaScript за допомогою bridge.

Отже, резюмуючи вищесказане, наведемо переваги React:

- для написання коду використовується JavaScript, зокрема React;
- є можливість швидко написати додаток під обидві платформи. Менше витрат – вигідніше бізнесу;
- велика бібліотека нативних і не нативних компонентів;
- для налагодження можна використовувати браузер, так само є hot-reload для швидкого перегляду змін. Немає необхідності збирати заново додаток при внесенні змін до коду;
- для відтворення компонентів використовуються нативні компоненти використовуваної системи (наприклад UIImage / UIImageView), відповідно продуктивність такого UI вище, ніж при використанні webview;

- просто написати свою бібліотеку для RN, що використовує нативний функціонал системи;
- ще одна причиною чому RN набрав популярність в останні роки є те, що його використовують такі гіганти як Facebook, Instagram, Skype, Tesla, Baidu, Bloomberg і т. ін.

У React Native є свої недоліки. Це відносно новий проект, і йому властиві проблеми всіх молодих бібліотек: відсутні деякі функції, ще не розроблені оптимальні методики використання. Від релізу до релізу впроваджуються серйозні зміни, хоча їх небагато і вони носять обмежений характер.

Проте, React Native вже досить зрілий проект, чії переваги вагоміші за недоліки [43]. За допомогою цього фреймворка можна використовувати єдину кодову базу для створення додатків під iOS і Android, не жертвуючи ні якістю, ні продуктивністю.

Отже, React Native – це потужна платформа, яка використовується підприємствами будь-якого розміру для створення мобільних додатків. Це швидкий, ефективний і відносно простий спосіб для створення додатків на JavaScript. Головною його перевагою є те, що він дозволяє швидко створювати додаток під кілька платформ одночасно, а також має невисокий поріг входження. React Native широко використовується великими брендами, що є гарантією його живучості на ринку.

2.3 Засоби обробки інформації

Розпізнавання зображень – це задача вводу зображення до нейронної мережі та присвоєння цьому зображенню певної мітки. Мітка, яку назначає мережа, відповідає заздалегідь заданому класу об'єктів. При цьому у випадку, коли є лише один клас, то має місце «розпізнавання», а якщо декілька, то це вже задача розпізнавання кількох класів, так звана «класифікація». І вже підмножина класифікацій зображень є визначенням об'єктів.

Тобто розпізнавання зображень – це ідентифікація екземплярів об'єктів залежно від їхньої приналежності до певного класу (люди, дерева, дорожні знаки тощо). Наприклад, задача ReCaptcha від Google, де з набору різних картинок необхідно обрати лише ті, які відносяться до вказаного класу.

Для виконання процедури розпізнавання або класифікації зображень, нейронна мережа повинна отримати ознаки – елементи даних максимального інтересу, наприклад групи пікселів (точки та лінії). Саме вони і будуть передаватися на вихід нейромережі. Тобто мережа повинна проаналізувати зображення за допомогою груп пікселів (ознак) на наявність патерну.

Відповідні ознаки нейромережа знаходить з метаданих або анотаціях, які містяться в самих зображеннях. Тому процес розпізнавання ознак (або вилучення ознак) зводиться до процесу отримання відповідних ознак з вхідного зображення для подальшого їх аналізу. [44].

Як стверджує автор у [45], то ще кілька років тому програмування нейромережі полягало у ручному завданні всіх шарів, функцій активації та функції втрат, аналітичного розрахунку їх похідних, а також перевірки коректності цих похідних за допомогою чисельного диференціювання. Ця процедура була схильна до помилок і мала високий поріг входу.

В останні роки необхідності в ручному завданні похідних немає, оскільки більшість бібліотек підтримують один з двох варіантів автоматизації цього процесу:

- Символьне диференціювання (Theano)
- Автоматичне диференціювання (Tensorflow, Torch).

Класичні нейронні мережі прямого розповсюдження (Feed-Forward Neural Networks, FNN) – вид нейронної мережі, в якій сигнали поширюються в одному напрямку, починаючи від вхідного шару нейронів, через приховані шари до вихідного шару і на вихідних нейронах отримується результат опрацювання сигналу. В мережах такого виду немає зворотних зв'язків.

Використання нейромережі полягає у наступному:

- на вхід нейромережі для кожного об'єкта подається вектор ознак;

- на виході нейронної мережі виходить результат, при чому перегляд результату визначається архітектурою нейронної мережі.

Наприклад, у разі багатокласової класифікації (N класів) вихідний шар містить N нейронів, кожен з яких видає «ступінь належності» до цього класу (якщо активація була сигмоїдальною). Наявність більш одного прихованого шару дає можливість будувати ієрархічні уявлення (рисунок 2.12).

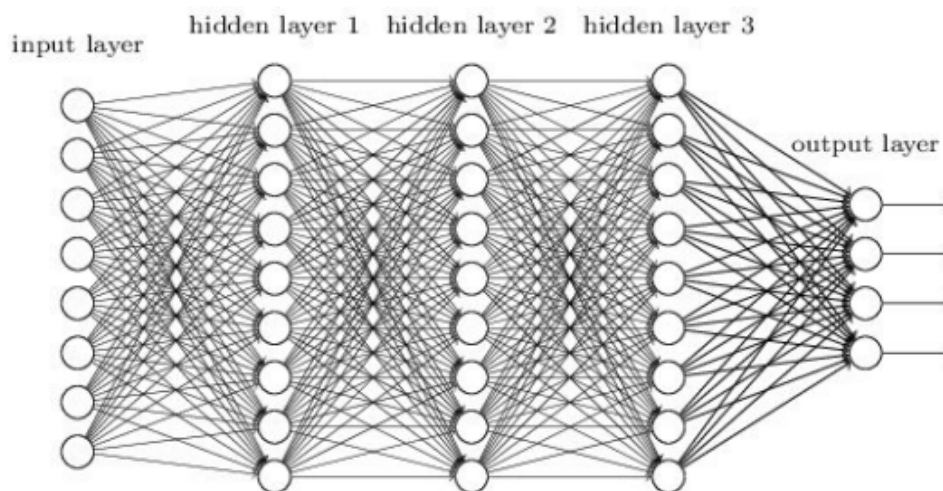


Рисунок 2.12 Нейронна мережа із трьома прихованими шарами

Класичні нейронні мережі прямого розповсюдження добре працюють для класифікації, але є деякі труднощі:

1. Багато параметрів – для мережі, біля якої на вході картинка 100×100 , три приховані шари по 100 нейронів кожен, і виходом на 10 класів, число параметрів буде приблизно 1 мільйон ($10000 \times 100 + 100 \times 100 + 100 \times 100 + 100 \times 10$);
2. Згасаючі градієнти (якщо багато шарів);
3. Такі мережі важко піддаються навчанню.

На даний момент можна виділити два підходи до розпізнавання емоцій людини, що базуються на використанні нейронних мереж.

Класичний підхід до вирішення задачі з розпізнавання емоцій, що ґрунтуються на класифікації точок людського обличчя. Розташування точок

фіксує жорсткі та нежорсткі деформації обличчя через рухи голови та міміки. Для отримання ключових точок обличчя можуть використовуватися такі алгоритми, як PDM, CML, AAM, DPM або CNN [46]. Наступний етап розпізнавання емоцій при класичному підході – це класифікація ключових точок.

На сьогоднішній день проблема пошуку ключових точок досить добре вивчена і є велика кількість алгоритмів, що дозволяють отримати ключові точки з точністю, достатньою для подальшої класифікації за цими точками емоцій людини. Але, для використання класичного підходу необхідно, щоб положення особи на зображенні було вирівняне [46].

Починаючи з 2011 року, передовою в мережах прямого поширення глибинного навчання була почерговість згорткових шарів та шарів максимізаційного агрегування, увінчаних декількома повно- або частково зв'язаними шарами, за якими йде рівень остаточної класифікації.

Тому альтернативою використанню класичного підходу є підхід, заснований на згорткових нейронних мережах.

Згорткова нейронна мережа (Convolutional Neural Networks, CNN) – це Feed-Forward мережа спеціального виду (рисунок 2.13).

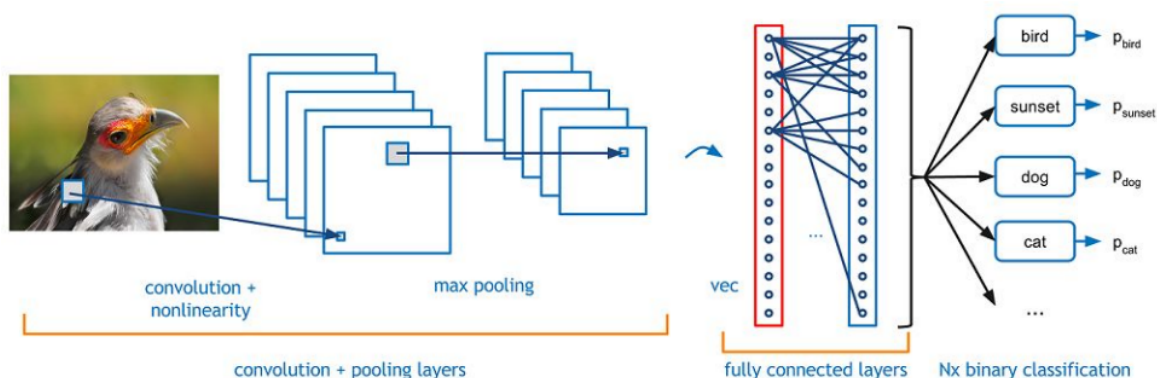


Рисунок 2.13 Згорткова нейронна мережа CNN [45]

Згорткова нейронна мережа – це архітектура штучних нейронних мереж, орієнтована на ефективне розпізнавання образів. Згорткові мережі є добрим базовим рішенням для класифікації різних візуальних даних (рисунок 2.14).

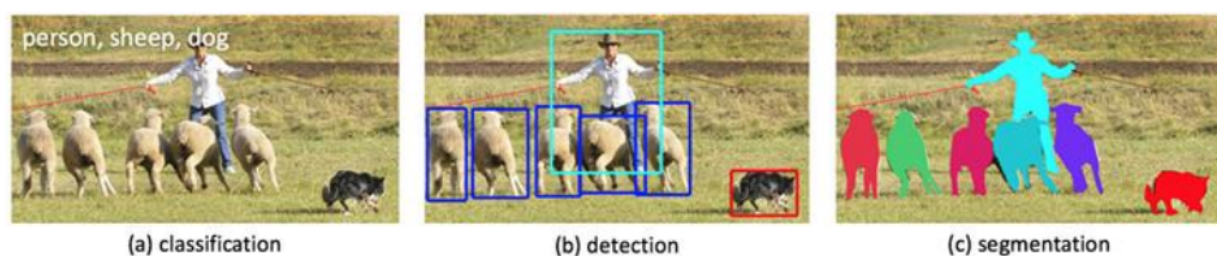


Рисунок 2.14 Класичні задачі для CNN [45]

Мережа CNN складається з наступних шарів [45]:

- Згорткові шари: кожна площина у згортковому шарі – це один нейрон, який реалізує операцію згортки (convolution) і є матричним фільтром невеликого розміру (наприклад, 5×5 або 3×3).
- Шари субдискретизації (subsampling, spatial pooling): зменшують розмір зображення (наприклад, у 2 рази).
- Повнозв'язні шари (MLP) на виході моделі (використовуються для класифікації).

Згорткові шари вчать ієрархічні ознаки для зображень, а spatial pooling дає деяку інваріантність до переміщень.

Якщо порівнювати CNN та FNN за параметрами, то CNN їх менше, ніж у FNN (таблиця 2.1)

Коротко розглянемо процес розпізнавання зображень нейронною мережею, який умовно можна розділити на три етапи:

- перший етап – це попередня фільтрація та підготовка зображення;
- другий етап – це логічна обробка результатів фільтрації;
- третій етап – це алгоритми прийняття рішень на основі логічної обробки.

Таблиця 2.1 Порівняння параметрів CNN та FNN [45]

CNN	FNN
Вхід: ч/б картинка 100×100	
Три згорткових шари по 100 площин кожний	Три прихованих шари по 100 нейронів кожний
Вихід: 10 класів	
Число параметрів приблизно 650 тис.: ($5 \times 5 \times 1 \times 100 + 5 \times 5 \times 100 \times 100 +$ $+ 5 \times 5 \times 100 \times 100 + 12 \times 12 \times 100 \times 10$)	Число параметрів приблизно 1 млн.: ($10\,000 \times 100 + 100 \times 100 + 100 \times 100 +$ $+ 100 \times 10$)

На першому етапі застосовуються методи, які дозволяють виділити на зображеннях області, що цікавлять, без їх аналізу. Більшість цих методів застосовує якесь єдине перетворення до всіх точок зображення. На перший шар нейронної мережі надходить вихідне зображення. Після того, як вихідні дані введені в мережу, до зображення застосовуються різні фільтри, що формують розуміння окремих частин зображення (рисунок 2.15).

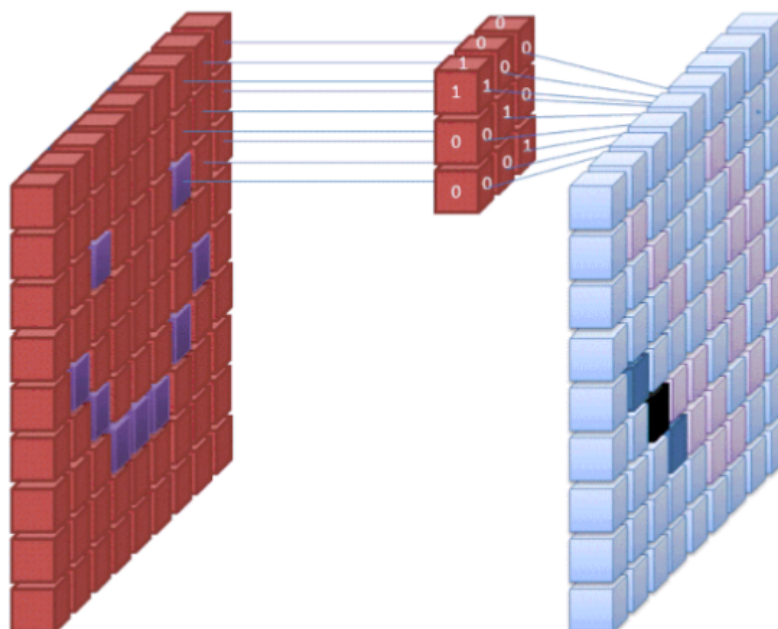


Рисунок 2.15 – Вилучення ознак за допомогою фільтрів

Хоча на цьому рівні не виконується аналіз зображення, проте точки, які проходять фільтрацію, можна розглядати як області з особливими характеристиками чи як вже говорилось вище – ознаками. Таким чином на цьому етапі відбувається отримання ознак об'єктів.

Фільтр – це колекція кернелів; іноді у фільтрі використовується один kernel. Кернел – це звичайна матриця чисел, що визначається вагами, які «навчаються» (підлаштовуються) з метою пошуку зображень конкретних характеристик.

Для кожної точки зображення вибирається вікно і перемножується з фільтром того самого розміру. Результатом такої згортки є нове значення точки (рисунок 2.16). Саме звідси і пішла назва «Згорткової нейронної мережі» – типу нейронної мережі, що в класифікації та розпізнаванні зображень використовується найчастіше. Таким чином мета операції згортки – витягти із вхідного зображення високорівневі ознаки.

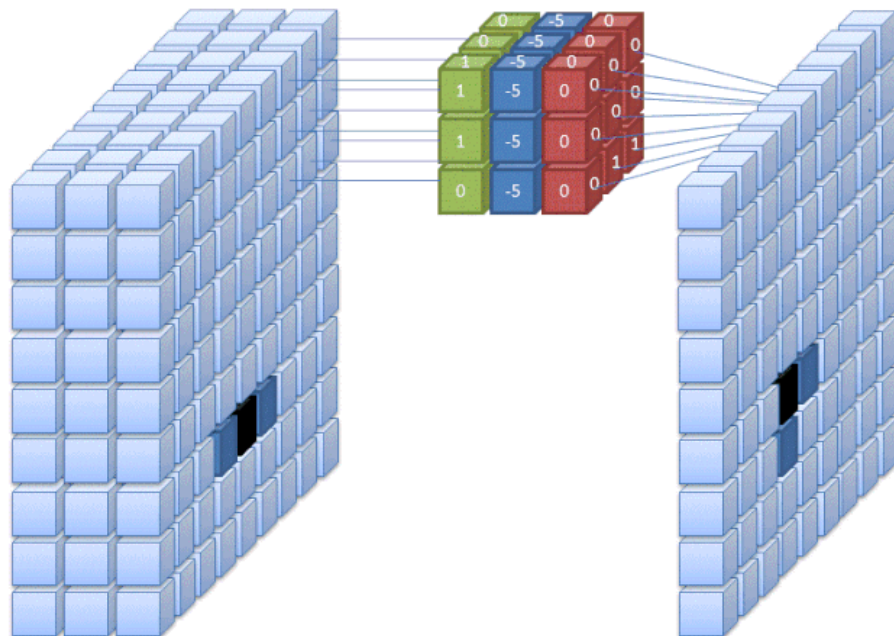


Рисунок 2.16 – Згорткова мережа із шарами ознак [44]

У згортковому шарі обчислюється результат скалярного добутку між певною областю вхідного зображення та ваговою матрицею (званою фільтром). Фільтр переміщуватиметься по всьому зображенню і повторюватиме ту саму операцію обчислення скалярного добутку. Результатом переміщення даного фільтра для зображення є матриця, що складається з результатів одиничних згорток.

Важливо, що кількість каналів фільтру має відповідати кількості каналів вихідного зображення – тільки тоді операція згортки буде мати необхідний ефект. Наприклад, якщо вихідне зображення складається з трьох каналів (RGB: червоний, зелений, синій), фільтр також повинен мати три канали.

Фільтр може переміщатися по матриці вхідних сигналів з кроком, що відрізняється від одиниць. Крок переміщення фільтра називається страйдом (крок). Страйд визначає, на скільки пікселів повинен поміщатися фільтр за один раз.

Найпростіші приклади фільтрів – це фільтр Гауса (реалізує підкреслення низьких частот) та фільтр Габора (підкреслення високих частот).

Якщо використовувати для згортки із сигналом якусь довільну характеристичну функцію, тоді це називатиметься «вейвлет-перетворення».

Це визначення вейвлетів не є зовсім коректним, але традиційно склалося, що у багатьох командах вейвлет-аналізом називається пошук довільного паттерна на зображенні за допомогою згортки з моделлю цього паттерна. Примітиви Хаара, про які раніше йшла мова (п. 2.1), належать до таких функцій для двовимірного простору. Прикладом використання розширеного трактування вейвлетів є задача пошуку відблиску в оці, для якої вейвлетом є сам відблиск (рисунок 2.17).

Окремим класом фільтрів є фільтрація функцій – математичні фільтри, які дозволяють виявити просту математичну функцію на зображенні (пряму, параболу, коло тощо). Будується акумулююче зображення, у якому до кожної точки вихідного зображення вимальовується безліч функцій, що її породжують. Найбільш класичним перетворенням є перетворення Хафа для прямих. Існує

модифіковане перетворення, яке дозволяє шукати будь-які фігури, але не завжди працює при обробці зображень внаслідок повільної швидкості роботи та високої чутливості до якості бінаризації.

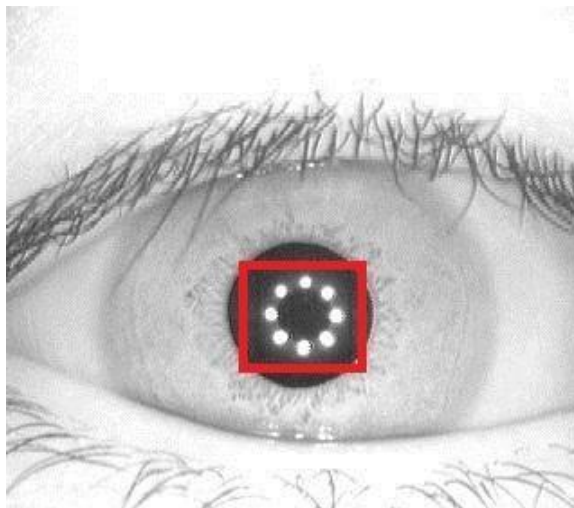


Рисунок 2.17 – Пошук відблиску в оці

Аналогом перетворення Хафа для прямих є перетворення Радону. Воно обчислюється через БПФ, що дає вигаш продуктивності в ситуації, коли точок дуже багато. До того ж його можна використовувати до небінаризованого зображення.

Ще один клас фільтрів – фільтрація меж та контурів, що використовується для вирішення задач із складними об'єктами. Тобто у випадку, коли сам об'єкт добре виділяється на зображенні, то задача зводиться виділення лише його контурів. Існує цілий ряд алгоритмів, що вирішують завдання фільтрації контурів:

- оператор Кенні;
- оператор Собеля;
- оператор Лапласа;
- оператор Прюїтт;
- оператор Робертса.

Найчастіше використовується оператор Кенні, реалізація якого є в OpenCV. Оператор Собіль також реалізовано в OpenCV, але він гірше шукає контури, ніж Кенні.

Вищенаведені фільтри та їх модифікації здатні вирішити 80-90% можливих завдань з розпізнавання зображень. Але крім них існують більш рідкісні фільтри: ітераційні фільтри (наприклад активна модель зовнішнього вигляду), а також ріджлет і курвлет перетворення, що є сплавом класичної вейвлет фільтрації та аналізом у полі радон-перетворення. Проте ці перетворення дуже специфічні та використовуються при вирішенні конкретних задач.

У згортковому шарі вага кожного вікна даних про підключення нейрона фіксована, і кожен нейрон чи фільтр при обробці зображень (наприклад, фільтр Собела для виявлення країв) звертає увагу лише на одну характеристику. Тобто кожен фільтр згорткового шару буде мати свою власну функцію зображення, таку як вертикальний край, горизонтальний край, колір, текстура і т. д. Сума всіх цих нейронів схожа на набір екстракторів ознак для всього зображення.

Фільтрування дає набір придатних для обробки даних. Але часто не можна просто взяти та використовувати ці дані без їхньої обробки. Тому після того, як формування переліку ознак зображення завершено, починається етап логічної обробки результатів фільтрації.

На цьому етапі значення, що представляють зображення передаються через функцію або шар активації. Оскільки дані, що формується з використанням операцій згортки, перебувають у лінійній формі, виникає недовіра, що дозволяє нейронній мережі чіткіше оцінювати ситуацію. Тому функція активації приймає вхідні значення і збільшує їх нелінійність. Це як носіння контактних лінз. Загалом це викликано тим, що вхідні дані за своєю природою нелінійні, тому потрібно навмисне шукати проміжні результати, щоб відповідь нейронної мережі була підходящою.

Часто в якості функції активації використовується функція ReLU – випрямлена лінійна одиниця. Також іноді використовуються інші функції, такі,

як: Sigmoid function, Softmax function, Softplus function, tanh function, SELU function, ELU function та exponential function [44].

Оскільки мережа повинна прийняти рішення відносно максимально важливих частин зображення, акцент йде на те, що вона вивчить тільки ті частини зображення, які дійсно є суть розглядуваного об'єкта, тобто частини із максимальною вагою. Це допомагає запобігти «перенавчанню» мережі – ситуація, коли мережа занадто добре вивчає всі аспекти навчального прикладу і вже не може узагальнювати нові дані, оскільки враховує нерелевантні відмінності.

Існує два найбільш часто використовуваних методи об'єднання – середній пул і максимальний пул, останній з яких використовується найчастіше. У згортковій нейронній мережі рівень об'єднання використовується для зменшення просторового виміру, але не зменшує глибину мережі. При використанні максимального рівня об'єднання беруться найбільш характерні точки у вхідній області (найчутливіша область на зображенні), а при використанні середнього рівня об'єднання – середнє значення вхідної області.

Операція максимального об'єднання полягає в тому, що регулювання обертів руху виконується так званим «вікном просіювання». З пікселів, що потрапляють у його поле зору, відображається маса і переміщується в матрицю результатів. Розмір для цього вікна може бути відмінним від одиничного – все залежить від того, якого розміру вихідну матрицю потрібно отримати, і з яким файлом потрібно «стиснути» дані.

Рисунок 2.18 ілюструє операцію максимального об'єднання: розмірність вікна 2×2 . Страйд має значення 2 – це крок, з яким рухається вікно за один етап. На кожну активність забирається максимальне значення. Вихідна матриця ознак просівається через сито, яка видає лише характерні пікселі. Завдяки максимальному об'єднанню зменшується кількість пікселів, що у свою чергу призводить до зменшення кількості операцій, що використовуються, і, відповідно, до економії обчислювальних ресурсів. У нашому прикладі це відсіює $3/4$ інформації.

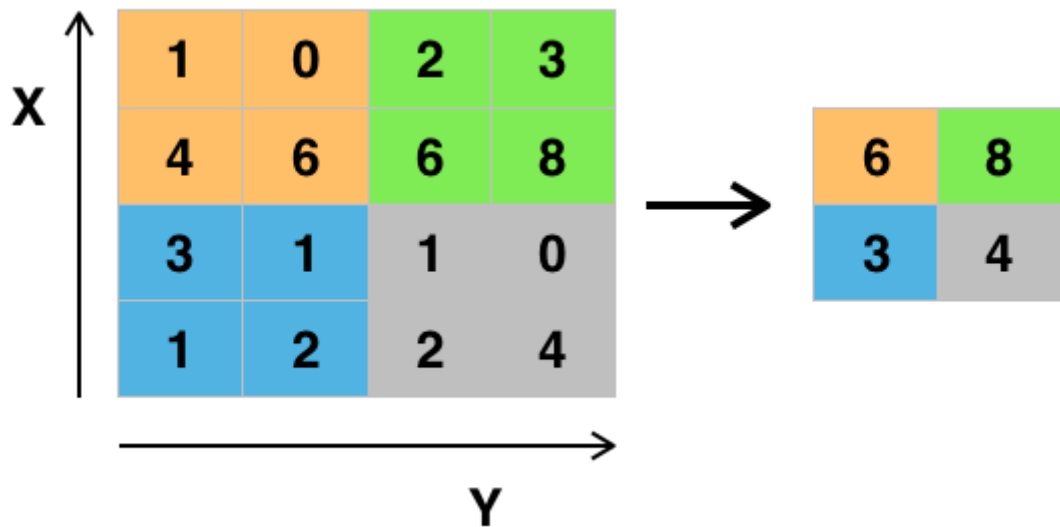


Рисунок 2.18 – Зріз максимальних значень пікселів [44]

Згортковий шар і пулінговий шар разом утворюють і-й шар згорткової нейронної мережі. Залежно від складності зображень кількість таких шарів може бути збільшено з метою більш точного захоплення деталей низького рівня, але ціною більшої обчислювальної потужності.

Виконання описаного вище процесу дозволяє моделі успішно досліджувати ознаки. Далі ми згладжуємо остаточний результат і передаємо його у звичайну нейронну мережу з метою класифікації.

На етапі прийняття рішень кінцеві вихідні шари CNN являють собою щільно пов'язані шари, які потребують представлення даних у формі вектора для можливості їх подальшої обробки. Для цього значення стискаються у довгий вектор або стовпець послідовно впорядкованих чисел.

У пов'язаному шарі згладжується виведення останнього згорткового шару і з'єднується кожен вузол поточного шару з вузлом наступного шару. Фактично повністю пов'язаний шар являє собою звичайну штучну нейронну мережу, як показано на рисунку нижче (рис. 2.19). Операція в повністю підключеному шарі така сама, як і в штучній нейронній мережі.

Fully connected neural network

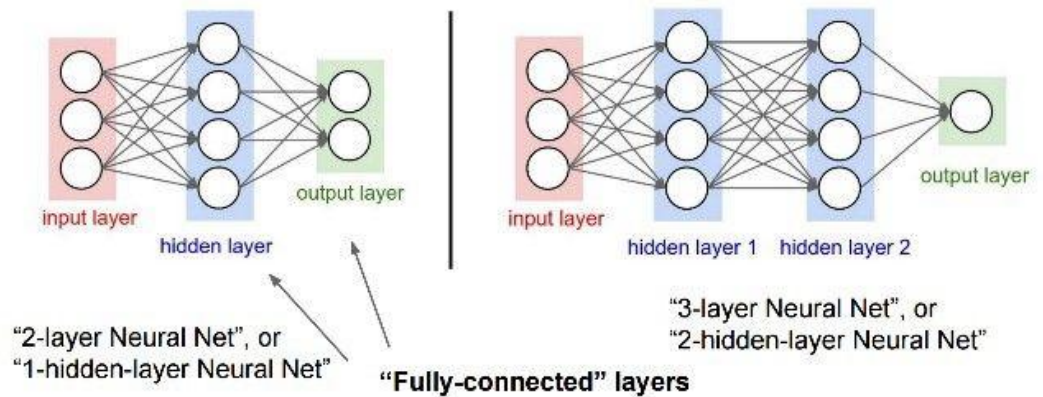


Рисунок 2.19 – Структура нейромережі CNN

Вихідні шари CNN ще називають штучною нейронною мережею (Artificial neural networks, ANN), основною функцією якої є аналіз вхідних ознак та об'єднання їх у різні атрибути, які допомагають у класифікації (рисунок 2.19). Ці шари утворюють набори нейронів, які представляють різні частини розглядуваного об'єкта, а набір нейронів може являти собою якийсь аспект (наприклад, людина з пістолетом у руках або жовтий дзьоб птаха). Коли достатня кількість цих нейронів активується у відповідь на вхідне зображення, то воно буде класифіковано як об'єкт.

Нейрони в середніх пов'язаних шарах будуть виводити двійкові значення, що відносяться до можливих класів. Якщо у наявності є чотири різні класи (скажімо, людина, автобус, яблуко та будівля), нейрон матиме значення «1» для класу, до якого, як він вважає, відноситься зображення, і значення «0» для інших класів.

Підхід, заснований на використанні машинного навчання, може бути поділений на дві основні категорії: статичні та динамічні методи. У статичних методах в якості інформації для розпізнавання використовується одиничне зображення людського обличчя. Тоді як динамічних методах використовується

послідовність зображень. Крім інформації про кожне зображення в послідовності, динамічні методи враховують часовий зв'язок між суміжними зображеннями [46].

Вищеназвані етапи дуже умовні, бо для вирішення завдання розпізнавання зображень нейронною мережею далеко не завжди потрібно застосовувати методи з кожного етапу, буває достатньо двох, а іноді навіть одного.

2.4 Висновки до другого розділу

Досліджено існуючі алгоритми розпізнавання емоцій людей на фото. Визначено особливості функціоналу для проектованої мережі.

Запропоновано алгоритмічне та програмне забезпечення сервісу прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото та обґрунтовано його ефективність у порівнянні з існуючими рішеннями.

За результатами порівняльного аналізу методів було обрано два алгоритми: метод Віоли-Джонса спільно з локальними бінарними шаблонами та згорткова нейронна мережа.

Обрано інструменти для розробки і проектування. У якості засобів розробки обрано відкриті й безкоштовні інструменти, підтримувані світовою спільнотою.

РОЗДІЛ 3. РЕАЛІЗАЦІЯ МЕТОДІВ ОБРОБКИ ІНФОРМАЦІЇ НА ОСНОВІ РЕАЛЬНИХ ДАНИХ.

3.1 Візуалізація вхідних даних проекту обробки інформації

3.1.1 Ідентифікація архетипу ІС

Проектований додаток можна охарактеризувати як Mobile Application (МА). Мобільний додаток розроблено під одну з платформ Android / iOS. Додаток виконується на пристроях з обмеженими апаратними ресурсами.

3.1.2 Інтерфейс користувача (UI View)

У процесі проектування користувацького інтерфейсу було розроблено макети вікон для візуалізації інтерфейсу Користувача – просте схематичне зображення екранів та інтерфейсу додатку (рисунок 3.1). Це чорно-білі каркасні малюнки, що дають лише загальне уявлення про майбутній дизайн додатку.



Рисунок 3.1 Візуальний інтерфейс користувача

Високоякісні UX прототипи (рис. 3.2). Показують усі візуальні і типографічні деталі дизайну створюваного додатку. При їх створенні враховуються фізичні розміри екранів пристроїв користувачів, щоб візуально зрозуміти, як додаток буде виглядати на цих пристроях.

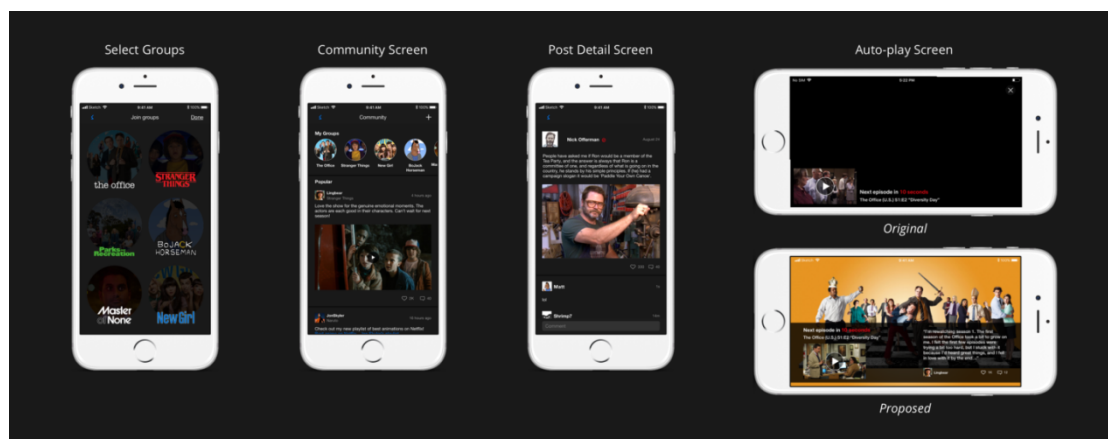


Рисунок 3.2. Високоякісні прототипи дизайну UX

Для більш детального розуміння структури переходів між вікнами інтерфейсу було спроектовано діаграму переходів (рисунок 3.3):

1. Головний екран додатку Application Window. При натисканні на кнопку «Take a shot» з'являється вікно вибору. Тут можливі два варіанти:
 - якщо Користувач обирає «Take a photo», то відкривається камера;
 - якщо «Choose from library» – відкривається галерея.
2. Коли користувач обирає/робить фото, на бекенд відправляється запит та відображається вікно завантаження (Camera/Gallery Window).

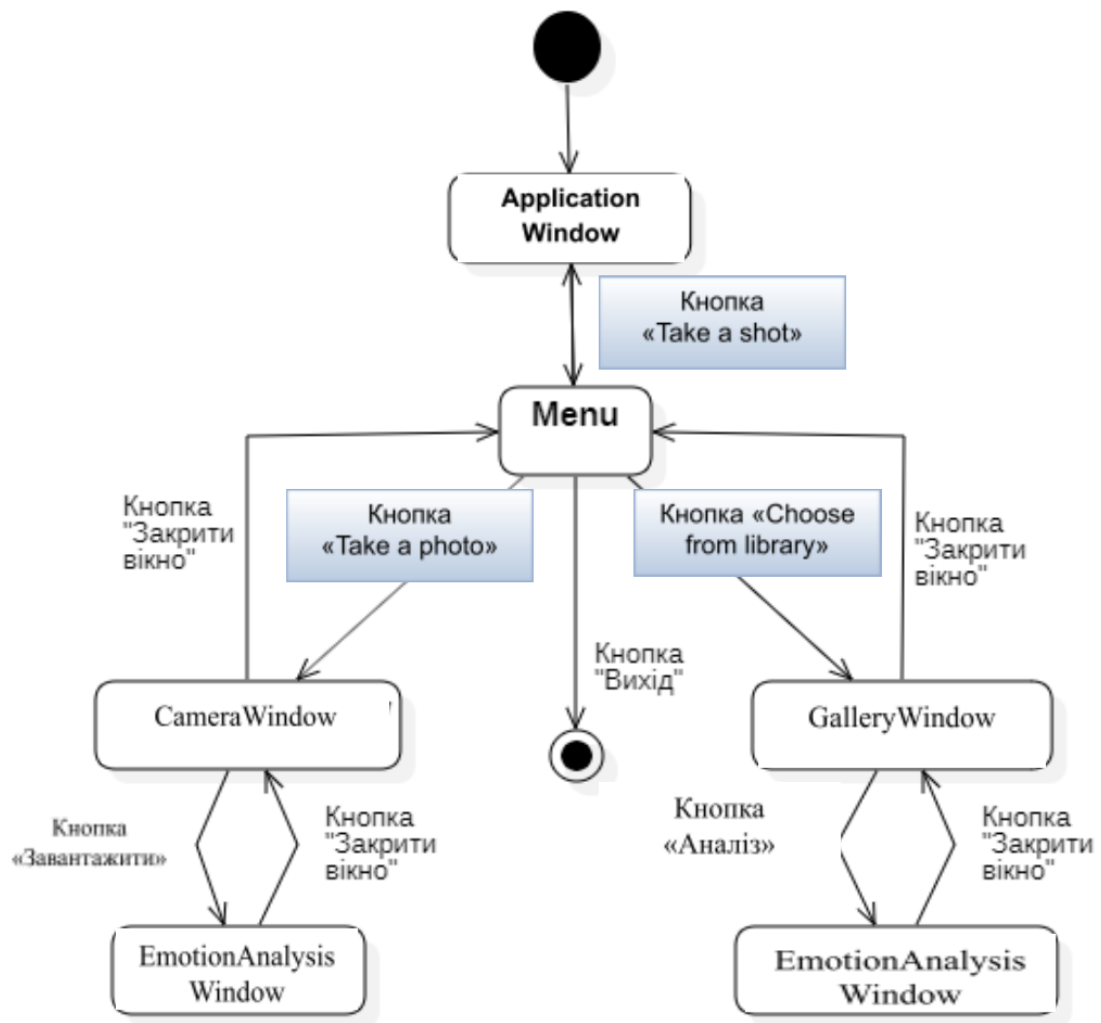


Рисунок 3.3 – Діаграма переходів між вікнами

3.1.3 Логічне уявлення ІС (Logical View)

Так як система побудована на основі React, який дозволяє легко створювати інтерфейси, розділяючи кожен сторінку на невеликі фрагменти і компоненти), то й найбільшу роль у логічному уявленні займає саме він: відображає інтерфейс, пов'язуючи дії користувача і відповідні пункти меню, завантажує екрани та об'єкти додатку.

На рисунку 3.4 наведено логічне уявлення проектованої інформаційної системи.

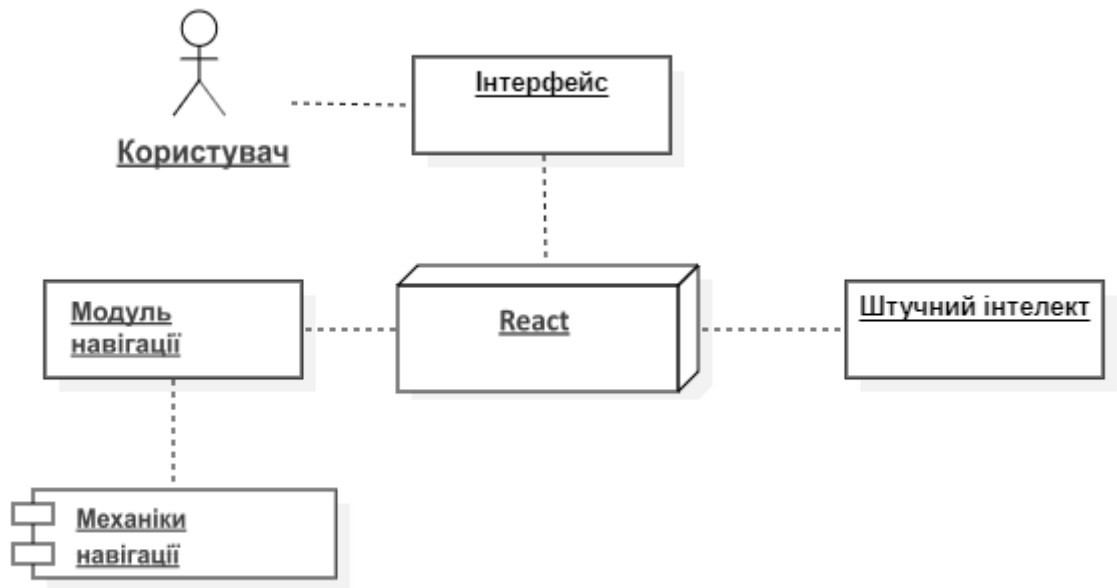


Рисунок 3.4 – Логічне уявлення ІС

Модуль навігації пов’язує дії користувача та модуль, що відповідає за механіки навігації додатком.

Модуль штучного інтелекту додатку містить алгоритми, за якими контролюється прогнозування фільмографічних уподобань користувача по аналізі емоцій з селфі фото.

3.1.4 Архітектура БД

Для реалізації серверної частини створюємо таблиці: *emotions*, *movies*.

Таблиця *emotions* – призначена для збереження інформації про існуючі емоції які розпізнає додаток.

Таблиця 3.1 – Структура таблиці БД *emotions*

Поле	Тип	Призначення
id	integer	Ідентифікатор емоції
title	string	Назва емоції

Таблиця *movies* – призначена для збереження інформації про існуючі фільми серед яких буде відбуватися пошук підходящих для користувача.

Таблиця 3.2 – Структура таблиці БД *movies*

Поле	Тип	Призначення
id	integer	Ідентифікатор фільму
tmdb_id	integer	Ідентифікатор фільму у базі TMDb
title	string	Назва фільму
release_date	date	Дата релізу фільму
imdb_rating	string	Рейтинг фільму по IMDb
poster_path	string	Посилання на картину постеру
original_language	string	Мова фільму
genre_ids	array	Масив жанрів фільму
imdb_id	string	Ідентифікатор фільму у базі IMDb
runtime	integer	Тривалість фільму у хвиликах
trailer	string	Посилання на трейлер фільму у YouTube
popularity	integer	Оцінка популярності фільму

3.1.5 Опис стеку технологій

Для написання програмного коду додатку використовується мова програмування PHP.

У якості середовища розробки (development environment) було обрано: Visual Studio IDE.

Основні технології розробки:

- PHP // server side – серверна частина додатку;
- React Native // client side – клієнта частина додатку.

Додаткові допоміжні бібліотеки:

- "@react-native-community/async-storage": "^1.12.1";

- "@react-native-firebase/app": "^11.5.0";
- "@react-navigation/native": "^5.9.4";
- "@react-navigation/stack": "^5.14.4";
- "prop-types": "^15.7.2";
- "react": "17.0.2";
- "react-native": "0.64.0";
- "react-native-circular-progress": "^1.3.7";
- "react-native-gesture-handler": "^1.10.3";
- "react-native-linear-gradient": "^2.5.6";
- "react-native-push-notification": "^7.3.1";
- "react-native-reanimated": "^2.1.0";
- "react-native-safe-area-context": "^3.2.0";
- "react-native-screens": "^3.1.1";
- "react-native-svg": "^12.1.1";
- "react-native-svg-transformer": "^0.14.3";
- "react-native-vector-icons": "^8.0.0";
- "react-native-video": "^5.1.1";
- "@babel/core": "^7.13.16";
- "@babel/runtime": "^7.13.17";
- "@react-native-community/eslint-config": "^2.0.0";
- "babel-jest": "^26.6.3";
- "eslint": "^7.25.0";
- "jest": "^26.6.3";
- "metro-react-native-babel-preset": "^0.66.0";
- "react-test-renderer": "17.0.2".

3.2 Представлення поетапної обробки даних засобами обробки інформації

Для визначення обличчя в реальному часі та зчитування/класифікації емоцій у проєктованому сервісі використовується набір даних fer2013 із моделлю keras CNN та openCV.

Keras – це високорівневий інтерфейс прикладного програмування (API-інтерфейс), який забезпечує взаємодію зі штучними нейронними мережами. Керівні принципи розробки Keras як бібліотеки з відкритим кодом – це розширюваність, зручність та модульність [47]:

1. Розширюваність – це можливість легко додавати нові класи, модулі та функції, що робить Keras відмінним засобом для проведення різноманітних досліджень.
2. Keras розроблялась для користувачів, а не для машин з використанням особливих методів: пропонує узгоджений та простий API, мінімізує кількість дій користувача, необхідних для вирішення найпоширеніших завдань. У разі виникнення помилок завжди можна звернутися на підтримку зворотного зв'язку.
3. Під модульністю розуміють послідовність або граф автономних, повністю налаштованих модулів, які можуть бути підключені без додаткових обмежень. Наприклад, це можуть бути нейронні шари, функції помилок, оптимізатори, схеми ініціалізації, функції активації та схеми регулювання – всі ці модулі можна комбінувати для створення моделі.

Тобто створювалась Keras для того, щоб максимально спростити процес створення нейронних мереж.

В основі Keras лежать моделі глибокого навчання, основний тип яких – послідовність, що представляє собою лінійний стек шарів.

Суть полягає у наступному: створюється послідовність і до неї додаються шари у тому порядку, у якому потрібно виконати обчислення. Після визначення

компілюється модель, що використовує базову платформу оптимізації обчислень.

Далі модель повинна відповідати даним – це можна зробити по одній партії даних за один раз або запусивши весь режим навчання моделі. Коли навчання буде завершено, можна використовувати модель для прогнозування нових даних.

Також є ще один тип моделей – це клас `Model`, який використовується з функціональним API.

Завдяки Keras модель або частину коду можна використовувати повторно та розширювати в майбутньому, або зберегти модель та тренувати її пізніше. Це робиться так: `model.save('model.h5')`.

Завдяки широкій сумісності Keras без перешкод працює на MacOS, Windows та Linux [47].

Keras є надбудовою над TensorFlow – платформи, що найбільш використовується для розробки алгоритмів глибокого навчання. Тому на практиці Keras дозволяє максимально просто реалізувати безліч функцій складного у освоєнні фреймворку TensorFlow.

TensorFlow – це відкрита бібліотека для машинного навчання від команди Google Brain, що в даний час є найвідомішою бібліотекою глибокого навчання у світі.

TensorFlow-бібліотека розроблена для побудови та тренування нейронних мереж. TensorFlow чудово підходить для автоматичного знаходження та класифікації образів, оскільки якість розпізнавання наближається до людського сприйняття.

TensorFlow може працювати у паралельному режимі на декількох процесорах: як CPU, і GPU. Для обчислень загального призначення графічних процесорах використовується CUDA. Це забезпечує високу швидкість навчання та роботи навчених моделей. [44].

OpenCV (Open Source Computer Vision Library) – це відкрита бібліотека для роботи з алгоритмами комп'ютерного зору, машинним навчанням та

обробкою зображень. OpenCV написана мовою високого рівня C/C++, але існує також Python, JavaScript, Ruby та інших мов програмування. Працює на Windows, Linux та MacOS, iOS та Android.

OpenCV може використовуватися скрізь, де потрібний комп'ютерний зір. Ця галузь IT працює з технологіями, які дозволяють пристрою «побачити», розпізнати та описати зображення. Комп'ютерний зір дає точну інформацію про те, що зображено на картинці, з описом, характеристиками та розмірами (з певним ступенем достовірності).

Також бібліотека працює з машинним навчанням – галуззю, яка навчає алгоритми діяти тим чи іншим чином [48].

OpenCV містить алгоритми для: інтерпретації зображень, калібрування камери за еталоном, усунення оптичних спотворень, визначення подібності, аналіз переміщення об'єкта, визначення форми об'єкта та стеження за об'єктом, 3D-реконструкція, сегментація об'єкта, розпізнавання жестів тощо.

Ця бібліотека дуже популярна завдяки своїй відкритості та можливості безкоштовно використовувати як у навчальних, так і комерційних цілях.

Фактично OpenCV – це набір типів даних, функцій та класів для обробки зображень алгоритмами комп'ютерного зору [48].

Для розпізнавання виразів обличчя був обраний розширений набір даних FER2013 (Facial Expression Dataset) з урахуванням того, що загалом, коли готується інформаційний збір більших обсягів в нейронних мережах, особливо в глибоких нейронних мережах, як правило, відмічається тенденція досягати кращих результатів

Набір даних fer2013 був представлений у 2013 році ICML Challenges in Representation Learning і являє собою суміш 35 000 фотографій осіб з сімома найважливішими артикуляціями, на кшталт – огида, здивування, злість, щастя, страх, сум і нейтральна [49].

Дані складаються із зображень осіб у градаціях сірого розміром 48×48 пікселів. Особи були автоматично зареєстровані, так що особа знаходиться

більш-менш по центру та займає приблизно однакову кількість місця на кожному зображенні.

Завдання полягає в тому, щоб класифікувати кожну особу на основі емоцій, що відображаються у вираженні особи, в одну з семи категорій (0 = агресія, 1 = огида, 2 = страх, 3 = радість, 4 = смуток, 5 = подив, 6 = нейтрально).

В нашій роботі ми розглядали такі сім важливих почуттів обличчя, як: 0, 1, 2, 3, 4, 5, 6.

Приклад семи виразів наведено на рисунку 3.5.



Рисунок 3.5 – Сім важливих почуттів обличчя [49]

Рисунок 3.6 демонструє дисперсію кожного виразу в наборі даних FER2013 [49].

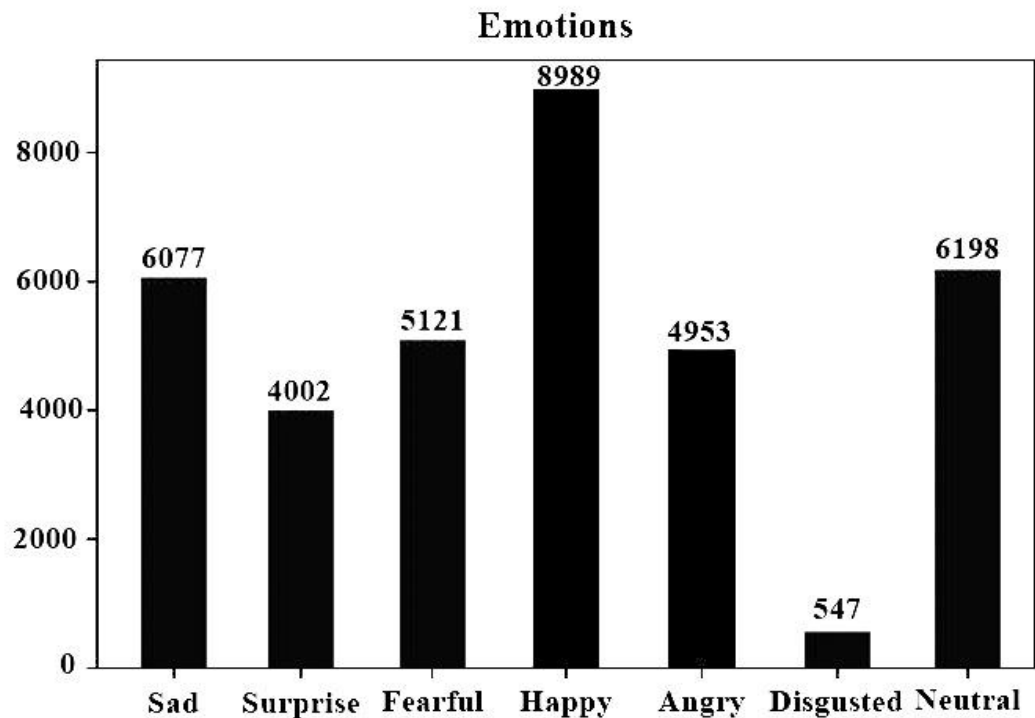


Рисунок 3.6 – Розподіл виразів обличчя в FER2013

Навчальний набір складається з 28709 прикладів. Публічний тестовий набір, який використовується для таблиці лідерів, складається з 3589 прикладів. Підсумковий тестовий набір, за яким визначається переможець конкурсу, складається з 3589 прикладів. Точність тесту класифікації емоцій fer2013 складає 66%.

Спочатку розпізнаються обличчя, використовуючи opencv, після цього використовується dlib для вилучення орієнтирів обличчя. Також використовуються функції HOG (Histogram of Oriented Gradients – Гістограма орієнтованого градієнта) – це дескриптор функції, який використовується для виявлення об'єктів у комп'ютерному зору та обробці зображень. Він складає елементи шляхом обчислення та підрахунку гістограми напрямку градієнта в локальній області зображення [50]. Схема алгоритму наведено на рисунку 3.7.

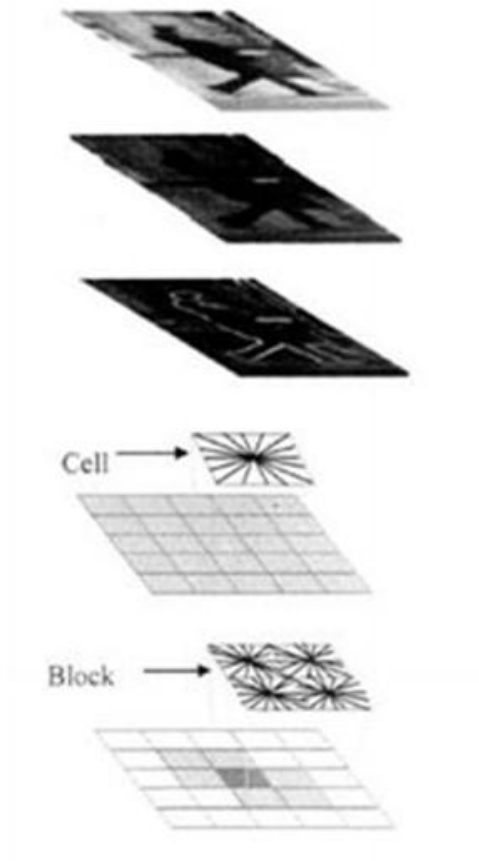


Рисунок 3.7 – Схема алгоритму вилучення функції HOG [50]

Метод вилучення функції HOG полягає в тому, щоб зробити зображення (ціль або вікно сканування, яке необхідно виявити):

1. Grayscale (зображення розглядається як тривимірне зображення x, y, z (відтінки сірого));
2. Використовується метод коригування гама для стандартизації колірного простору вхідного зображення (нормалізація); мета полягає в тому, щоб відрегулювати контраст зображення, зменшити вплив локальних тіней та змін світла у зображенні, а також придушити шумові перешкоди;
3. Розраховується градієнт кожного пікселя зображення (включаючи розмір та напрямок); в основному для захоплення інформації про контури, при подальшому ослабленні інтерференції світла.
4. Поділ зображення на дрібні осередки (наприклад, 6×6 пікселів/осередок);

5. Підрахунок гистограми градієнта кожного осередку (кількість різних градієнтів), щоб сформувати дескриптор кожного осередку;
6. Об'єднання кількох осередків у блок (наприклад, 3×3 осередки/блок), та об'єднання дескрипторів ознак усіх осередків у блоці, щоб отримати дескриптор ознак HOG блоку. Тобто дескриптор блоку (вектор) – дескриптор HOG.
7. Підключення дескрипторів функції HOG всіх блоків зображення, щоб отримати дескриптор функції HOG зображення (мета, яку необхідно виявити). Це останній крок – зібрати всі блоки, що перекриваються, у вікні виявлення для об'єктів HOG і об'єднати їх в остаточний вектор ознак для класифікації.

Далі вводяться необроблені вихідні дані зображення разом із орієнтирами обличчя та HOG у згорткову нейронну мережу, забезпечивши безпосередньо шари Fully Connected (FC).

Набір даних fer2013 доступний на веб-сайті Kaggle [51].

Для набору даних fer2013 реалізовано SPI (Strictly Person Independent) протокол, що дозволяє оцінювати алгоритми розпізнавання людських емоцій. Метриками для вимірювання продуктивності систем розпізнавання емоцій є точність, повнота та специфічність.

Оскільки, завдання полягає в тому, щоб підібрати фільми під настрій користувача, для цього потрібно мати широкий вибір фільмів у вільному доступі з повною інформацією про фільм. Найбільш великою бібліотекою фільмів на даний час являється IMDb [52]. Але IMDb не надає доступ до відкритого API аби отримати список всіх фільмів, тому потрібно використати парсинг сайту.

Фрагмент коду програми, що виконує парсинг інформації з сайту за результатами аналізу емоції обличчя з селфі-фото наведено в додатку Б, візуалізація проєктованого додатку – у п. 3.3. Також наведено результати роботи мережі (п. 3.3).

Найважливішим кроком програми є підбір фільмів під проаналізований настрій користувача. Після отримання результату емоцій на фото користувача відбувається розрахунок підходящих жанрів фільму для відповідного емоційного стану. Якщо переважає емоція радості це показник того, що користувачеві наразі можна запропонувати фільм з жанру комедії, якщо переважає емоція суму, то можна запропонувати фільм з жанру роману, тощо. Після визначення жанру фільму відбувається фільтрація списку фільмів по жанру та популярності фільму. Остаточний результат фільтрації фільмів і є результатом підбору фільмів.

3.3 Візуалізація результату проекту

Користувацька частина розробленого додатку прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото, що має назву «MyFilm», наведено на рисунку 3.10. Саме так виглядає головний екран додатку який дає змогу завантажувати фотографії обличчя та отримати список рекомендованих фільмів для перегляду під заданий настрій. На головну екрані присутня лише одна кнопка “Take a shot”, яка дає для користувача як саме він бажає завантажити фото. Додаток підтримує два варіанти вибору фото:

- фото з галереї;
- зробити фото.

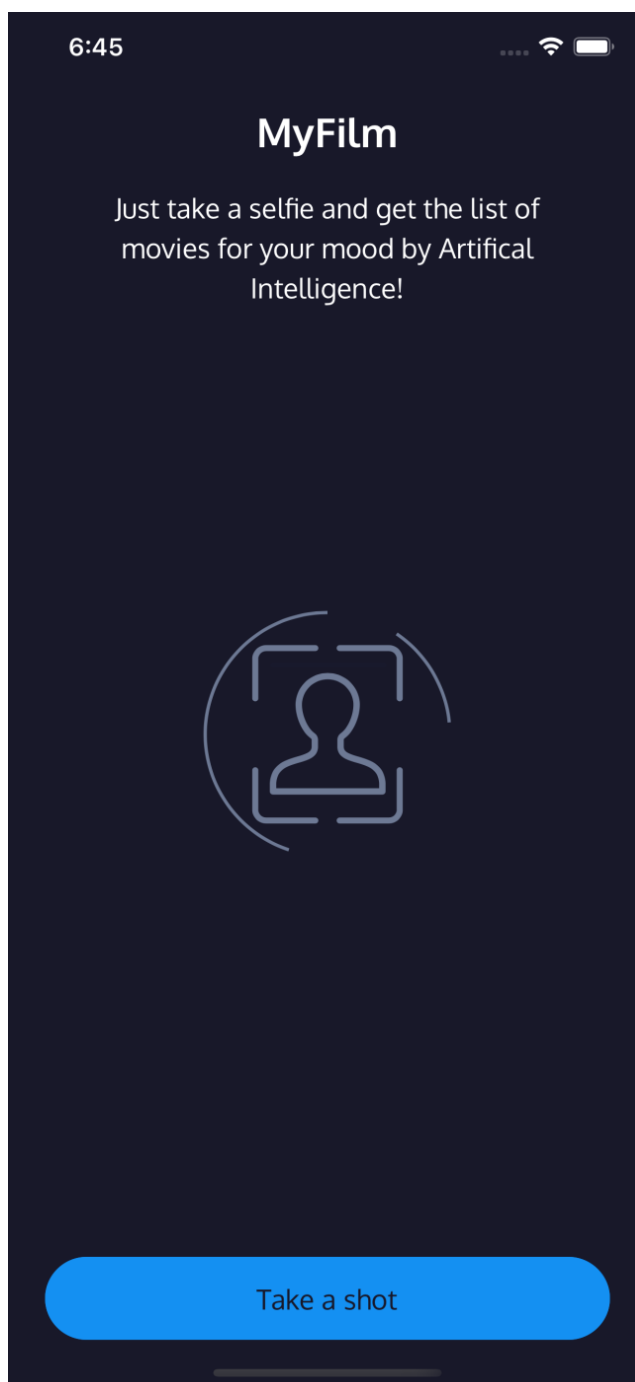


Рисунок 3.10 – Головний екран додатку

На рисунку 3.11 зображено вікно вибору способу завантаження фото у додаток при натисканні на кнопку «Take a shot».

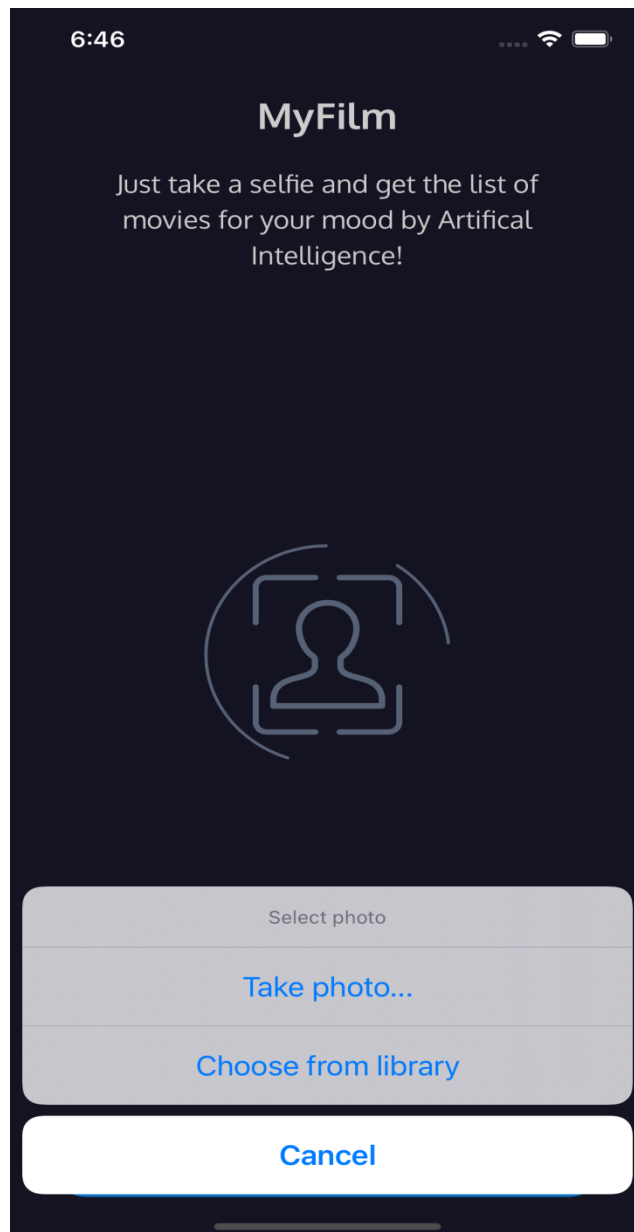


Рис. 3.11 Вікно вибору способу завантаження фото у додаток

При натисканні кнопки «Choose from library» користувачу надається доступ до галереї телефону, де він вибирає бажану фотографію. Після вибору фото на бекенд відправляється запит та відображається вікно завантаження фото (рисунок 3.12).

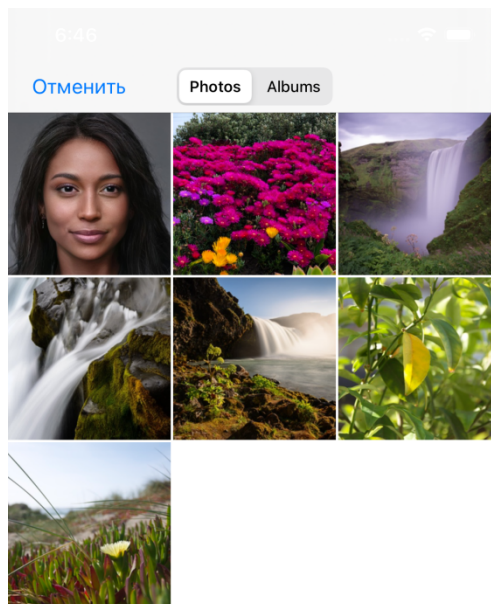


Рисунок 3.12 Вікно завантаження фото з галереї

Після завантаження фото підлягає аналізу (рисунок 3.13). Результатом аналізу є класифікація емоції та створення списку рекомендованих до перегляду фільмів у відповідності до отриманого результату класифікації емоції.

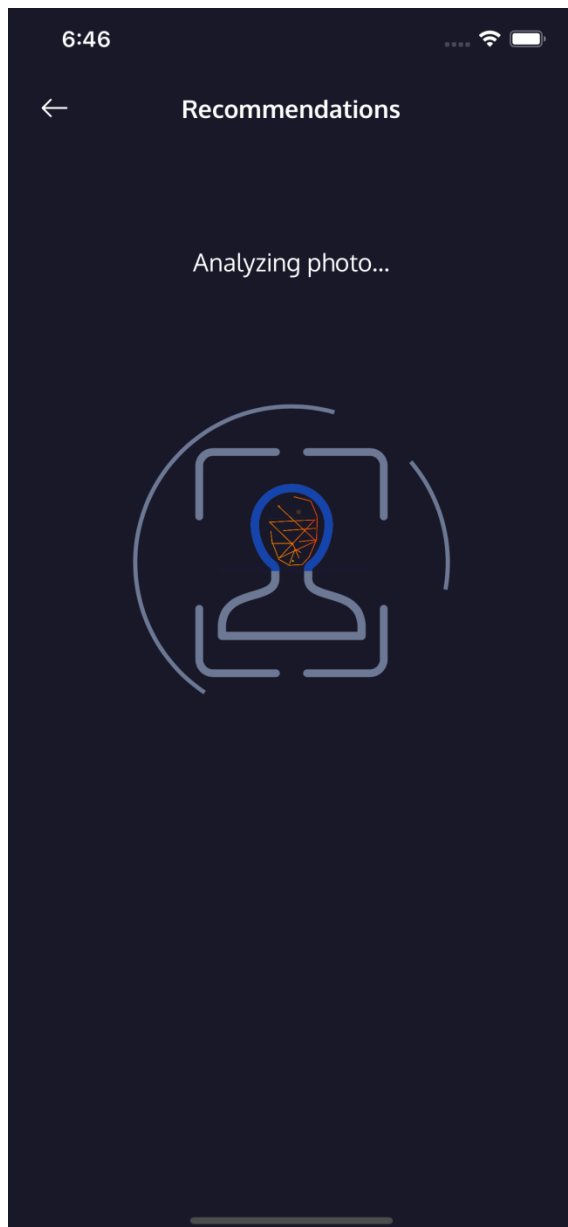


Рисунок 3.13 Аналіз фото

Після цього на екрані відображається вікно результатів у вигляді списку фільмів, які відповідають емоції на завантаженому фото з попереднього етапу. Список відображається у вигляді піктограми постеру фільму (рисунок 3.14).

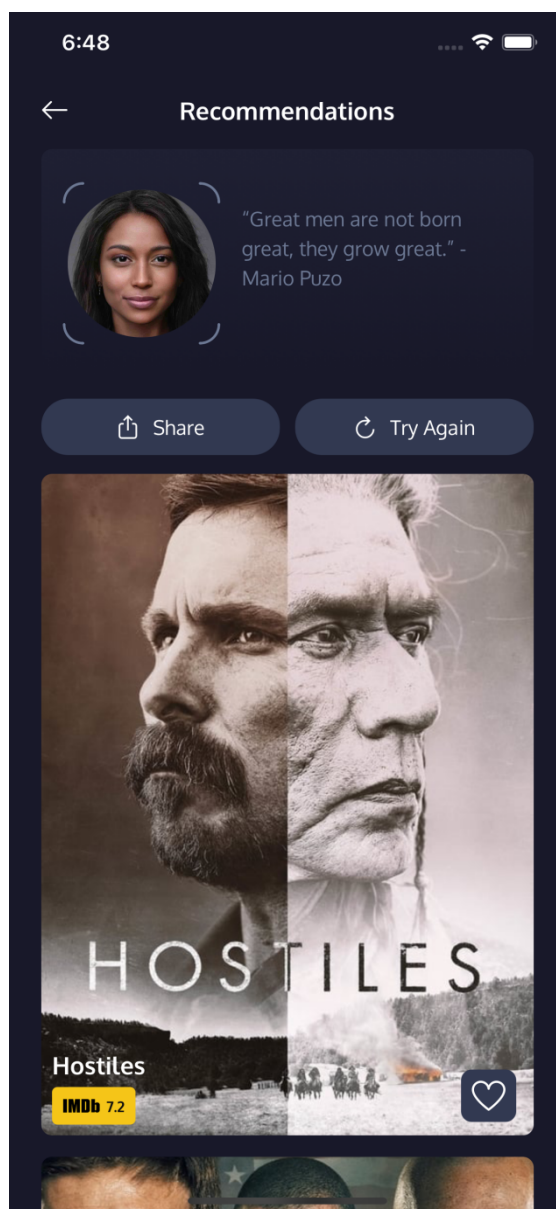


Рисунок 3.14 Перелік рекомендацій

При виборі Користувачем конкретного фільму відбувається перехід до екрану додатку із повною інформацією про фільм – акторський склад, рік виходу на екрани, прізвище режисера тощо. Відповідне вікно додатку зображено на рисунку 3.15.

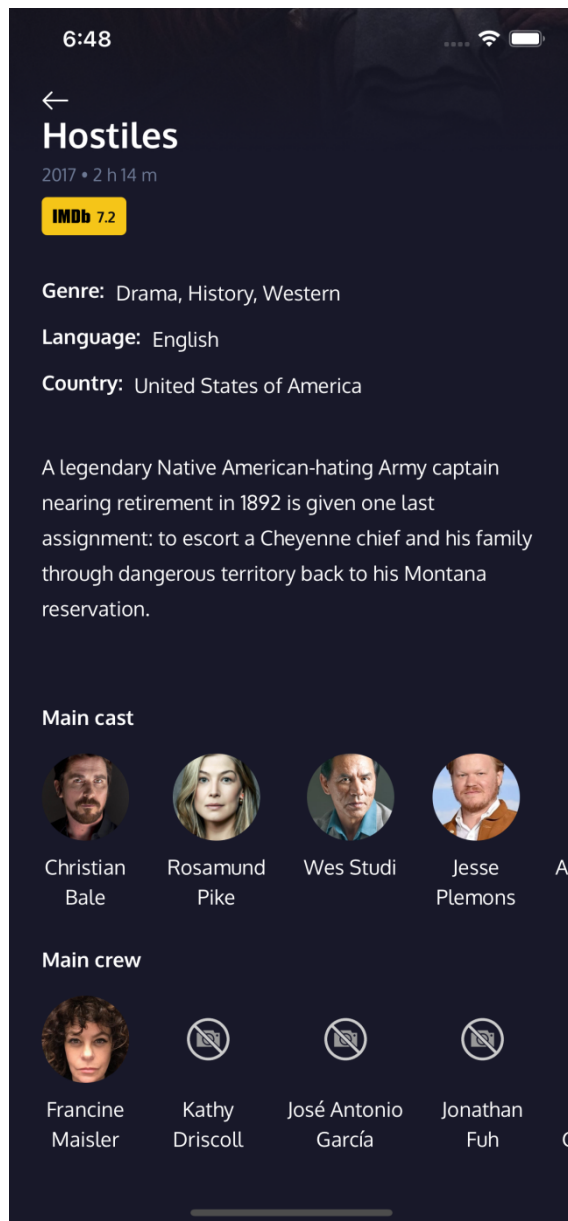


Рисунок 3.15 Інформація про фільм

Якщо Користувач обирає «переглянути трейлер», то відкривається вікно перегляду у YouTube. Це відображено на рисунку 3.16.



Рисунок 3.16 Вікно додатку із можливістю переглянути трейлер

Отже, проектована модель працює за описаним вище алгоритмом у розділі 2, п. 2.1) за схемою, наведеною на рисунку 3.17.

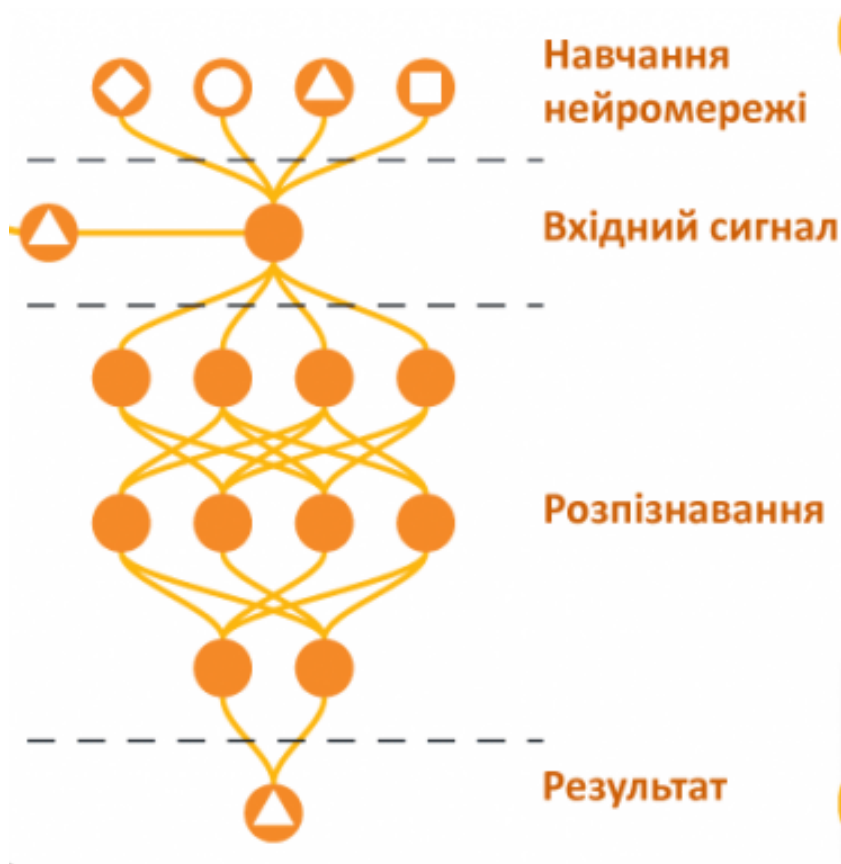


Рисунок 3.17 Схеми роботи мережі

Вихідні дані для моделі наведено на рисунку 3.18.



Рисунок 3.18 – Вихідні дані для проектованої мережі

Під час аналізу даної фотографії програма видає ваги (сумарно) кожної з можливих емоцій – neutral | happy | sad | angry | fearful | disgusted | surprised.

Результат аналізу фотографії наведено на рисунку 3.19.

```
{"neutral":0.610073447227478,  
"happy":0.38992342352867126,  
"sad":1.5906839223589486e-7,  
"angry":9.211696294642024e-8,  
"fearful":4.781164006373295e-11,  
"disgusted":3.6275160653076455e-10,  
"surprised":3.0405296911339974e-6}
```

Рисунок 3.19 – Результат аналізу фото

Відповідь по фото з рисунку 3.18 означає, що на обличчі 61% емоції типу neutral и 39% happy.

Для такої емоції програма рекомендує до перегляду певний список фільмів. Нижче наведено результат роботи мережі (рисунок 3.20).

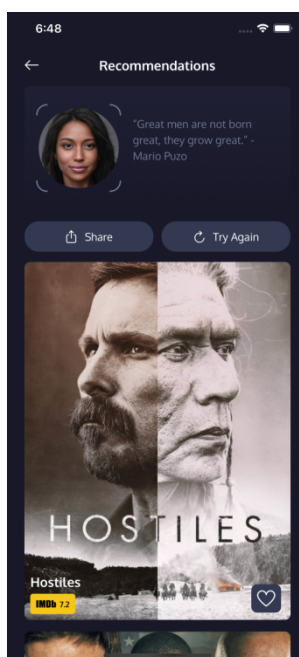


Рисунок 3.20 – Результат роботи мережі

Отримані результати дозволяють зробити висновок, що мету навчання досягнуто.

3.4 Висновки до третього розділу

1. Запропоновано алгоритмічне та програмне забезпечення для реалізації мобільного додатку прогнозування фільмографічних уподобань користувача по аналізі емоцій з селфі фото;
2. Розроблено дизайн, інтерфейс та макет додатку згідно визначених особливостей та вимог;
3. Реалізовано функціонал додатку у програмному середовищі;
4. Проведено експериментальне дослідження моделі прогнозування фільмографічних уподобань користувача по аналізі емоцій з селфі фото та виконано тестування функціоналу проектованої мережі.

ВИСНОВКИ

Дану роботу присвячено актуальній темі розробки програми для прогнозування фільмографічних уподобань користувача по аналізу емоцій з селфі фото.

Результати роботи такі:

1. Досліджено існуючі алгоритми визначення людей на фото;
2. Визначено особливості функціоналу для проектованої мережі;
3. Обрано інструменти для розробки і проектування;
4. Реалізовано програмну модель визначення людей на фото з використанням сучасних методів та засобів розробки. У якості засобів розробки обрано відкриті й безкоштовні інструменти, підтримувані світовою спільнотою;
5. Виконано тестування функціоналу проектованої мережі.

ПЕРЕЛІК ВИКОРИСТАНИХ ІНФОРМАЦІЙНИХ ДЖЕРЕЛ

1. Мобильное приложение [Электронный ресурс] Режим доступа до ресурсу:
<https://www.calltouch.ru/glossary/chto-takoe-mobilnoe-prilozhenie-i-zachem-ono-mozhet-potrebovatsya/>
2. [Cam Cullen](#) Sandvine releases COVID-19 Global Internet Phenomena Report [Электронный ресурс] Режим доступа:
<https://www.sandvine.com/press-releases/sandvine-releases-covid-19-global-internet-phenomena-report>
3. The Global Internet Phenomena Report [Электронный ресурс] Режим доступа: https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/Internet%20Phenomena/Internet%20Phenomena%20Report%20Q32019%2020190910.pdf
4. Типы мобильных приложений [Электронный ресурс] Режим доступа до ресурсу: <https://punicapp.com/blog/pages/1046/typy-mobilnyh-prilozhenij>
5. Типы мобильных приложений [Электронный ресурс] Режим доступа до ресурсу:
<https://training.qatestlab.com/blog/technical-articles/types-of-mobile-applications/>
6. Упская О. Автореферат магистерской работы «Исследование и программная реализация алгоритмов распознавания лиц» [Электронный ресурс] Режим доступа:
https://knastu.ru/media/files/page_files/page_391/magistr_referat/Avtoreferat_Upskaya_O.K._6VSm-1.pdf
7. Системы распознавания лиц Facial recognition technology (FRT) [Электронный ресурс] Режим доступа:
[https://www.tadviser.ru/index.php_\(Facial_recognition\)](https://www.tadviser.ru/index.php_(Facial_recognition))
8. Что такое искусственный интеллект [Электронный ресурс] Режим доступа: <https://www.oracle.com/ru/artificial-intelligence/what-is-ai/>

9. Нейробиологи считают неэффективным использование искусственного интеллекта для распознавания человеческих эмоций [Электронный ресурс] – Информационно-аналитический журнал РУБЕЖ / Новости. – Режим доступа до ресурсу:

<https://ru-bezh.ru/kompanii-i-ryinki/news/21/05/06/nejrobiologi-schitayut-neeffectivnyim-ispolzovanie-iskusstvennog>

10. Лучшие алгоритмы распознавания лиц

<https://www.dssl.ru/publications/stati/luchshie-algoritmy-dlya-raspoznavaniya-lits/>

11. Томас Д. Распознавание эмоций и лиц: помощь полиции или путь в антиутопию? / Би-би-си. Русская служба. – 17.07.2018 [Электронный ресурс] Режим доступа до ресурсу: <https://www.bbc.com/russian/news-44858147>

12. T. Chambel Being Happy, Healthy and Whole Watching Movies That Affect Our Emotions [Электронный ресурс] Teresa Chambel, Eva Oliveira, Pedro Martins / S. D’Mello et al. (Eds.): ACII 2011, Part I, LNCS 6974, pp. 35–45, 2011 // Springer-Verlag Berlin Heidelberg 2011 – Режим доступа до ресурсу: http://www.di.fc.ul.pt/~tc/papers/2011_ACII_happy.pdf

13. Эмоции в кино: результаты большого исследования [Электронный ресурс] Режим доступа до ресурсу: <https://vizual.club/2018/08/28/emocii-v-kino/>

14. Fear & Loathing in Cinema Theatre: our favourite genres and emotions in IMDb Top 250 [Электронный ресурс] Режим доступа до ресурсу: <https://readymag.com/u52884737/fear-and-loathing/>

15. T. Grodal How film genres are a product of biology, evolution and culture – an embodied approach [Электронный ресурс] Torben Grodal / PALGRAVE COMMUNICATIONS | 3:17079 |DOI: 10.1057 // palcomms. 2017.79. – Режим доступа до ресурсу:

<https://www.nature.com/articles/palcomms201779.pdf>

16. N. Redfern Emotion, Genre, and the Hollywood Paranoid Film [Электронный ресурс] Redfern Nick / Research Into Film. An Empirical Approach to Film Studies. – Режим доступа до ресурсу:

<https://nickredfern.wordpress.com/2009/05/07/emotion-genre-and-the-hollywood-paranoid-film/>

17. Издательство Атлантик [Электронный ресурс] Режим доступа до ресурсу:

<https://www.theatlantic.com/technology/archive/2021/04/artificial-intelligence-misreading-human-emotion/618696/?mod=djemAIPro>

18. Bardzell, S., Bardzell, J., Pace, T.: Understanding Affective Interaction: Emotion, Engagement, and Internet Videos. In: Proc. of IEEE ACII, Amst., NL (September 10-12, 2009).

19. CARLOS A. GOMEZ-URIBE The Netflix Recommender System: Algorithms, Business Value, and Innovation [Электронный ресурс] – Carlos A. Gomez-Uribe and Neil Hunt. 2015The / Netflix recommender system: Algorithms, business value, and innovation // ACM Trans. Manage. Inf. Syst. 6, 4, Article 13 (December 2015), 19 pages. – Режим доступа до ресурсу:

<https://dl.acm.org/doi/epdf/10.1145/2843948>

20. Ahlberg, C., Truvé, S.: Tight coupling: Guiding user actions in a direct manipulation retrieval system. In: People and Computers: Proc. of HCI 1995, pp. 305–321 (1995).

21. Rocha, T., Chambel, T.: VideoSpace: a 3D Video Experience. In: Proceedings of Artech 2008, 4th International Conference on Digital Arts, Porto, Portugal (November 2008).

22. Martinho, J., Chambel, T.: ColorsInMotion: Interactive Visualization and Exploration of Video Spaces. In: Proc. of Academic MindTrek 2009, Tampere, Finland (SeptemberOctober 2009).

23. Сервис по подбору фильмов на основе эмодзи попал в App Store и Google Play [Электронный ресурс] Режим доступа до ресурсу:
<https://ruvod.com/servis-po-podboru-filmov-na-osnove-emozhi-popal-v-topy-app-store-i-google-play/>

24. Р. И. Ахметшин, А. П. Кирпичников, М. П. Шлеймович
РАСПОЗНАВАНИЕ ЭМОЦИЙ ЧЕЛОВЕКА НА ИЗОБРАЖЕНИЯХ

[Электронный ресурс] Режим доступа до ресурсу:
<https://cyberleninka.ru/article/n/raspoznavanie-emotsiy-cheloveka-na-izobrazheniyah/viewer>

25. Метод Виолы-Джонса (Viola-Jones) как основа для распознавания лиц [Электронный ресурс] Режим доступа до ресурсу:
<http://habrahabr.ru/post/133826/>.

26. Си Я. Автоматическое распознавание эмоций пользователя для организации интеллектуального интерфейса // Электронный журнал «Молодежный научно-технический вестник», 2013, № 9. – Режим доступа до ресурсу: <http://sntbul.bmstu.ru/doc/616498.html>.

27. Internet Movie Database [Электронный ресурс] Режим доступа до ресурсу: https://ru.wikipedia.org/wiki/Internet_Movie_Database

28. Р. Андреев Что такое парсинг и как правильно парсить [Электронный ресурс] Режим доступа до ресурсу:
<https://blog.calltouch.ru/chto-takoe-parsing/>

29. Обнаружение объектов с 10 строчками кода [Электронный ресурс] Режим доступа:
<https://medium.com/nuances-of-programming/%D0%BE-%D0%BA%D0%BE%D0%B4%D0%B0-953bd0e22a2>

30. Питон – что это такое [Электронный ресурс] Режим доступа:
<https://mchost.ru/articles/chto-takoe-python/>

31. Язык программирования Питон [Электронный ресурс] Режим доступа:
<https://web-creator.ru/articles/python>

32. Программирование на Питон [Электронный ресурс] Режим доступа:
<https://ru.hexlet.io/blog/posts/programmirovanie-na-python-osobennosti-obucheniya-perspektivy-situatsiya-na-rynke-truda>

33. Рейтинг популярности мови серед користувачів GitHub [Электронный ресурс] Режим доступа: <https://octoverse.github.com/>

34. The RedMonk Programming Language Rankings: June 2020 [Электронный ресурс] Режим доступа: <https://redmonk.com/sograzy/2020/07/27/language-rankings-6-20/>
35. TIOBE Index for November 2021 [Электронный ресурс] Режим доступа: <https://www.tiobe.com/tiobe-index/>
36. Рейтинг языков программирования 2021: доля Python падает, а TypeScript обошел C++, в лидерах JavaScript, Java, C# [Электронный ресурс] Режим доступа: <https://habr.com/ru/post/543346/>
37. Что такое Php? [Электронный ресурс] Режим доступа: <http://www.php.su/php/?php>
38. SEO-википедия. PHP [Электронный ресурс] Режим доступа: <https://blog.ingate.ru/seo-wikipedia/php/>
39. Принимая PHP всерьёз [Электронный ресурс] Режим доступа: <https://habr.com/ru/post/314970/>
40. ЧТО ТАКОЕ MYSQL? [Электронный ресурс] Режим доступа: https://freehost.com.ua/faq/wiki/chto-takoe-mysql/?gclid=Cj0KCQiAip-PBhDVARIsAPP2xc1ss4Jg2Ra5FkRA6yPfdMa4g9aBSdw4J1qP16PeLiFsnftiyZLQl7YaAn-NEALw_wcB
41. Что Такое MySQL: Объяснение MySQL Для Начинающих [Электронный ресурс] Режим доступа: <https://www.hostinger.com.ua/rukovodstva/shto-takoje-mysql/>
42. Как разработать своё первое приложение на React Native [Электронный ресурс] Режим доступа до ресурсу: <https://tproger.ru/articles/your-first-app-in-react-native/>
43. Создание кроссплатформенных приложений с помощью React Native [Электронный ресурс] Режим доступа до ресурсу: <https://habr.com/ru/company/nix/blog/324562/>
44. Распознавание изображений на Python с помощью TensorFlow и Keras [Электронный ресурс] Режим доступа до ресурсу: <https://evileg.com/ru/post/619/>

45. Сапунов И. Deep Learning [Электронный ресурс] Режим доступа до ресурсу: http://edu.bioinf.me/files/vvedenie_v_deep_learning.pdf

46. Кессо П.И. РАСПОЗНАВАНИЕ ЭМОЦИЙ ЧЕЛОВЕКА ПО ФОТО / 55-я юбилейная научная конференция аспирантов, магистрантов и студентов БГУИР, 2019 г. – с. 39-40

https://libeldoc.bsuir.by/bitstream/123456789/35529/1/Kesso_Raspoznavaniye.pdf

47. Научитесь строить общие нейронные сети, такие как CNN RNN с Keras [Электронный ресурс] Режим доступа до ресурсу: <https://russianblogs.com/article/2184513289/>

48. OpenCV. Словарь терминов. [Электронный ресурс] Режим доступа до ресурсу: <https://blog.skillfactory.ru/glossary/opencv/>

49. Распознавание эмоций на лице для оценки социальных и психологических факторов [Электронный ресурс] Режим доступа до ресурсу: file:///D:/%/pritom_m-_model-raspoznavaniya-emociy-po-vyrajeniyu-lica-dlya-ocenk-i-socialno-psihologicheskikh-faktorov_140125.pdf

50. Функция HOG извлечения признаков изображения для обнаружения цели [Электронный ресурс] Режим доступа до ресурсу: <https://russianblogs.com/article/44111516002/>

51. FER-2013 Learn facial expressions from an image [Электронный ресурс] Режим доступа до ресурсу: <https://www.kaggle.com/msambare/fer2013>

52. Internet movie database (IMDb) [Электронный ресурс] Режим доступа до ресурсу: <https://www.imdb.com/>

```
<?php

namespace App\Services;

class MovieSelection
{
    const ANIME = '210024';
    const COMEDY = '35';
    const ACTION = '28';
    const DRAMA = '18';
    const HORROR = '27';
    const SCIENCE_FICTION = '878';
    const MYSTERY = '9648';
    const HISTORY = '36';
    const CRIME = '80';
    const WAR = '10752';
    const ROMANCE = '10749';

    const COMEDY_CRIME = '35,80';
    const DRAMA_ACTION = '18,28';
    const DRAMA_COMEDY = '18,35';

    /**
     * @param string $emotion
     * @return string
     */
    public static function getGenreToMovieDB(string $emotion): string
    {

        switch ($emotion) {
            case 'happy':
                $genres = self::COMEDY; //375
                break;
            case 'angry':
                $genres = self::ACTION; //359
```

```

        break;
    case 'sad':
        $genres = self::ROMANCE; //247
        break;
    case 'fearful':
        $genres = self::HORROR; //100
        break;
    case 'disgusted':
        $genres = self::SCIENCE_FICTION; //170
        break;
    case 'surprised':
        $genres = self::MYSTERY; //145
        break;
    case 'criminal':
        $genres = self::CRIME; //305
        break;
    case 'dramaAction':
        $genres = self::DRAMA_ACTION; //116
        break;
    case 'drama':
        $genres = self::DRAMA; //898
        break;
    default:
        $genres = self::HISTORY; //101
        break;
}

return $genres;
}

/**
 * @param string $genre
 * @return string|null
 */
public static function getReverseGenre(string $genre): ?string
{
    switch ($genre) {
        case self::ACTION:
        case self::HORROR:

```

```
    case self::CRIME:
        $genres = self::COMEDY;
        break;
    default:
        $genres = null;
}

return $genres;
}
}
```

```
<?php
declare(strict_types=1);
class ParseImdbAutoCommand extends Command
{
    const PAGES_START_MOVIES = ['0', '250', '500', '750', '1000', '1250', '1500', '1750',
'2000', '2250'];

    const PAGES_START_SERIES = ['0', '250'];

    protected static $defaultName = 'app:imdb:parse:auto';
    /**
     * @var ImdbApi
     */
    private $imdbApi;
    /**
     * @var ImdbDataHandler
     */
    private $imdbDataHandler;

    /**
     * @var EntityManagerInterface
     */
    private $entityManager;

    /**
     * @var TmdbService
     */
    private TmdbService $tmdbService;

    /**
     * ParseImdbCommand constructor.
     * @param string|null $name
     * @param ImdbApi $imdbApi
     * @param TmdbService $tmdbService
     * @param ImdbDataHandler $imdbDataHandler
     * @param EntityManagerInterface $entityManager
     */
    public function __construct(string $name = null, ImdbApi $imdbApi, TmdbService
$tmdbService, ImdbDataHandler $imdbDataHandler,
        EntityManagerInterface $entityManager ) {
        parent::__construct($name);
        $this->imdbApi = $imdbApi;
        $this->imdbDataHandler = $imdbDataHandler;
```

```

$this->entityManager = $entityManager;
$this->tmdbService = $tmdbService; }

/**
 * @param InputInterface $input
 * @param OutputInterface $output
 * @return void
 */
protected function execute(InputInterface $input, OutputInterface $output): void
{
    $output->writeln('Start parsing...' . PHP_EOL);
    $countMovieFromDb = $this->countMovies();
    $output->writeln('Movies in database - ' . $countMovieFromDb . PHP_EOL);
    foreach (self::PAGES_START_MOVIES as $pageStart) {
        $type = '1';
        $url =
'https://www.imdb.com/search/title/?title\_type=feature&release\_date=2000-01-01,&user\_rating=6.1,&num\_votes=40000,&count=250&start=' . \$pageStart . '&ref\_=adv\_nxt'';
        $output->writeln(PHP_EOL . 'movie ' . $url);
        $this->parse($output, $url, $type);
    }
    $countMovies = $this->countMovies();
    $newMovies = $countMovies - $countMovieFromDb;
    $output->writeln( PHP_EOL . 'New movies - ' . $newMovies . PHP_EOL);
    foreach (self::PAGES_START_SERIES as $pageStart) {
        $type = '2';
        $url =
'https://www.imdb.com/search/title/?title\_type=tv\_series,tv\_miniseries&release\_date=2000-01-01,&num\_votes=40000,&count=250&start=' . \$pageStart . '&ref\_=adv\_nxt'';
        $output->writeln(PHP_EOL . 'series ' . $url);
        $this->parse($output, $url, $type);
    }
    $output->writeln('Parsing completes!' . PHP_EOL);
}
private function parse(OutputInterface $output, $url, $type)
{
    try {
        $moviesFromImdbEn = $this->imdbApi->getMovieByUrl($url, 'en-EN');
        $moviesFromImdbRu = $this->imdbApi->getMovieByUrl($url, 'ru-RU');
        $moviesFromImdbUk = $this->imdbApi->getMovieByUrl($url, 'uk-UK');
        $countMovies = count($moviesFromImdbEn);
        $output->writeln(PHP_EOL . 'Movies - ' . $countMovies . PHP_EOL);
        $movieRepository = $this->entityManager->getRepository(Movie::class);
        $progressBar = new ProgressBar($output, $countMovies);
        foreach ($moviesFromImdbEn as $key => $movieFromImdb) {
            $isExistMovie = $movieRepository->findOneBy(['imdb_id' =>
            $movieFromImdb['id']]);

```