

РЕФЕРАТ

на магістерську дисертацію

виконану на тему: Використання конвуляційних нейронних мереж для генерації структурної моделі обличчя

студентом: Савельєвим Юрієм Дмитровичем

Робота виконана на 98 сторінках, містить 33 ілюстрації, 25 таблиць. При підготовці використовувалася література з 71 різних джерел.

Актуальність

Задача детекції облич є добре вивченою проблемою у комп'ютерному зорі. Сучасні детектори облич можуть легко знайти та обробити його фронтальні зображення. Нещодавні дослідження у цій області фокусуються на менш контрольованому процесі детекції, де число можливих факторів, такі як зміна пози, вираження емоцій і незвичайне освітлення може призвести до великих візуальних варіацій у вигляді обличчя, що значно погіршує якість роботи детектора облич. Найбільш важкою задачею комп'ютерного зору залишається проблема вирішення неоднозначності, що виникає при перенесенні трьовимірних об'єктів реального простору на плоскі зображення. Дуже багато аспектів залежить від важкопрогнозованих фактів.

Механізм детекції облич на зображеннях знаходить використання у різноманітних додатках, а задача є однією з основоположних у комп'ютерному зорі. Область використання систем розпізнавання осіб, вже не обмежується верифікацією особи і спостереженням. Все більше додатків використовують розпізнавання осіб як перший крок до інтерпретації дій людини, його намірів і поведінки. Інакше кажучи, до реалізації тих можливостей, які будуть грати центральну роль в інтелектуальному середовищі наступного покоління.

Мета

Метою роботи є розроблення системи для детекції облич на зображенні та генерації структурної моделі знайдених облич, пошук варіантів використання цих даних у системах розпізнавання облич. Провести аналіз існуючих підходів до використання нейронних мереж у розпізнаванні облич.

Завдання

Для досягнення мети були поставлені наступні завдання:

- Зробити аналіз існуючих методів, моделей, алгоритмів і програмно-апаратних систем для розпізнавання облич;
- На основі аналізу розробити архітектуру та вимоги до системи розпізнавання облич
- Розробити модель на основі нейронної мережі для детекції облич на зображенні та їх ідентифікації;
- Спроектувати програмні засоби детекції, розпізнавання та генерації структурної моделі облич;

Об'єкт дослідження

Об'єктом дослідження було обрано використання нейронних мереж для обробки зображень з людськими обличчями.

Предмет дослідження

Предметом дослідження в даній роботі є можливості використання нейронних мереж для детекції і генерації структурної моделі обличчя, а також варіанти аналізу цих даних.

Методи досліджень

Проведення аналізу ефективності використання нейронних мереж на різних етапах у процесі розпізнавання облич, від етапу детекції обличчя, до вирішення задачі ідентифікації.

Наукова новизна

Наукова новизна роботи полягає в:

1. Виробленні рекомендацій щодо вибору підходів і стратегій до побудови систем для аналізу людських облич
2. Виробленні набору модулів, для вирішення задачі отримання універсальних даних людського обличчя в умовах реального світу, що можуть бути використані для широкого спектру задач, як розпізнавання настрою, стану, активності, емоцій.

Практична цінність

В ході роботи був розроблений модульний додаток, що вирішував задачу детекції, побудови структурної моделі обличчя, а на основі цих даних міг ідентифікувати людину, що наявна в базі даних. Модульність архітектури не потребує втручання у інші модулі для зміни роботи класифікатору.

Ключові слова

Штучні нейронні мережі, конвуляційні нейронні мережі, комп'ютерний зір, розпізнавання патернів, розпізнавання облич, SVM, Caffe, Torch7, Tensorflow.

РЕФЕРАТ

на магистерскую диссертацию

выполненную на тему : Использование конвуляционных нейронных сетей для генерации структурной модели лица

студентом: Савельевым Юрием Дмитриевичем

Работа выполнена на 98 страницах, содержит 33 иллюстрации, 25 таблицы. При подготовке использовалась литература из 71 разных источников.

Актуальность

Задача детекции является хорошо изученной проблемой в компьютерном зрении. Современные детекторы лиц могут легко найти и обработать его фронтальные изображения. Последние исследования в этой области фокусируются на менее контролируемом процессе детекции, где число возможных факторов, такие как изменение позы, выражения эмоций и необычное освещение может привести к большим визуальным вариациям в виде лица, что значительно ухудшает качество работы детектора лиц. Наиболее трудной задачей компьютерного зрения остается проблема решения неоднозначности, возникающей при переносе трёхмерных объектов реального пространства на плоские изображения. Очень многие аспекты такого переноса зависят от трудно прогнозируемых фактов.

Механизм детекции лиц на изображениях находит применение в различных приложениях, а задача является одной из основополагающих в компьютерном зрении. Область применения систем распознавания лиц, уже не ограничивается верификацией личности и наблюдением. Все больше приложений используют распознавание лиц как первый шаг к интерпретации действий человека, его намерений и поведения. Иначе говоря, к реализации тех возможностей, которые будут играть центральную роль в интеллектуальной среде следующего поколения.

Цель

Целью работы является разработка системы для детекции лиц в кадре и генерации структурной модели найденных лиц, поиск вариантов использования этих данных в системах распознавания лиц. Провести анализ существующих подходов к использованию нейронных сетей в распознавании лиц.

Задание

Для достижения цели были поставлены следующие задачи:

- Сделать анализ существующих методов, моделей, алгоритмов и программно-аппаратных системах для распознавания лиц;
- На основе анализа разработать архитектуру и требования к системе распознавания лиц
- Разработать модель на основе нейронной сети для детекции лиц в кадре и их идентификации;
- Спроектировать программные средства детекции, распознавания и генерации структурной модели лиц;

Объект исследования

Объектом исследования было выбрано использование нейронных сетей для обработки изображений с человеческими лицами.

Предмет исследования

Предметом исследования в данной работе является возможность использования нейронных сетей для детекции и генерации структурной модели лица, а также варианты анализа этих данных

Методы исследований

Проведение анализа эффективности использования нейронных сетей на различных этапах в процессе распознавания лиц, от этапа детекции лица, к решению задачи идентификации.

Научная новизна

Научная новизна работы заключается в:

1. выработка рекомендаций по выбору подходов и стратегий к построению систем для анализа человеческих лиц

2. Выработка набора модулей, для решения задачи получения универсальных данных людского лица в условиях реального мира, которые могут быть использованы для широкого спектра задач, как распознавание настроения, состояния, активности, эмоций.

Практическая ценность

В ходе работы было разработано модульное приложение, которое решало задачу детекции, построения структурной модели лица, а на основе этих данных мог идентифицировать человека, имеющаяся в базе данных. Модульность архитектуры не требует вмешательства в другие модули для изменения работы классификатору.

Ключевые слова

Искусственные нейронные сети, конвуляционные нейросети, компьютерное зрение, распознавание шаблонов, распознавание лиц, SVM, Caffe, Torch7, Tensorflow.

ABSTRACT

on master's thesis

on topic: Convolutional neural networks for facial landmark generation

Student: Savelyev Yuriy Dmitrovich

Work carried out on 98 pages containing 33 figures, 25 tables. The paper was written with references to 71 different sources.

Relevance

The task of detection is a well-studied problem in computer vision. Modern face detectors can easily find and process its front images. Recent research in this area focuses on a less controlled detection process, where the number of possible factors, such as changing postures, expressing emotions and unusual lighting can lead to large visual variations in the form of a face, which significantly degrades the quality of the face detector. The most difficult task of computer vision remains the problem of solving the ambiguity arising when moving three-dimensional objects of real space to flat images. Many aspects of this transfer depend on hard-to-predict facts.

The mechanism of face detection in images is used in various applications, and the task is one of the fundamental in computer vision. The scope of face recognition systems is no longer limited to verification of identity and observation. More and more applications use face recognition as the first step to interpreting human actions, their intentions and behavior. In other words, to realize those opportunities that will play a central role in the next generation intellectual environment, recognizing and generating a structural model of individuals;

Goal

The aim of the work is the development of a system for detecting persons in the frame and generating a structural model of the found persons, searching for options for using these data in face recognition systems. Conduct an analysis of existing approaches to the use of neural networks in face recognition.

Task

To achieve the goal, the following tasks were set:

- Make an analysis of existing methods, models, algorithms and software and hardware systems for face recognition;
- Based on the analysis, develop the architecture and requirements for the face recognition system
- Develop a model based on the neural network for the detection of persons in the frame and their identification;
- Design software tools for detecting

Object of study

The object of the study was the use of neural networks for image processing with human faces.

Subject of study

The subject of the study in this paper is the possibility of using neural networks for the detection and generation of a structural model of a person, as well as options for analyzing these data

Research methods

Conducting an analysis of the efficiency of the use of neural networks at various stages in the process of recognizing persons, from the stage of detection of a person, to solving the problem of identification.

Scientific novelty

The scientific novelty of the work is:

1. Development of recommendations for the selection of approaches and strategies for the construction of systems for the analysis of human faces
2. Development of a set of modules for solving the problem of obtaining universal data of a human face in real-world conditions, which can be used for a wide range of tasks, like the recognition of mood, condition, activity, emotions.

Practical value

In the course of the work, a modular application was developed that solved the task of detecting, constructing a structural model of a person, and on the basis of this data it could identify a person available in the database. Modularity of the architecture does not require intervention in other modules to change the work of the classifier.

Keywords

Artificial neural networks, convolutional nerves, computer vision, pattern recognition, face recognition, SVM, Caffe, Torch7, Tensorflow.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ	12
ВСТУП	13
1 АВТОМАТИЗОВАНЕ РОЗПІЗНАВАННЯ ОБЛИЧ	16
1.1 Питання автоматизованого розпізнавання облич	16
1.2 Основні складнощі	17
1.2.1 Варіації пози	17
1.2.2 Наявність/відсутність певних елементів/оклюзій(перекрить)	18
1.2.3 Зміни емоцій на обличчі	18
1.2.4 Старіння обличчя	19
1.2.5 Варійовані умови освітлення	19
1.2.6 Роздільна здатність зображення	20
Висновки	24
2 НЕЙРОННІ МЕРЕЖІ	25
2.1 Штучні нейронні мережі. Навчання з вчителем	25
2.2 Конвуляційні нейронні мережі	30
2.3 Функції активації	35
Висновки	39
3 АНАЛІЗ МЕТОДІВ РОЗПІЗНАВАННЯ ОБЛИЧ	40
3.1 Структура системи розпізнавання облич	40
3.2 PCA	41
3.3 PCA з використанням штучних нейронних мереж	44
3.4 Deep Dense Face Detector	45
3.5 Radial Basis Function Neural Networks	45
3.6 Convolutional Neural Network Cascade	46
3.7 Bilinear CNNs	46
3.8 Back Propagation Network та Radial Basis Function Network	47
3.9 Retinal Connected Neural Network (RCNN)	48
3.10 Rotation Invariant Neural Network (RINN)	48

	11
Висновки	49
4 РОЗРОБКА ДОДАТКУ ДЛЯ ГЕНЕРАЦІЇ СТРУКТУРНОЇ МОДЕЛІ ОБЛИЧЧЯ З ВИКОРИСТАННЯМ КОНВУЛЯЦІЙНИХ НЕЙРОННИХ МЕРЕЖ	50
4.1 Загальний алгоритм роботи додатку	50
4.2 Пошук облич на зображенні.....	51
4.3 Аналіз обличчя. Генерація структурної моделі	55
4.4 Пошук людини по базі даних.....	58
Висновок.....	63
5 РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ	64
5.1 Інформаційна карта проекту	64
5.2 Команда стартап-проекту	65
5.3 Маркетингова стратегія та маркетинговий план стартапу	65
5.3.1 Опис ідеї проекту (товару, послуги, технології)	65
5.3.2 Технологічний аудит ідеї проекту	67
5.3.3 Аналіз ринкових можливостей запуску стартап-проекту.....	69
5.4 Розроблення ринкової стратегії проекту.....	76
5.5 Розроблення маркетингової програми стартап-проекту	81
Висновки	88
ВИСНОВКИ	89
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	91

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ

HOG (англ. - Histogram of Oriented Gradients)	– гістограма орієнтованих градієнтів
AFR (англ. - Automated face recognition)	– автоматизоване розпізнавання облич
Tanh	– гіперболічний тангенс
ReLU (англ. - Rectified Linear Unit)	– випрямлений лінійний елемент
PCA	– метод головних компонент
DAG (англ. - directed acyclic graph)	– орієнтований ациклічний граф
BPN (англ. - Back Propagation Network)	– мережа оберненого розповсюдження похибки
RBF (англ. - Radial Basis Function Network)	– радіально базисна мережа
CNN (англ. Convolutional Neural Network)	– конвуляційна нейронна мережа
ANN (англ. Artificial Neural Network)	– штучна нейронна мережа
RPN (англ. - Region proposal network)	– мережа пропозицій областей
NMS (англ. - non-maximum suppression)	– алгоритм подавлення не максимумів
FLE (англ. - face landmark estimation)	– визначення розмітки обличчя

ВСТУП

Задача детекції облич є добре вивченою проблемою у комп'ютерному зорі. Сучасні детектори облич можуть легко знайти та обробити його фронтальні зображення. Нещодавні дослідження у цій області фокусуються на менш контрольованому процесі детекції, де число можливих факторів, такі як зміна пози, вираження емоцій і незвичайне освітлення може призвести до великих візуальних варіацій у вигляді обличчя, що значно погіршує якість роботи детектора облич. Найбільш важкою задачею комп'ютерного зору залишається проблема вирішення неоднозначності, що виникає при перенесенні трьовимірних об'єктів реального простору на плоскі зображення. Дуже багато аспектів залежить від важкопрогнозованих фактів.

Складність у детекції облич в основному диктується двома аспектами:

- велика кількість візуальних змін облич людини на різноманітному фоні
- великий розмір простору пошуку можливих позицій облич та їх розмірів

Перший потребує детектору облич точно вирішувати проблему класифікації, коли другий нав'язує проблеми з потребами у швидкодії.

Механізм детекції облич на зображеннях знаходить використання у різноманітних додатках, а задача є однією з основоположних у комп'ютерному зорі. Область використання систем розпізнавання осіб, вже не обмежується верифікацією особи і спостереженням. Все більше додатків використовують розпізнавання осіб як перший крок до інтерпретації дій людини, його намірів і поведінки. Інакше кажучи, до реалізації тих можливостей, які будуть грати центральну роль в інтелектуальному середовищі наступного покоління. Багато дій та особливості поведінки людини піддаються інтерпретації тільки в тому випадку, якщо ідентифікуються його особу і людей, які її оточують. Прикладами можуть слугувати розпізнавання постійного покупця магазину, спостереження за поведінкою пацієнтів, інтерфейси командного управління на військових і промислових об'єктах.

З часів фундаментальної роботи Віоли та Джонса [2], прискорений каскад з простими ознаками залишається найбільш популярним та ефективним підходом у розробці програм для практичної детекції облич. Проста природа ознак дозволяє швидко оцінювати і рано відкидати невірні результати пошуку. Тим часом, прискорений каскад створює групу простих ознак для досягнення точної класифікації облич з не-обличчями. Оригінальний детектор Віоли-Джонса використовує ознаки Хаара, які швидко вираховуються, та їх достатньо для опису фронтальних зображень облич. Тим не менш, через простоту ознак Хаара, алгоритм відносно слабкий у неконтрольованому середовищі.

Згодом, у 2005 році, з'явився більш універсальний і зручний метод, що базується на HOG дескрипторах [3]. Метод демонструє гарні результати і активно використовується навіть зараз для детекції облич. Зокрема алгоритми пошуку облич у сучасних камерах допомагають краще визначити область фокусу.

В останні роки спостерігається сплеск уваги до нейронних мереж. Посилений інтерес виник у 2012 році, коли Алекс Крижевський за допомогою конвуляційних нейронних мереж переміг у конкурсі ImageNet [1], понизивши рекорд помилок у класифікації з 26% до 15%, що тоді стало проривом. Сьогодні «глибоке навчання» стоїть у в основі багатьох систем великих компаній: Facebook використовує нейронні мережі для алгоритмів автоматичного виставлення тегів, Google — для пошуку серед фотографій користувача, Amazon — для генерації рекомендацій товарів, Pinterest — для персоналізації домашньої сторінки користувача, а Instagram — для пошукової інфраструктури.

Фокус на генерації структурної моделі пояснюється великою кількістю можливостей використання цієї інформації. Ця проблема є більш загальною та важливою в контексті обробки зображень облич людини. Інформація може бути використана як для розпізнавання облич, емоцій людини та інших ознак обличчя, так і для обробки зображення.

Отже, актуальною науково-прикладною проблемою є створення підходу до пошуку облич людини та ключових точок на ньому на двовимірному

зображенні за допомогою систем штучного інтелекту для побудови структурної моделі обличчя, що може бути використана для розпізнавання людини.

Об'єктом дослідження – нейронні мережі для обробки зображень з людськими обличчями.

Предмет дослідження – генерація структурної моделі обличчя на основі методів штучного інтелекту.

Мета і задачі дослідження. Метою роботи є розроблення системи для детекції облич на зображенні та генерації структурної моделі знайдених облич, пошук варіантів використання цих даних у системах розпізнавання облич. Детекція відбувається за рахунок використання методів штучного інтелекту таких як нейронні мережі.

Для досягнення поставленої мети необхідно розв'язати такі задачі:

- проаналізувати існуючі методи, моделі, алгоритми і програмно-апаратні системи для розпізнавання облич;
- розробити алгоритм для детекції облич, який буде працювати в складних умовах та мати високу якість розпізнавання;
- розробити модель на основі нейронної мережі для детекції облич на зображенні;
- спроектувати програмні засоби розпізнавання облич;
- розробити стратегію стартап проекту який дозволить реалізувати описану технологію в якості конкурентоспроможного продукту.

1 АВТОМАТИЗОВАНЕ РОЗПІЗНАВАННЯ ОБЛИЧ

1.1 Питання автоматизованого розпізнавання облич

Автоматизоване розпізнавання облич отримало велику кількість уваги від дослідників та представників індустрії протягом декількох останніх поколінь як через великий спектр наукових викликів [5], так і через велику кількість способів застосування у комерційних додатках [6], частково у контексті біометрії та питань безпеки. З недавніх пір технології знайшли використання у сфері мультимедіа та соціальних медіа [4].

Розпізнавання облич це область, що намагається дати відповідь на питання: «Кому належить це обличчя?». Для цієї цілі люди мають природні здібності за допомогою власних перцептивних і когнітивних систем, тоді як машини потребують комплексні системи, що включають в себе передові алгоритми та великі бази даних облич. Вивчення, розробка та проектування таких методів і технологій – це проблеми області автоматизованого розпізнавання облич.

Автоматизоване розпізнавання облич може згодом стати частиною систем комп'ютерної автоматизованої ідентифікації облич та комп'ютерної автоматизованої верифікації облич. З одного боку, автоматизоване ідентифікації обличчя полягає у один-до-багатьох (1:N) пошуку обличчя серед бази даних, що містить зображення облич багатьох різних людей для відповіді на питання «Чи відоме це обличчя?» [7]. З іншого боку, автоматизована верифікація – це пошук один-до-одного для пошуку відповіді на питання «Чи належить це обличчя до ...?» [8].

Окрім того, автоматизоване розпізнавання облич може бути базисом для рішення питання «Хто знаходиться на зображенні?», що допоможе вирішити проблему автоматизованого виставлення міток облич/найменування облич [9].

Загалом процес автоматизованого розпізнавання облич зображений на рисунку 1.1. Спочатку вирішується питання «Чи присутнє обличчя на

зображенні?» (детекція облич). Наступне питання – «Де знаходиться обличчя?». Після цього виконується процес автоматизованого розпізнавання [10].



Рисунок 1.1 - Загальний процес автоматизованого розпізнавання облич [4]

1.2 Основні складнощі

1.2.1 Варіації пози

Рух голови, що може бути описаний у формі егоцентричного кута оберту, або зміна точку огляду камери може призвести до суттєвих змін у вигляді або формі та створити велику кількість варіацій, що проілюстровані на рис. 1.2, що робить розпізнавання обличчя у позах важкою задачею.



Рисунок 1.2 - Ілюстрація можливих поз голови в залежності від площини нахилу [4]

Оскільки автоматизоване розпізнавання обличчя чутливе до варіацій, корекція пози є важливим елементом і може бути досягнуте використанням різноманітних технік націлених на вирівнювання обличчя [11].

1.2.2 Наявність/відсутність певних елементів/оклюзій(перекрить)

Різноманітність варіантів одного обличчя також може виникати через відсутність певних структурних елементів, або ж через наявність таких компонентів як борода чи/та вуса, головних уборів, сонцезахисних окулярів, і т.д., або ж через оклюзій лиця елементами фону або переднього плану.



Рисунок 1.3 - Ілюстрація відсутності (а) або наявності (b-d) структурних елементів, а саме бороди та вус(б), кепки(с), сонцезахистних окулярів(д) та часткової оклюзії(е) [4]

1.2.3 Зміни емоцій на обличчі

Деякі варіації у вигляді обличчя спричинені змінами через вираження емоцій обличчям у варійованих емоційних станах, що зображені на рисунку 1.4.

Ефективне автоматичне розпізнавання різних виразів обличчя важливе як для еволюції емоційних станів, так і для автоматизованого розпізнавання облич. Загалом, людські вирази обличчя складаються з макро-виразів, що можуть виражати злість, страх, сум, радість, здивування та інші швидкі вирази обличчя, тобто мікро-вирази [4]. Усі ці вирази створюють рух обличчя. Так динаміка обличчя може бути врахувана за допомогою методів optical flow [12].



Рисунок 1.4 - Ілюстрація змін обличчя при вираженні емоцій, таких як злість(а), відраза(б), сум(с) та щастя(д) [4]

1.2.4 Старіння обличчя

Інша причина зміни вигляду обличчя полягає у старінні людського обличчя, що може мати вплив на процес автоматичного розпізнавання облич, якщо час поміж кожним зображенням є суттєвим [13].



Рисунок 1.5 - Ілюстрація процесу людського старіння, де одна і та сама людина була сфотографована у різному віці [4]

Для вирішення проблеми старіння обличчя у автоматизованому розпізнаванні обличчя, методи мають обробляти патерни старіння обличчя. Окрім цього, з часом змінюються не лише форма та лінії обличчя, а також і інші аспекти, як зачіска, наприклад [14].

1.2.5 Варійовані умови освітлення

Великі зміни у освітленні можуть суттєво знизити якість роботи систем автоматизованого розпізнавання облич. Зокрема, за низького рівня освітлення переднього або заднього плану, детекція та розпізнавання облич відбуваються значно гірше, оскільки тіні на обличчі можуть зробити певні патерни обличчя, які не будуть відрізнятися. З іншого боку, занадто яскраве освітлення може призвести до переекспонування фотографії і також зробити певні особливості [4].

Швидкі автоматизовані детекція та розпізнавання облич через занадто варійоване освітлення потребують використання таких технік обробки зображення як нормалізація освітлення через виведення гістограми, або методи машинного навчання, що працюють з дійсною загальною інтенсивністю зображення [15,16].



Рисунок 1.6- Ілюстрація варіантів змін освітлення камери, що призводять до переекспонування(a), глибоких тіней(b) або часткового затемнення(c)

1.2.6 Роздільна здатність зображення

Інші звичайні фактори, що впливають на якість автоматизованого розпізнавання облич відносяться до якості та роздільної здатності зображення обличчя або до налаштувань можливостей цифрового спорядження для створення зображень. Для цієї цілі ISO/IEC 19794-5 стандарт [17] був розроблений для специфікації сцени та фотографічних потреб у ролі формату зображення обличчя для автоматизованої системи розпізнавання облич, особливо в контексті біометрії. Однак, реальні ситуації отримання зображень людини включають використання різноманітного заліза для фотографування, що можуть бути ширококутними, мати специфічний нахил-поворот-масштаб, чутливі специфічні сенсори, фотометричне стерео. Камери можуть працювати у спектрі видимого світла, або у інфрачервоному діапазоні, що спричинює складнощі для систем автоматизованого розпізнавання облич. Обличчя отримані в умовах реального світу створюють нові виклики для систем AFR [18].



Рисунок 1.7 - Ілюстрація варіацій масштабу зображень та якості [4]

Наприклад, на рис. 1.7 в деяких ситуаціях обличчя може бути сфотографоване на відстані, що призведе до меншого розміру обличчя у порівнянні зображенням у великому масштабі. Також, деякі камери можуть мати малу роздільну здатність, що призведе до низької якості фотографії обличчя, за якої AFR дуже важко провести [19, 20]. Щоби розібратися з цими обмеженнями, були запропоновані алгоритми реконструкції зображень з низькою роздільною здатністю [21, 22].

1.2.7 Доступність і якість наборів даних облич

Кожна AFR технологія потребує доступний, якісний і реалістичний набір даних облич для виконання 1:N та 1:1 пошуку облич. Якості як повнота(містить різноманітні пози, освітлення, вирази облич), точність (патерни для різного віку) і характеристики(різноманітні формати файлів, рівень кольорів/сірого, роздільна здатність, умови) є ключовими для процесу AFR. Також, при роботі з даними облич, людські права та привітність мають бути враховані оскільки AFR системи мають бути сумісними з Data Protection Act 2010 [23].

Для дослідницьких цілей, декілька наборів даних були розроблені і є публічно доступними. Слід звернути увагу на наступні:

- ORL являє собою набір даних на 400 зображень з 40 різних об'єктів, PGM форматі і з розділенням 92×112 , 8-бітного рівня сірого. Чоловічі і жіночі обличчя сфотографовані на однорідному темному фоні, при різних умовах освітлення. Суб'єкти нахиляють голову в вгору-вправо, фронтально, з

- варіаціями в лицьових виразів, деталей особи і обертають в межах $\pm 20\%$ [24].
- The Face Recognition Technology (FERET) – база даних, що складається з 14,126 зображень облич 1199 людей, для кожного 5–11 зображень у градаціях сірого. Кожний набір містить зображення, що зняти під різним кутом з різними виразами обличчя, з використанням декількох камер та різних умов освітлення [25].
 - BioID Face database містить 1521 фронтальних зображень облич 23 людей. Розмір зображень 384×286 пікселів у png форматі, зроблених у реальних умовах, тобто, присутнє різноманіття розмірів, освітлення та інших умов фотографії [26].
 - Yale face database містить 165 gif зображень 15 індивідів. На кожну особу приходиться 11 зображень, по одному на кожен вираз обличчя і конфігурацію світла ліворуч/праворуч/зверху, з та без окулярів [27].
 - Caltech 10,000 web faces складається з 10,524 зображень облич людини у різних наборах (портрети, групові фото і т.д.) від сервісу Google Image. Координати очей, носу та центру роту для кожного фронтального зображення наведені для використання перевірки роботи алгоритмів розпізнавання облич, або для детекції облич перед AFR [28].

Деякі набори даних містять 2D та 3D інформацію про обличчя, такі як Face Recognition Grand Challenge (FRGC) датасет [29], що містить 50,000 контрольованих та не контрольованих зображень від 4003 суб'єктів.

Інші датасети містять декілька модальностей, як наприклад XM2VTSDB мульти-модальний датасет, що є розширенням M2VTS датасету. Це великий, мультимодальний набір зроблений за допомогою високоякісного відео. Він містить 4 записи, кожен з головою, що розмовляє, та головою, що обертається, 295 суб'єктів, збирався він протягом 4 місяців. Ця база включає в себе високоякісні фотографії, 32 kHz 16-bit звукові файли, відео послідовності та 3D моделі [4, 30].

Інша мультимодальна база даних - Surveillance Cameras Face (SCFace) датасет. Він містить записи 4160 статичних людських облич 130 суб'єктів, у видимому та інфрачервоному спектрі, у кімнатних умовах, знятих на набір камер з п'яти відеокамер, що записують різноманітні якості міміки в умовах реального світу [31].

Останні розробки датасетів фокусуються на зборі даних в реальних умовах, тобто невимушених умовах. Наприклад, Face Detection Data Set and Benchmark (FDDb) це набір з 2845 зображень, кольорових та у градаціях сірого, з 5171 обличчями у реальних умовах, що включають в себе оклюзії, варіації поз, низьку роздільну здатність та погане фокусування на обличчях [32].

Labelled Faces in the Wild (LFW) датасет – це популярний набір даних для вивчення різнопланових облич. Він містить записи 13,233 зображень облич на передньому плані; Інші обличчя асимілюються до фону. Він містить 5749 окремих облич різних людей, що можуть мати одне або декілька фотографій на кожного, окрім того присутні варіації у присутніх позах, освітленні, виразах облич, фоні, расі, етнічній приналежності, віку, полу, одязі, типах зачісок, якості камер, типі освітлення, фокусі і так далі. Зображення 250×250 пікселів та у форматі jpeg, більшість - кольорові, лише декілька у відтінках сірого [33].

Деякі інші датасети зібрані для спеціальних цілей. Наприклад, Spontaneous Micro-expression датасет (SMIC) використовується для розпізнавання мікро-емоцій, коли Acted Facial Expression in the Wild (AFEW) датасет, що містить напівавтоматична зібрані зображення зіграних емоцій у фільмах, присвячений для розпізнавання макро-емоцій у умовах близьких до реальних. З іншого боку, FG- NET Ageing database (FG-NET) може використовуватися для розпізнавання віку, розробки інваріантних до віку AFR та вивченню зміни при віці [34,35].

Висновки

Задача автоматичного розпізнання облич на сьогоднішній день є досить актуальною, як через велику кількість наукових досліджень даної задачі так і через великий потенціал використання даної технології у комерційних проектах.

Задачу слід розглядати як комплексну проблему, наприклад розділюючи детекцію та ідентифікацію. Рішення в одній підзадачі матиме вплив на вирішення іншої, але тим не менш їх можна вирішувати окремо.

Однак дана технологія має деякі складності, які не дають можливості застосовувати дану технологію на 100%:

- Варіації пози людини (нахил або поворот голови)
- Наявність/відсутність певних елементів/оклюзій(перекрив)
- Зміни емоцій на обличчі
- Старіння обличчя
- Варійовані умови освітлення
- Роздільна здатність зображення

Процес вирішення задачі ідентифікації людини полягає у послідовній відповіді на наступні питання:

- Чи присутнє обличчя на зображенні?
- Де знаходиться обличчя?
- Кому належить це обличчя?

2 НЕЙРОННІ МЕРЕЖІ

2.1 Штучні нейронні мережі. Навчання з вчителем

Один з основних підходів, найбільш широко використовувався в області розпізнавання зображень, являє собою застосування класичних моделей-класифікаторів, що навчаються з учителем. Для навчання таких моделей використовуються маркована вибірка даних, що складається з масиву зображень і відповідного їм масиву міток, що визначають категорію, до якої відноситься зображення. В процесі навчання масив даних розділяється на дві нерівні частини - навчальну вибірку і тестову вибірку, потім за допомогою специфічного для конкретного алгоритму правила навчання параметри моделі налаштовуються з використанням навчальної вибірки таким чином, щоб отримавши в якості вхідних даних зображення, модель на виході виробляла б мітку відповідного класу. Цей підхід представлений безліччю моделей, серед яких найбільш широко використовуваними є регресивна модель, штучна нейронна мережа (багатошаровий перцептрон), метод опорних векторів, а також дерева прийняття рішень і моделі-ансамблі, що представляють собою поєднання деяких перерахованих моделей [36]

Спроби відтворити здатність навчатися і виправляти помилки призвели до створення штучних нейронних мереж. Штучні нейронні мережі являють собою сімейство моделей, побудованих по принципу організації і функціонування біологічних нейронних мереж — мереж нервових клітин живого організму. Поняття штучної нейронної мережі було запропоновано ще у 1943 році У. Маккалоком і У. Піттсом в статті [37]. Зокрема, ними була запропонована модель штучного нейрона. Щоб відобразити суть біологічних нейронних систем, штучний нейрон будується наступним чином. Він отримує вхідні сигнали (вихідні дані або вихідні сигнали інших нейронів нейронної мережі) через кілька вхідних каналів. Кожен вхідний сигнал проходить через з'єднання, що має певне політичне значення. З кожним нейроном пов'язане значне порогове значення. Обчислюється зважена сума входів, з неї віднімається граничне значення і в

результаті виходить величина активації нейрона. Сигнал активації перетворюється за допомогою функції активації і в результаті виходить вихідний сигнал нейрона.

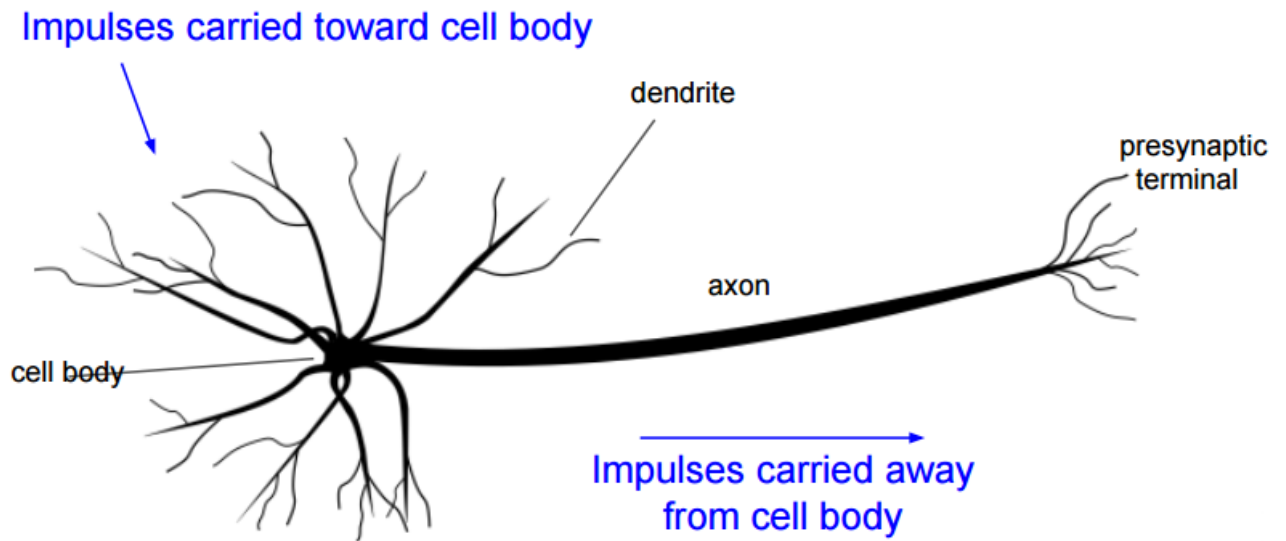


Рисунок 2.1 - Модель нейрону живого організму [53]

Багатошарові перцептрони, які навчаються методом зворотного поширення помилки, широко використовуються для розпізнавання різних категорій зображень, таких як рукописні цифри [16], почерк [39], людські обличчя [40] і дані зорових сенсорів робототехнічних систем [41]. Модель багатошарового перцептрона являє собою сукупність штучних нейронів - обчислювальної одиниці моделі - об'єднаних в рівні (шари), задані в ієрархічному порядку [36].

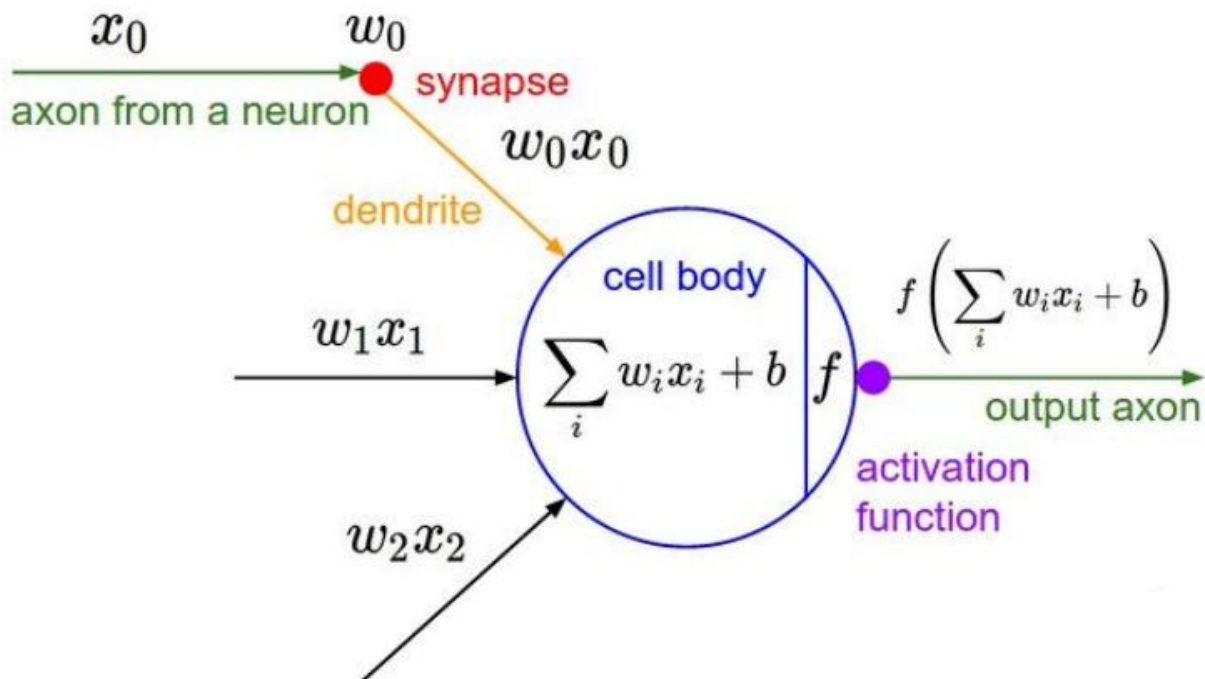


Рисунок.2.2 - Модель штучного нейрону, x_i – вхідний сигнал, w_i – вага вхідного сигналу, $f(\cdot)$ – функція активації [53]

Поведінка нейрона будується наступним чином: нехай є $m + 1$ входів, значення яких дорівнюють $x_0, x_1, x_2, \dots, x_m$, а значення їх ваг рівні, w_0, w_1, \dots, w_m , при цьому перший вхідний елемент, як правило, являє собою фіксоване значення зміщення $x_0 = 1$. Тоді вихідне значення нейрона є значення функції активації від зваженої суми його вхідних значень:

$$y = \varphi\left(\sum_{i=0}^m w_i x_i\right)$$

При об'єднанні штучних нейронів в мережу вхідні значення нейрона шару l являють собою вихідні значення нейронів попереднього шару $l - 1$. При цьому нейрони першого (вхідного) шару отримують в якості вхідних значення безпосередньо дані, що підлягають розпізнаванню, які в разі розпізнавання зображення представляють собою значення інтенсивності складових його пікселів (точкових елементів). Вихідний шар мережі може варіюватися в залежності від завдання, але класична архітектура має на увазі формування його числом нейронів, рівній кількості класів розпізнавання, при цьому вихідне

значення кожного нейрона нормується по інтервалу $\{0,1\}$, і являє собою ймовірність приналежності вхідного зображення до відповідного класу. Як відзначають дослідники, такі багатошарові нейронні мережі здатні інкапсулювати будь-яку математичну функцію за допомогою довільного набору нейронів [36].

Оскільки сформулювати аналітично правило класифікації зображень за категоріями розпізнавання часто представляється скрутним, здатність навчатися на базі вибірки робить нейронні мережі та споріднені з ними моделі придатними для розпізнавання природних зображень навколишнього світу, що відрізняються нечіткою структурою і безліччю варіацій в межах класу [36].

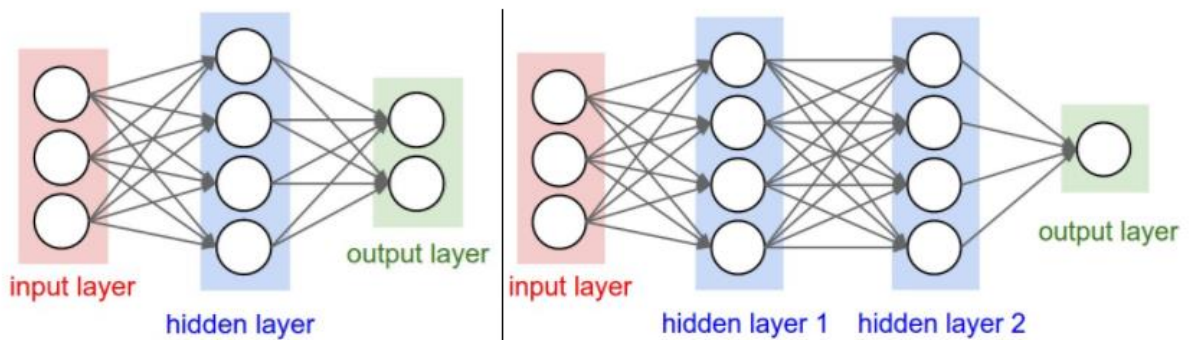


Рисунок 2.3 - Ліворуч: двошарова нейронна мережа, один прихований шар з чотирьох нейронів. Праворуч: 3-шарова мережа з двома прихованими шарами[53]

Навчання мережі методом зворотного поширення полягає в наступному: нехай є деяка невідома функція розпізнавання $g : X \rightarrow Y$, аргументом котрої є зображення $x_n \in X$, представлені у вигляді вектора довжиною n , а значеннями функції – множина класів (категорій) $y \in Y$. Навчальна вибірка складає собою підмножину значень цієї функції $D = \{(x_0, y_0), (x_1, y_1), \dots, (x_m, y_m)\}$ [36].

Задача навчання моделі розпізнавання полягає у пошуку такої функції $h : X \rightarrow Y$, яка б апроксимувала функцію g на всій області її визначення, в тому числі в значеннях, що не включені в D , і представляють собою додаток теорії оптимізації [36].

Нехай $h(x)$ – вихідне значення мережі, отримане шляхом послідовної активації нейронів кожного шару, а $g(x)$ – значення апроксимованої функції для одного й того ж зображення. Далі виконується крок зворотнього поширення, що полягає у вирахуванні часткової похідної для кожного нейрону мережі по відношенню до його вагових коефіцієнтів [36]:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_i} \frac{\partial o_i}{\partial net_j} \frac{\partial net_j}{\partial w_{ij}}$$

$$E = \frac{1}{2} (h(x) - g(x))^2,$$

де E – середньоквадратичне відхилення мережі. Далі на кожному кроці навчання ваги нейронів інкрементуються значеннями часткових похідних в відповідності до методу градієнтного спуску, або інших градієнтних методів навчання. Варіації навчального алгоритму включають в себе використання додаткових параметрів регуляризації з метою захисту від перенавчання, і використання різних оптимізаторів - методу Ньютона, методу імітації відпалу, L-BFGS і інших [36].

Нейронна мережа – це дуже потужна та швидка техніка класифікації що може бути використана не лише для передбачення відомих даних, а й для невідомої інформації. Вона гарно працює як для лінійно-розділимих даних, так і для лінійно-нерозділних. Нейронні мережі використовуються у багатьох областях [36]. Наведені вище переваги нейромережевої обробки даних визначають сфери їх застосування:

- обробка і аналіз зображень;
- розпізнавання мови незалежно від диктора;
- обробка високошвидкісних цифрових потоків;
- автоматизована система швидкого пошуку інформації;
- класифікація інформації в реальному масштабі часу;
- планування, застосування сил і засобів у великих масштабах;
- вирішення трудомістких задач оптимізації;
- адаптивне управління і передбачення.

В наш час основними напрямками реалізації мереж є:

- програмна реалізація на цифрових ЕОМ традиційної архітектури;
- програмно-апаратна реалізація у вигляді співпроцесорів до ЕОМ загального призначення;
- апаратна реалізація шляхом створення нейрокомп'ютерів на базі нейроплат у вигляді паралельних нейроподібних структур.

Ранні варіанти реалізації нейронних мереж відносяться до перших двох із вказаних напрямів. Перший напрям характеризується універсальністю, дешевизною і низькою швидкістю навчання і функціонування нейронних мереж. Для другого напрямку характерна висока швидкість моделювання функціонування мереж, але при цьому існують серйозні фізичні обмеження числа модельованих елементів і зв'язків між ними, а також можливостей навчання і донавчання. Із розвитком елементної бази ЕОМ став можливим самостійний розвиток третього напрямку, який поклав початок індустрії нейрокомп'ютерів, що подають сукупність апаратних і програмних засобів для реалізації моделей нейронних мереж [36].

2.2 Конвуляційні нейронні мережі

Проблеми, що виникли в процесі використання моделей, які формують цілісні репрезентації, сприяли розвитку нової групи алгоритмів, що використовують локальні ознаки зображень. Необхідність такого підходу була продиктована властивістю стаціонарності природних зображень - об'єкти, присутні на зображенні, могли вільно переміщатися в межах поля зору, при цьому бажаним результатом розпізнає алгоритму залишалось співвіднесення безлічі таких інваріантних репрезентацій об'єкта до одного класу [36].

Крім іншого, використання локальних ознак при розпізнаванні зображень було підкріплено свідченнями з області нейробіології. У класичній праці Д. Хьюбела і Т. Візела [42], що візуальна кора головного мозку являє собою складний комплекс клітин, кожна з яких чутлива тільки до обмеженого ділянки

поля зору. Такі ділянки, інакше звані рецептивних полями, стикуються разом, забезпечуючи перекриття всього поля зору. Відповідні клітини при цьому виконують роль локальних фільтрів вхідних даних, реагуючи на присутність у власному рецептивній поле деяких примітивних структур, таких як краю і кордони. Було виявлено також існування так званих «складних клітин», мають більш широкі рецептивні поля, і демонстрували інваріантність по відношенню до точного розташування об'єкта в полі зору [36].

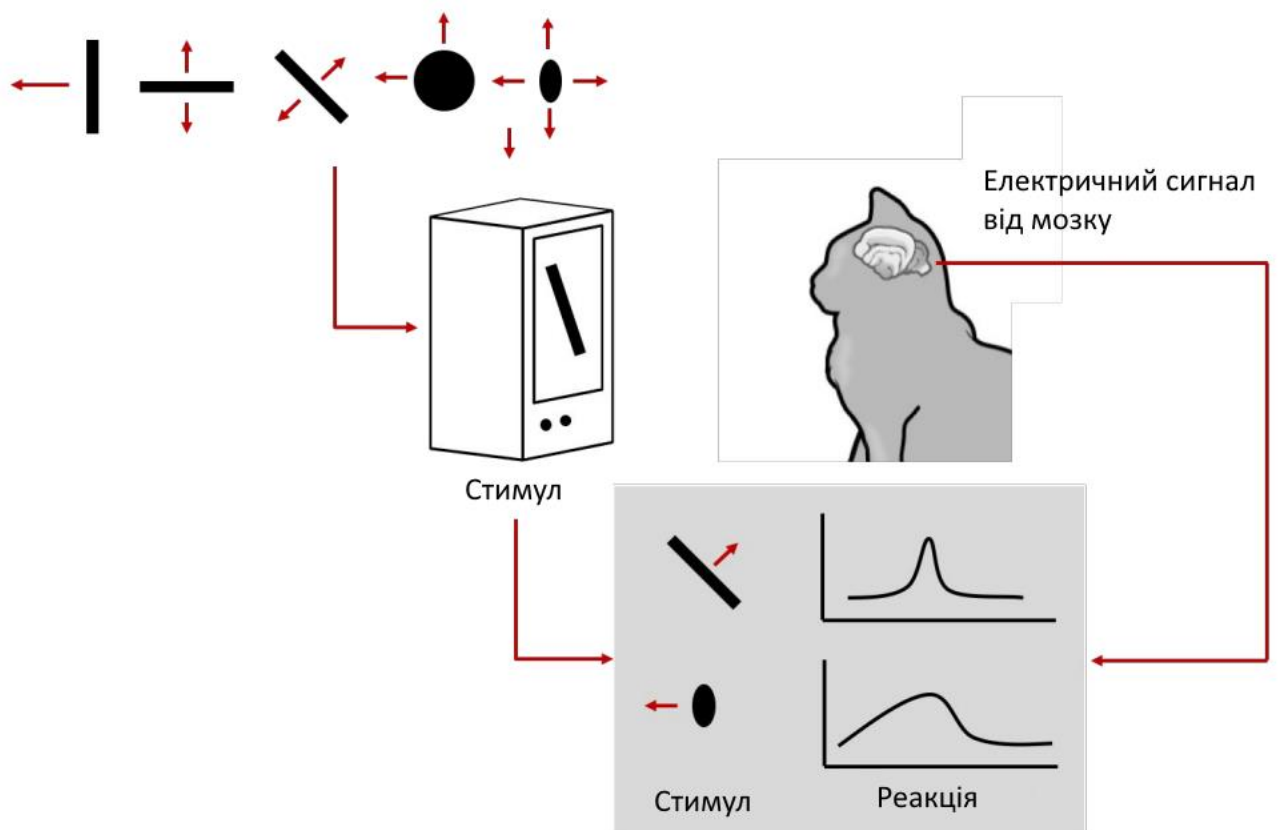


Рис. 2.4 - Експеримент Хьюбела і Візела

З урахуванням того, що візуальна кора головного мозку являє собою найбільш потужну і гнучку зорову систему з існуючих на даний момент, поява моделей [43,44,45], емулюючих її поведінку, виглядало природнім кроком. Однією з найбільш успішних моделей, що вважається визнаним лідером [46] в області розпізнавання зображень, є конвуляційна нейронна мережа [36].

Згорткові мережі являють собою варіацію архітектури багат шарового перцептрона, і включають в себе згорткові шари, шари підвибірки

(субдискретизація), і повнозв'язні шари. Архітектура конвуляційної мережі використовує переваги двовимірної структури вхідних даних - зображень за допомогою методу локальної зв'язності, обмежуючи кількість зв'язків між нейронами прихованого конвуляційного шару і вхідними даними. Конкретно, кожен нейрон прихованого шару пов'язаний тільки з обмеженою локальною ділянкою зображення [36].

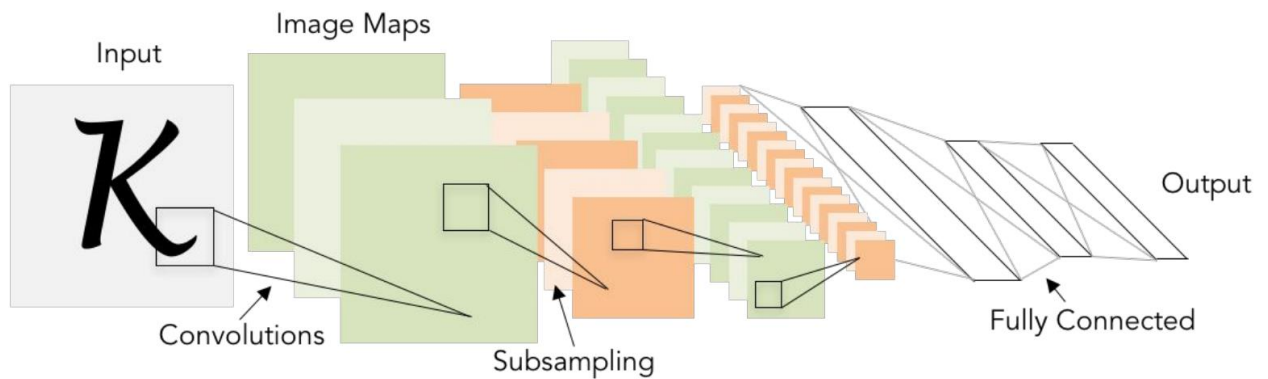


Рисунок 2.5- Архітектура конвуляційної нейронної мережі [53]

Крім цього, нейронна мережа використовує загальні, або розподілені вагові коефіцієнти, накладаючи штучне обмеження на алгоритм навчання зворотним поширенням помилки, так, щоб кожен нейрон прихованого шару мав набір ваг, спільний з іншими нейронами цього шару. При прямому поширенні така мережа здійснює математичну операцію згортки вхідного зображення набором фільтрів, які подаються вагами нейронів прихованого шару. Проміжними результатами мережі є так звані «карти ознак» - двовимірні матриці, що представляють собою результат згортки окремим фільтром [36].

Шар субдискретизації виконує операцію групування мап ознак, розглядаючи регіони $p * p$ і агрегуючи значення, отримані в результаті згортки. Основне призначення субдискретизації – знизити варіативність даних, забезпечуючи стійкість к трансляціям локальної ознаки в рамках даного регіону. Таким чином, в умовах, коли одна і та сама ознака опиняється здвигнутою на деяке значення ($\Delta x, \Delta y$) в межах p , відповідний нейрон, що інкапсулює локальну субдискретизовану ознаку зображення, буде все ще активним. Таким чином

мережа забезпечує інваріантність до просторових спотворень. В якості агрегуючої функції шару субдискретизації зазвичай розглядається функція вирахування середнього чи максимального значення [36].

Таким чином, значення за рахунок використання декількох поперемінних шарів згортки і субдискретизації конвуляційна нейронна мережа дозволяє отримувати уявлення, незалежні від конкретного розташування локального ознаки в зображенні, і однаково чинно реагувати на об'єкти, що цікавлять (наприклад, людські обличчя), присутні на будь-якій ділянці фотографії [36].

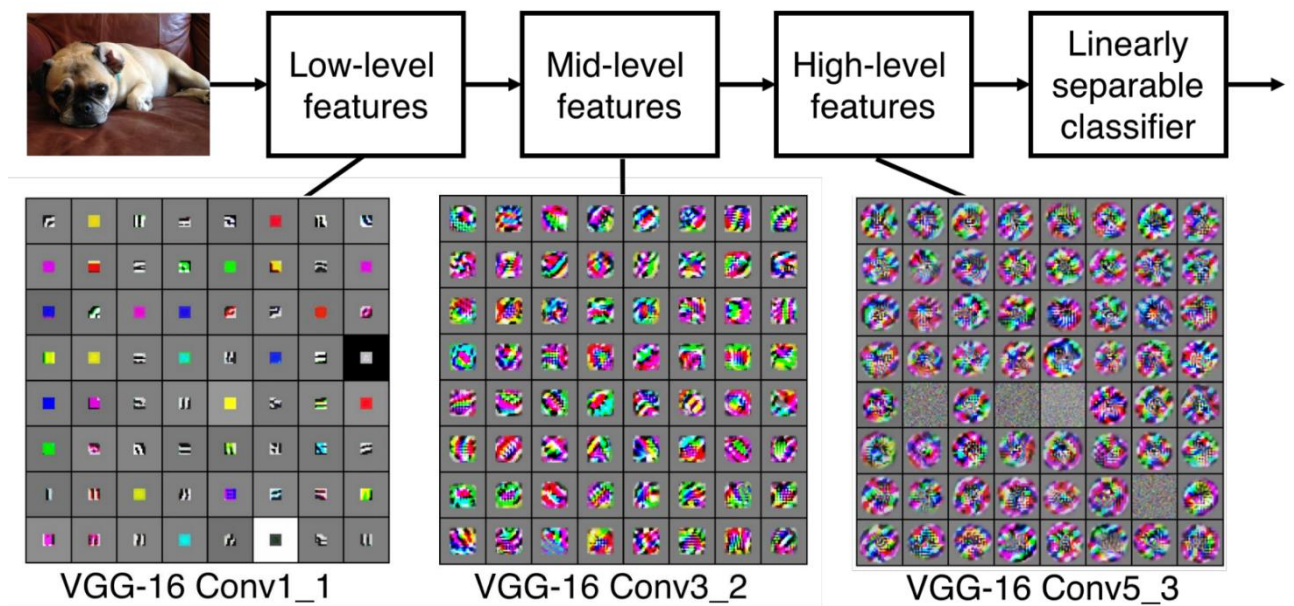


Рисунок 2.6 - Візуалізація активацій фільтрів конвуляційної нейронної мережі[53]

Згорткові мережі - один з найбільш успішних існуючих на сьогоднішній день підходів до розпізнавання зображень. Моделям, які реалізують відповідну архітектуру, належать перші місця в змаганнях алгоритмів розпізнавання, таких як ImageNet: згорткові мережі розпізнають рукописні цифри вибірки MNIST з 0.23% помилок [47], людські обличчя з 2.4% помилок [48] (використовувалася вибірка розміром в 5600 фотографій обличчя більш ніж 10 осіб); результат змагання ImageNet для переможця 2014 року - конвуляційної мережі GoogLeNet компанії Google становить 0.06656% помилкових розпізнавань [49], що являє собою найкращий результат на поточний момент і можна порівняти з помилкою, що допускається людиною на вибірці ImageNet. Серед недоліків виділяють

труднощі при обробці маленьких об'єктів, і нездатність справлятися з спотвореннями, такими як розмиття або сильний шум (такі спотворення присутні в навколишньому світі, наприклад, при погляді через товсте скло). При цьому згорткові мережі порівняно легко справляються з проблемами високоточного розпізнавання, які викликають труднощі у людей - наприклад, розпізнавання окремих моделей машин або порід собак, і інші завдання, що вимагають виділення вузькоспецифічних ознак [36].

Операція згортки забезпечує інваріантність по відношенню до трансляції локальних ознак (здвигу по осях x і y). Архітектура конвуляційних мереж не передбачає стійкості до інших афінних перетворень, таким як обертання, дзеркальне відображення і масштабування. Для вирішення цієї проблеми, як правило, використовуються евристичні методи (вирівнювання зображення по лінії горизонту, використання просторових пірамід і різних відображених копій оригіналу) [50]. Перетворення, які не є афінними по відношенню до плоского зображення, такі як обертання об'єкта в тривимірній площині, аналогічним чином не можуть бути оброблені за допомогою операції згортки. Для об'єктів, представлених в різних орієнтаціях, моделі доводиться навчатися різним, непересічним наборам ознак [51]. Згорткові нейронні мережі та інші підходи, які використовують локальні інваріантні ознаки, в процесі розпізнавання періодично вимушено навчаються ознакам, максимізуючим інваріантність в рамках представленої вибірки, але не забезпечують дискримінативного вигляду об'єкта в цілому [52].

Альтернативні підходи до виділення локальних ознак включають в себе методи класичного комп'ютерного зору, які не використовують навчальні моделі. Ці методи здійснюють пошук на зображенні характерних ділянок, що відповідають алгоритмічно явно заданим умовам.

2.3 Функції активації

Кожна функція активації додає нелінійність в навчання нейронної мережі. Вона бере окреме число, що є результатом роботи попереднього шару і виконує фіксовану математичну операцію з ним табл.2.1.

Таблиця 2.1 – Класифікація функцій активацій [54]

Назва	Формула	Область значень
Порогова	$\Psi = F(\psi) = \begin{cases} 0, \psi < \theta, \\ 1, \psi \geq \theta. \end{cases}$	$(0, 1)$
Знакова	$\Psi = F(\psi) = \begin{cases} 1, \psi > 0, \\ -1, \psi \leq 0. \end{cases}$	$(-1, 1)$
Сигмоїдна	$\Psi = F(\psi) = \frac{1}{1 + e^{-\psi}}$	$(0, 1)$
ReLU	$\Psi = F(\psi) = \begin{cases} \psi, \psi > 0, \\ 0, \psi \leq 0. \end{cases}$	$(0, \infty)$
Лінійна	$F(\psi) = \psi$	$(-\infty, \infty)$
Радиальна базисна	$\Psi = F(\psi) = e^{-\psi^2}$	$(0, 1)$
Напівлінійна	$\Psi = F(\psi) = \begin{cases} 0, \psi \leq 0, \\ \psi, 0 < \psi < 1, \\ 1, \psi \geq 1. \end{cases}$	$(0, 1)$
Линейная з насиченням	$\Psi = F(\psi) = \begin{cases} -1, \psi \leq -1, \\ \psi, -1 < \psi < 1, \\ 1, \psi \geq 1. \end{cases}$	$(-1, 1)$
Tanh	$\Psi = F(\psi) = \frac{e^{\psi} - e^{-\psi}}{e^{\psi} + e^{-\psi}}$	$(-1, 1)$
Трикутна	$\Psi = F(\psi) = \begin{cases} 1 - \psi , \psi \leq 1, \\ 0, \psi > 1. \end{cases}$	$(0, 1)$

Найбільш часто у практиці зустрічаються наступні:

Сигмоїдна. Суть її роботи полягає в тому, що вона бере значення і «втискає» його в межі від 0 до 1. По суті, великі негативні числа стають близькими до 0, а великі позитивні ближче до 1.

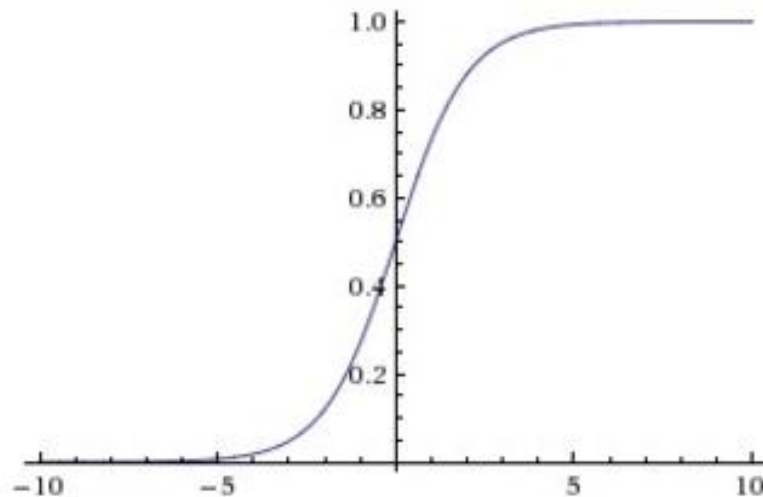


Рисунок 2.7 - Графік сигмоїдної функції [53]

Сигмоїдна функція історично часто використовується через подібність до рівня активації нейрона: від зовсім неактивного (0), до повністю активованого (1). На практиці, сигмоїдна нелінійність не так давно вийшла з моди і досить рідко використовується. Вона має два суттєві недоліки:

- Сигмоїда подавляє та вбиває градієнти. Зовсім непотрібна особливість сигмоїдного нейрону полягає в тому, що коли активізація близька до обох «кінців» біля 1 або 0, градієнт у цих областях близький до 0. Під час зворотного розповсюдження похибки локальний градієнт буде помножений на вихідний. Таким чином, якщо локальний градієнт дуже малий, то він зробить і глобальний дуже малим, і розповсюдження помилки також буде малим, рекурсивно поширюючи цю проблему. Окрім того, слід слідкувати за ініціалізацією вагових коефіцієнтів.
- Вихід сигмоїди не централізований відносно нуля. Це є проблемою, оскільки нейрони у глибших шарах нейронної мережі отриматимуть не також не центровані відносно нуля значення. Це матиме вплив на динаміку під час градієнтного спуску, тому що дані на вході нейрону будуть завжди

позитивними, тоді градієнт вагових коефіцієнтів буде або цілком позитивним, або цілком негативним. Це призведе не зиг-загообразної динаміки оновлень градієнту. Хоча ця проблема може бути вирішеною використанням батчів, під час яких градієнти сумуються, а лише потім йде оновлення вагових коефіцієнтів, все одно це призводить до незгучностей [53].

Tanh. Tanh нелінійність «стискає» дійсне число у проміжок $(-1,1)$. Подібно до сигмоїдної функції її активації демонструють насичення, але вона є центрованою відносно 0. На практиці завжди обирають Tanh у порівнянні з сигмоїдною функцією. Окрім того, Tanh фактично видозмінена сигмоїда.

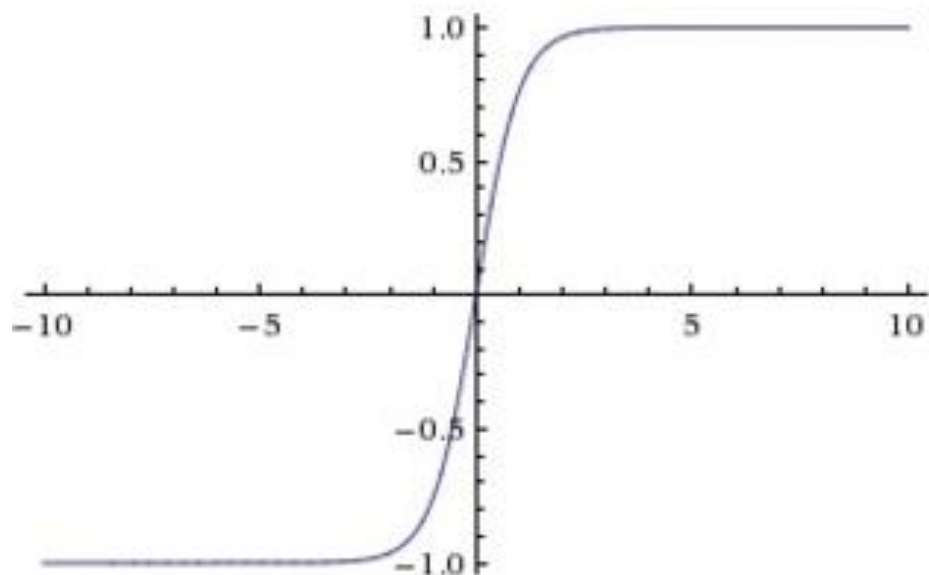


Рисунок 2.8 - Графік Tanh [53]

ReLU. Ця функція активації стала дуже популярною в останні роки. Фактично вона робить рівними нулю всі значення менше нуля. Вона має декілька переваг та недоліків:

- Вона значно пришвидшує швидкість стохастичного градієнтного спуску у порівнянні з сигмоїдною або Tanh функціями [53].
- У порівнянні з сигмоїдною або Tanh функціями, що містять складні операції(експонента і т.д.), ReLU легко імплементується та швидко вираховується [53].

– Нажаль, ReLU об'єкти можуть бути хрупкими і вбивати навчання. Наприклад, великий градієнт, що проходить через ReLU нейрон може спричинити оновлення вагових коефіцієнтів у такий спосіб, що нейрон не активується за жодного набору даних. Якщо це виникає, тоді градієнт, що проходить через цей нейрон завжди буде рівний нулю. Через цей ефект ви можете з'ясувати, що 40% вашої мережі не активізується, якщо ваша інтенсивність навчання буде занадто високою. За правильних налаштувань навчання ця проблема зустрічатиметься рідше[53].

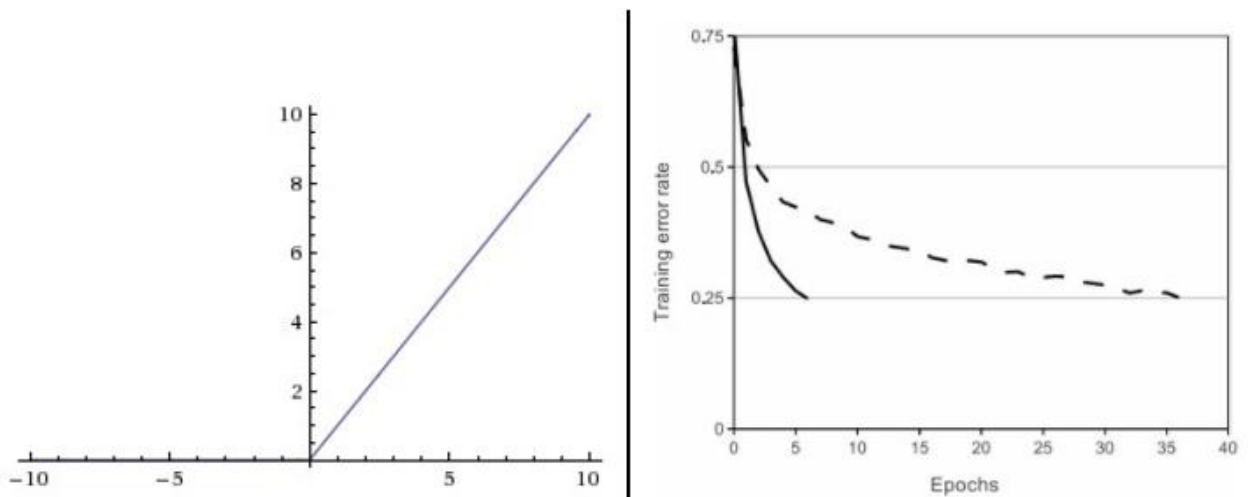


Рисунок 2.9 - Ліворуч: графік ReLU. Праворуч: графік що демонструє швидкість навчання у порівнянні з Tanh [53]

Leaky ReLU. Спроба вирішити проблему вмираючого ReLU. Замість присвоювання 0 від'ємним значенням, Leaky ReLU матиме дуже малий коефіцієнт, на який значення буде домножуватися. Формула:

$$f(x) = 1(x < 0)(\alpha x) + 1(x \geq 0)(x)$$

Деякі дослідники кажуть, що досягли успішних результатів з використанням цієї функції, але результати не завжди вдалі[53].

Висновки

Нейронні мережі зарекомендували себе як потужний механізм для класифікації зображень, але протягом останнього часу вони активно використовуються і в інших областях комп'ютерного зору. Їм належать перші місця в змаганнях алгоритмів розпізнавання, таких як ImageNet, а якість розпізнавання близька, а іноді і перевищує людську.

Окрім того, конвуляційні мережі відносно легко вирішують проблеми високоточного розпізнавання, що визивають проблеми у людей, наприклад, визначення породи собаки чи марки автомобіля, та інші задачі, що потребують виділення вузькоспецифічних ознак.

Серед недоліків можна виділити проблеми при обробці елементів малого розміру, і відсутність можливості справлятися зі спотвореннями, такими як розмиваючий фільтр або шум. Також недоліком є проблема вибору набору даних, який задовольнить проблему, що вирішується та буде досить якісним і великим для навчання нейронної мережі.

3 АНАЛІЗ МЕТОДІВ РОЗПІЗНАВАННЯ ОБЛИЧ

3.1 Структура системи розпізнавання облич

Будь-який з розглянутих алгоритм розпізнавання облич базується на архітектурі, в основі якої наступні чотири етапи:

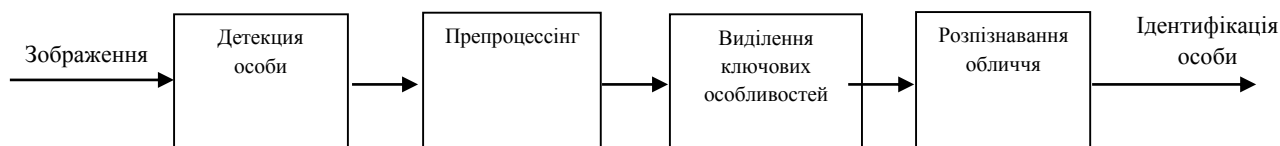


Рисунок 3.1 – Архітектура алгоритмів розпізнавання обличчя [55]

Першим кроком є детекція особи з вхідного зображення. Суть етапу полягає в підтвердженні чи має вхідне зображення обличчя і локалізація цієї області. Після цього ця область подається на препроцесінг.

Під час препроцесінгу зображення позбавляється небажаного шуму, розмиття, різних умов освітленості, тіней за допомогою різних технік. Необхідність його застосування визначається підходом, використовуваним для виділення ключових особливостей.

На третьому етапі відбувається виділення ключових особливостей за допомогою обраного для цього підходу. Результатом цього найчастіше є вектор фіксованого дозволу або набір певних точок з відповідним розташуванням.

На четвертому етапі відбувається аналіз відображень кожної особи. Тут відбувається ідентифікація особи на основі наявної бази даних[55].

Порівняння середньої точності на різних наборах даних наведених нижче методів і інших популярних підходів [55].

Таблиця 3.1 - Порівняння середньої точності популярних підходів ідентифікації обличчя [55]

PCA with ANN Face recognition system	95.45%
DDFD	91.79%
RBNNs	97.56%
CNN	85.1%
B-CNN	95.3%
BPN+RBF	98.88%
RCNN	90.3%
RINN	90.6%
MRC & MLP Neural Network	91.6%
Gabor Wavelet Faces with ANN	93%
WNN	89.22%

3.2 PCA

Метод головних компонент (PCA) - математичний метод, який перетворює ряд ймовірно-корельованих змінних в меншу кількість некоррелірованих змінних, які називаються основними компонентами. PCA є дуже відомим підходом, який використовується для розрахунку набору функцій для розпізнавання особи. Це може відноситися до будь-якої особи:

1. Особа може бути представлено мінімальним набором характеристик[56].
2. Може бути побудовано з використання невеликої колекції власних фотографій. Зображення особи проектується на деяку кількість шаблонів осіб, зване власними обличчями, потім обчислюється різниця між цими обличчями, яка буде вважатися характерною рисою для зображення. Коли набір власних обличчя обчислений, зображення особи може бути реконструйовано зваженою комбінацією власних обличчя. Ваги утворюють вектор ознак для подання обличчя і розпізнавання. Коли подається нове тестове зображення, ваги вираховані

завдяки проекції зображення на вектор власних осей. Класифікація між зображеннями виходить з порівняння відстаней між векторами ваг тестового зображення і навчальної вибірки (вхідних даних). Є можливість реконструювати вихідне зображення з власних облич так, що вхідне зображення повинно точно відповідати оригінальному зображенню, використовуючи всі власні особи, витягнуті з оригінальних зображень. PCA є математичним методом, який заснований на перетворенні змінних, за допомогою якого можна перетворити набір корельованих змінних в набір некоррелірованих змінних. Змінні, які некоррелюють один з одним - основні компоненти. Кількість основних компонентів означає, що некоррелірованих має бути менше або дорівнює кількості вихідних змінних. У такого роду трансформації перші головні компоненти будуть мати найвищий пріоритет, який покаже максимальне відхилення. Це допоможе обчислити точність даних. Якщо набір даних розподілено нормально, тільки тоді головні компоненти можуть бути незалежними від інших змінних [56].

Алгоритм PCA

Нехай є такий набір осіб для навчання: $\Gamma_1 \Gamma_2 \Gamma_3 \dots \dots \Gamma_M$. Середнє обличчя набору може бути визначено[56]:

$$\psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n$$

Кожне обличчя відрізняється від середнього вектором ознак[56]:

$$\Phi_i = \Gamma_i - \psi$$

Матриця коваріацій будується завдяки вектору[56]:

$$C = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T = AA^T$$

Де матриця $A = (\Phi_1 \Phi_2 \Phi_3 \dots \dots \Phi_M)$. До множини великих векторів застосовують PCA $U_1 U_2 \dots \dots U_M$. Отримуємо вектор ваг Ω зображення обличчя. Обличчя проєцирується на простір облич[56]:

$$W_k = U_k^T (\Gamma - \psi)$$

Для $k = 1, M$, де $M' \leq M$ номер власного обличчя, яке використовується для ідентифікації. Ваги формують вектор $\Omega = (w_1, w_2, w_3 \dots \dots w_M)$ [56].

Навчальний процес багатошарової нейромережі зі зворотним поширенням помилки вимагає набори вхідних і бажаних вихідних даних. Вихідний вектор "о" порівнюється з бажаним вихідним. У разі виявлення різниці між цими векторами, ваги підлаштовуються для досягнення мінімізації різниці. Спочатку випадкові ваги підлаштовуються до нейромережі. Ці ваги оновлюються під час кожної ітерації з метою мінімізувати середньоквадратичну помилку між вихідним і бажаним вихідним векторами. Вхідні дані для прихованого шару надаються таким чином[56]:

$$\text{net}_m = \sum_{z=1}^n x_z w_{mz}$$

Значення вихідного вектора прихованого шару після проходження функції активації надаються таким чином[56]:

$$h_m = \frac{1}{1 + \exp(-\text{net}_m)}$$

і значення вихідного вектора вихідного шару надаються таким чином[56]:

$$o_k = \frac{1}{1 + \exp(-\text{net}_k)}$$

Для поновлення ваг необхідно обчислити помилку таким чином[56]:

$$E = \frac{1}{2} \sum_{i=1}^k (o_i - t_i)^2$$

Це становить реальний вихід і бажаний на і-тому нейрону вихідного шару. Якщо помилка менше заданого порогу, навчання припиняється, інакше ваги повинні бути оновлені. Для ваг між прихованим і вихідним шарами зміна ваг відбувається так[56]:

$$\Delta w_{ij} = \alpha \delta_i h_j$$

Де швидкість навчання в межах $[0.01, 1.0]$, вихід з нейрона j в прихованому шарі, може бути отримана таким чином[56]:

$$\delta_i = (t_i - o_i) o_i (1 - o_i)$$

Аналогічно, зміна ваг між прихованим і вихідним шарами відбувається таким чином [56]:

$$\Delta w_{ij} = \beta \delta_{ni} x_j,$$

де β - коефіцієнт швидкості навчання в межах $[0.01, 1.0]$, вихід з нейрона j у вхідному шарі, може бути отримана таким чином [56]:

$$\delta_{ni} = x_i(1 - x_i) \sum_{j=1}^k \delta_j w_{ij},$$

вихід на нейроні i у вхідному шарі, і сумарно становить зважену суму всіх значень, що відповідають нейронам вихідного шару, отриманих в обчисленні. Після обчислень вага змінюється на всіх шарах, вона може бути оновлена таким чином [56]:

$$w_{ij}(\text{new}) = w_{ij}(\text{old}) + \Delta w_{ij}$$

Цей процес повторюється поки похибка не буде мінімізована [56].

3.3 PCA з використанням штучних нейронних мереж

PCA це метод зменшення розмірності заснований на перетворенні Карунена-Лоєва. У задачі розпізнавання осіб його застосовує головним чином для представлення зображення особи за допомогою вектора малої розмірності (головні компоненти), який порівнюється з еталонними векторами, закладеними у базу даних [57].

Головною метою методу головних компонент є значне зменшення розмірності простору ознак таким чином, щоб вони якомога краще описували "типові" риси, які притаманні великій кількості осіб. Використовуючи цей метод можна виявити різні особливості в навчальній вибірці зображень облич і описати ці особливості в базисі декількох ортогональних векторів, які називаються власними [55].

В даному підході для кожної особи відбувається навчання нейронної мережі на основі власного обличчя, яка в подальшому буде визначати схожість

осіб. Метод має високу точність, але недоліком є те, що для будь-якої людини потрібна своя нейронна мережа [55].

3.4 Deep Dense Face Detector

Даний алгоритм базується на глибокій згортковій нейронній мережі і використовується для детекції осіб. Основними ідеями є висока здатність мереж в класифікації і виділення особливостей для побудови єдиного класифікатора для детекції осіб і мінімізація обчислювальної складності за рахунок спрощення архітектури детектора. Автор методу використовував архітектуру AlexNet і навчив її для завдання детекції саме облич. Обличчя одного і того ж класу (людини) в різних умовах були навчальною вибіркою. Їхні обличчя були модифіковані за допомогою різних методів аугментації даних для збільшення розмірів самої вибірки [58].

Детектор аналізує оцінки DDFD детектора і шукає кореляції між цими оцінками і розподілом позитивних прикладів в навчальній вибірці. Метод показав себе добре в багатьох випадках і умовах, був стійкий до різних типів освітлення, хоча відсоток точності розпізнавання значно впав при сильному повороті особи. Хоча автор вказав можливі шляхи вирішення даної проблеми за допомогою різних підходів до роботи з вхідними даними. Недоліком є необхідність повного перенавчання мережі для додавання нових осіб [55].

3.5 Radial Basis Function Neural Networks

Даний підхід відмінно себе зарекомендував і має одні з найвищих показників в точності. Як частина препроцесінга, використовується гібридний підхід на базі Моделі Активної Форми (Active Shape Models) і PCA (Principal Component Analysis). За допомогою ASM відбувається виділення контуру та

форм особи для створення персонального профілю, а після використовується PCA алгоритм, щоб зменшити розмірність відображень осіб [55]. У розпізнавальній частині використовуються Radial Basis Function Neural Networks (RBF NN) для ідентифікації унікального шаблону асоційованого з кожною людиною. Така архітектура складається з трьох модулів займаються фазою стану, фазою виведення і фазою укладення, яка працює за допомогою правил нечіткої логіки в замість стандартних if-then умов[59].

3.6 Convolutional Neural Network Cascade

Naohiang Liy, Zhe Linz, Xiaohui Shenz, Jonathan Brandtz, Gang Hua [60] запропонували згорткову нейронну мережу «Cascade» для розпізнання осіб. Цей метод має потужні здатності розділення ознак, зберігаючи при цьому високу продуктивність. Запропонований CNN каскад швидко відкидає фонові області на швидких етапах з низьким дозволом, і ретельно оцінює невелику кількість кандидатів в останній стадії. Для підвищення ефективності, а також зменшення кількості кандидатів на більш пізніх етапах, автор ввів етап калібрування CNN. У результаті генерування калібрувальної сітки отримується точніша локалізація обличчя, використовуючи більші вікна сканування по меншій кількості параметрів [55].

3.7 Bilinear CNNs

Aruni RoyChowdhury Tsung-Yu Lin Subhranshu Maji Erik Learned-Miller [61] запропонували метод Bilinear CNN для ідентифікації особи, який показав різке зростання продуктивності на деяких дрібнозернистих зразках. Модель ліквідує розрив між текстурами і деталями на основі алгоритмів CNN (зазвичай використовується дві CNN).

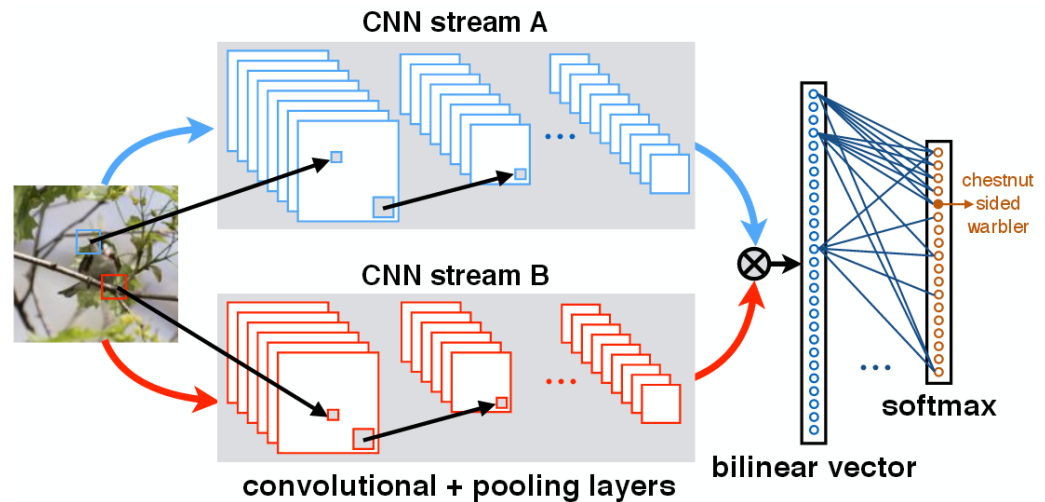


Рисунок 3.2 - Структура Bilinear CNNs [57]

Архітектура являє собою орієнтований ациклічний граф (DAG), обидва мережі можуть бути навчені одночасно за допомогою градієнтного back-propagating. Замість того, щоб навчати CNN для розпізнавання осіб з нуля, що потребує пошук оптимальної архітектури нейронної мережі та масивні бази даних, B-CNN може використовувати заздалегідь підготовлені мережі і адаптувати їх до задачі розпізнавання особи.

3.8 Back Propagation Network та Radial Basis Function Network

M.Nandini, P.Bhargavi, G.Raja Sekhar [62] запропонували комбінацію Back Propagation Network та Radial Basis Function Network для розв'язку задачі розпізнавання осіб. Цей метод пропонує новий підхід для розпізнавання людських облич. Розпізнавання здійснюється шляхом порівняння характеристик нового обличчя з уже відомими. Навчання нейронної мережі (NN) здійснюється за допомогою Back Propagation і Radial Basis Function (RBF) мереж. Radial Basis Function нейронні мережі є вельми привабливими для інженерних завдань. Вони мають дуже компактну топологію, універсальні наближення, дуже велику швидкість навчання. Нейронна мережа Back Propagation має трьохшарову архітектуру (вхідний шар, прихований шар і вихідний шар). RBF використовується як ідентифікатор у системі розпізнавання особи і входить в цій

мережі є результати, отримані з VPN. Точність розпізнавання за допомогою такої моделі дуже висока. Що дає можливість розширити і використовувати цей метод для рухомих зображень і зображень з різних фоном [55].

3.9 Retinal Connected Neural Network (RCNN)

Henry A. Rowley, студент [63] запропонував, RCNN, що досліджує невеликі вікна зображення і визначає, чи містить кожне вікно людину. Перший компонент системи є фільтр, який приймає в якості вхідних даних пікселі області 20×20 зображення, і генерує вихідний сигнал в межах від 1 до -1, що означає наявність або відсутність особи, відповідно. Для виявлення особи в будь-якому місці на вході, фільтр застосовується в кожній точці в зображенні. Для виявлення особи, розмір яких перевищує розмір вікна, вхідне зображення багаторазово зменшується. Цей фільтр повинен мати деяку інваріантність в положенні і масштабі. Кількість інваріантності визначає число ваг і положень, в яких він повинен бути застосований [55].

Основним недоліком існуючої системи є те, що вона виявляє тільки в вертикальному положенні особи, дивлячись на камеру. Окремі версії систем можуть бути навчені для кожної орієнтації голови, і результати можуть бути об'єднані за допомогою арбітражних методів [55].

3.10 Rotation Invariant Neural Network (RINN)

Henry A. Rowley, Shumeet Baluja, Takeo Kanade [64] запропонував обертальний інваріантний метод навчання нейронної мережі для системи розпізнавання осіб, яка використовується для ідентифікації об'єктів будь-якого ступеня обертання в площині зображення. На відміну від своїх аналогів, які розпізнають об'єкти тільки у вертикальному положенні. розпізнає обличчя в будь-якого ступеня обертання в площині зображення [55].

Висновки

Був проведений аналіз різних методів у розпізнаванні облич. Окремо можна виділити загальну структуру, що притаманна всім методам. Першим кроком зазвичай є детекція особи на вхідному зображенні. Суть етапу полягає в підтвердженні чи має вхідне зображення обличчя і локалізація цієї області. Після цього ця область подається на препроцесінг. Під час препроцесінгу зображення позбавляється небажаного шуму, розмиття, різних умов освітленості, тіней за допомогою різних технік. Необхідність його застосування визначається підходом, використовуваним для виділення ключових особливостей. На третьому етапі відбувається виділення ключових особливостей за допомогою обраного для цього підходу. Результатом цього найчастіше є вектор фіксованого дозволу або набір певних точок з відповідним розташуванням. На четвертому етапі відбувається аналіз відображень кожної особи.

У огляді методи на нейронних мережах суттєво краще демонструють себе в тестах і мають кращу точність, але тим не менш, важливо відмітити, що одні з найкращих методів є модифікаціями класичних підходів з використанням нейронних мереж.

Незважаючи на гарні результати на тестових даних, деякі алгоритми мають властивість погано демонструвати себе на інших наборах, або в реальних умовах. Тобто проблема універсальної системи все ще є відкритою.

4 РОЗРОБКА ДОДАТКУ ДЛЯ ГЕНЕРАЦІЇ СТРУКТУРНОЇ МОДЕЛІ ОБЛИЧЧЯ З ВИКОРИСТАННЯМ КОНВУЛЯЦІЙНИХ НЕЙРОННИХ МЕРЕЖ

4.1 Загальний алгоритм роботи додатку

Ідея написання програми полягала у тому, щоб написати систему для генерації структурної моделі обличчя, яка б використовувала конвуляційні нейронні мережі у своїй роботі, що мало покращити результати роботи у умовах реального світу. Тобто зробити його більш універсальним у порівнянні з додатками, що демонструють гарну роботу лише в штучних умовах.

Додаток складається з трьох модулів:

- Модуль, що займається детекцією облич.
- Модуль, що відповідає генерацію структурної моделі обличчя. Він вирішує локальні задачі, такі як розуміння пози обличчя, якості освітлення та проводить нормалізацію. Після цього він видає унікальні особливості, які можна використовувати для відрізнення однієї людини від іншої.
- Модуль, що шукає людину в базі даних та з певною впевненістю каже кому належить кожне обличчя.

На рисунку 4.1 можна побачити алгоритм роботи додатку для розпізнавання людини. Після отримання зображення для аналізу, в дію вступає модуль, що відповідає за пошук зображень на обличчі. В результаті роботи цього модуля ми знаємо кількість та розташування облич на зображенні. Аналіз обличчя та генерація структурної моделі наступним етапом. На основі даних отриманих від цього модуля нейронна мережа генерує особливості обличчя, які вже потім використовуються для генерації пропозиції. У ролі модуля для генерації пропозиції використовується система для розпізнавання облич, але можливі і інші варіанти використання, такі як розпізнавання емоцій, віку і т.д.



Рисунок 4.1 – Алгоритм роботи додатку

4.2 Пошук облич на зображенні

Перший крок у нашій послідовній системі робить система детекції облич. Очевидно, що нам необхідно знати розташування облич на зображенні, перед аналізом цих самих облич. Можлива відсутність облич на зображенні робить виклик наступних модулів марним, оскільки ми не знатимемо зон, на яких запускати їх.

Для вирішення задачі детекції було обрано архітектуру Faster R-CNN. У 2015 році вона перемогла у змаганні COCO. Останні модифікації Faster R-CNN

демонструють вражаючі результати в задачах класифікації (S. Ren, K. He, R. B. Girshick, and J. Sun. Faster R-CNN: towards real-time object detection with region proposal networks. In NIPS, pages 91–99, 2015). Тому було вирішено використати її для задачі детекції облич. В основу лягли преднавчані шари від VGG16, що навчена для класифікації об'єктів. Основною перевагою і особливістю Faster R-CNN є вдале поєднання швидкості та перенесення задачі класифікації на високоточну задачу детекції.

Ключовими аспектами Faster R-CNN є використання преднавчаної мережі для класифікації, на базі якої будується RPN, що складається з конвуляційних шарів 3×3 , два шари 1×1 додані для гілок класифікації та регресії. Розмір входу мережі відповідає VGG16 і становить 228×228 . Архітектура мережі наведена на рисунку 4.2.

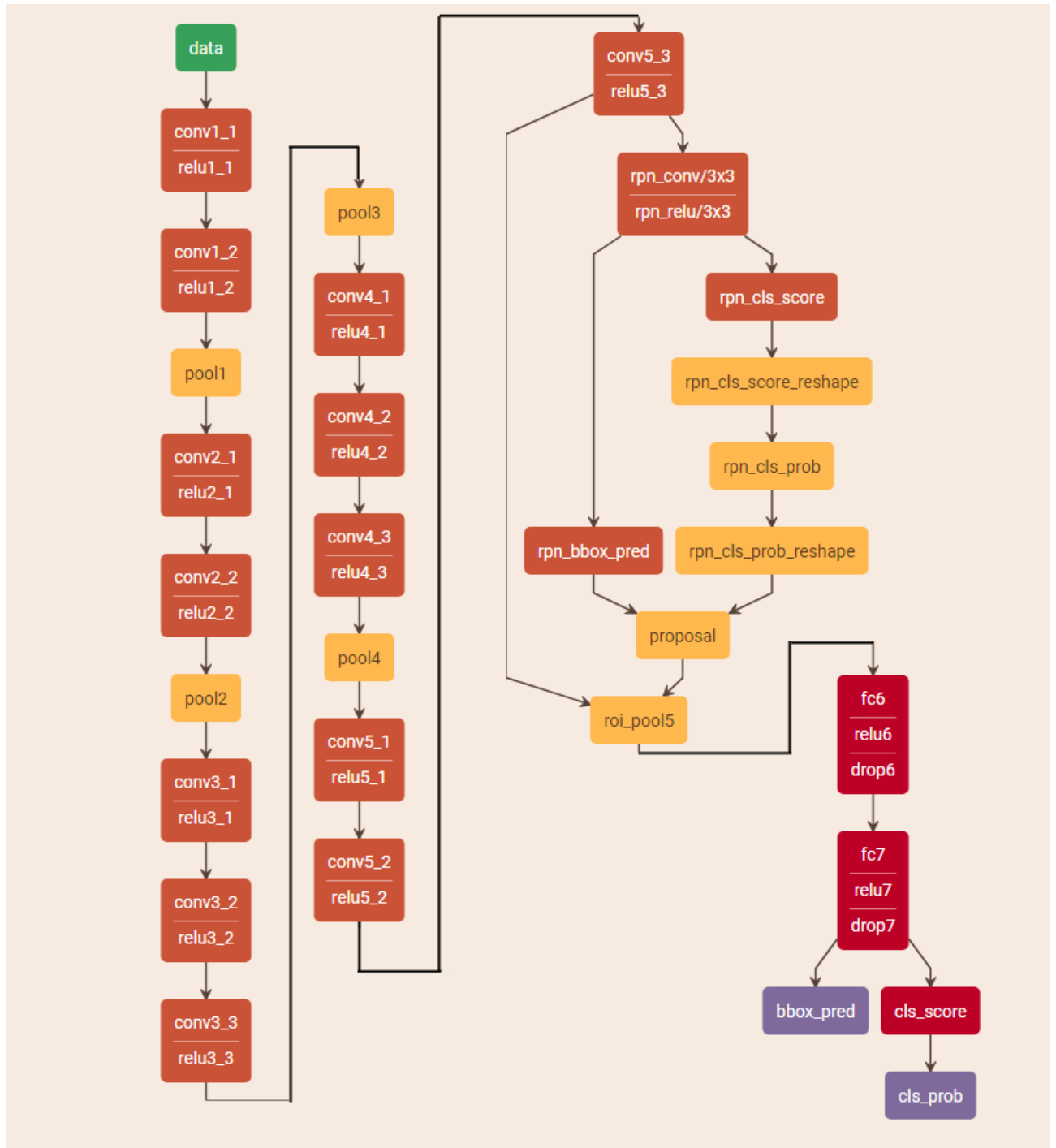


Рисунок 4.2 - Архітектура Faster-R-CNN для детекції облич

Для навчання був обраний датасет WIDER. Він містить 32,203 зображення з 393,703 розміченими людьми з великим різноманіттям масштабів, поз, оклюзій. Метрика евалюації аналогічна PASCAL VOC.



Рисунок 4.3 - Приклади зображень з датасету WIDER

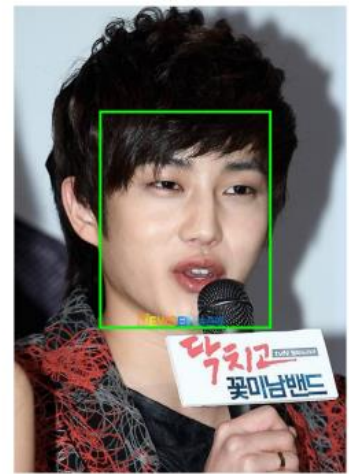
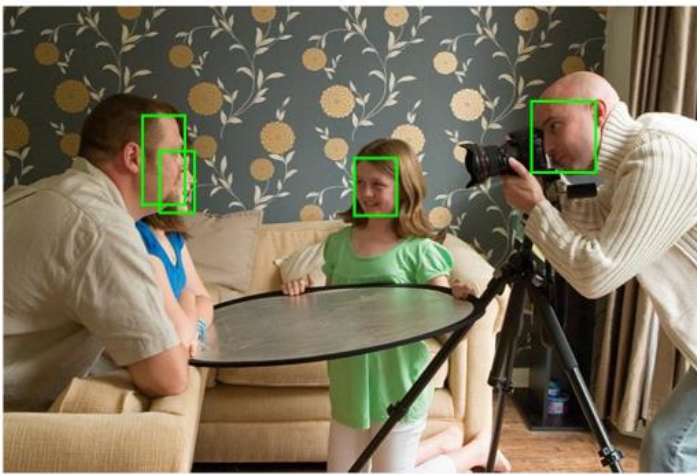


Рисунок 4.4 - Приклади зображень з датасету WIDER



Рисунок 4.5 - Приклади зображень з датасету WIDER

Детектор дає точність 95.2% на тестовій вибірці, що можна вважати гарним результатом на фоні інших state-of-the-art підходів. Недоліком слід вважати велику досить велику кількість false-positive помилок. Ця особливість є недоліком архітектури. Алгоритм NMS дозволяє суттєво зменшити кількість false-positive, але зазвичай він подавляє інколи і вірні детекції. Існують більш точні архітектури, але вони працюють значно довше. Процес тренування цієї архітектури займає приблизно 2 дні.

Досі для детекції обличчя використовують HOG методи, через їх швидкість та простоту.

4.3 Аналіз обличчя. Генерація структурної моделі

Отож, ми ізолювали обличчя на нашому зображенні. Але тепер нам необхідно розібратися з проблемою оберту і нахилу обличчя. Для цього ми будемо проводити нормалізацію зображення так, щоб очі та губи завжди були на одному місці. Це спростить задачу порівняння облич.

Для цього було використано алгоритм FLE [66]. Існує безліч декілька його реалізацій. Ідея полягає в тому, що ми визначаємо 68 специфічних точок, що характеризують обличчя. Схема цих точок зображена на рисунку 4.6

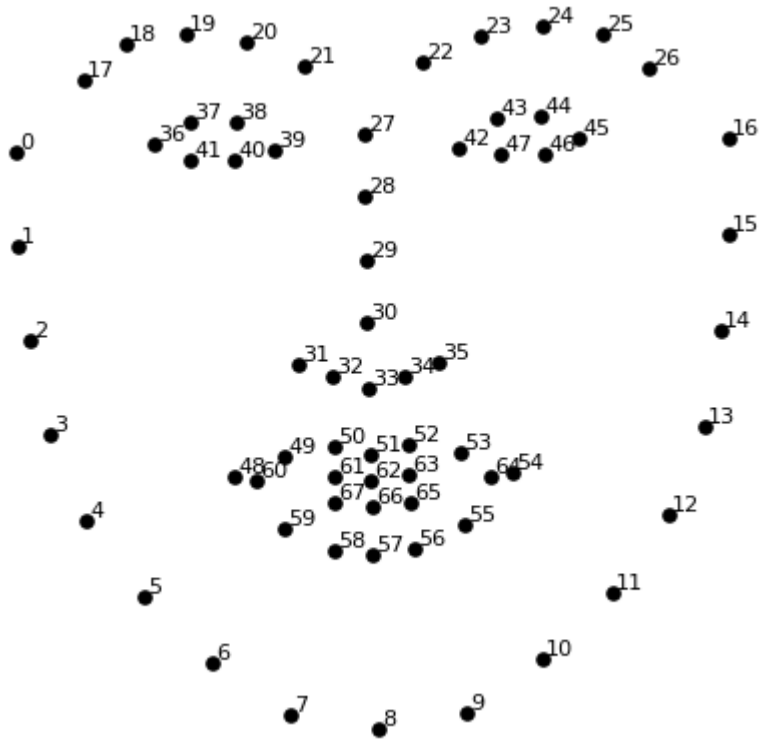


Рисунок 4.6 – 68 точок, що будуть визначені на кожному обличчі[67]

Тоді коли ми маємо позицію очей і рота ми просто «вирівнюємо» його. Не важливо куди було повернуто обличчя, ми грубо зможемо відцентрувати у приблизно необхідне положення за рахунок операцій оберту, масштабування і т.д. Нижче наведені приклади накладення розмітки на обличчя.

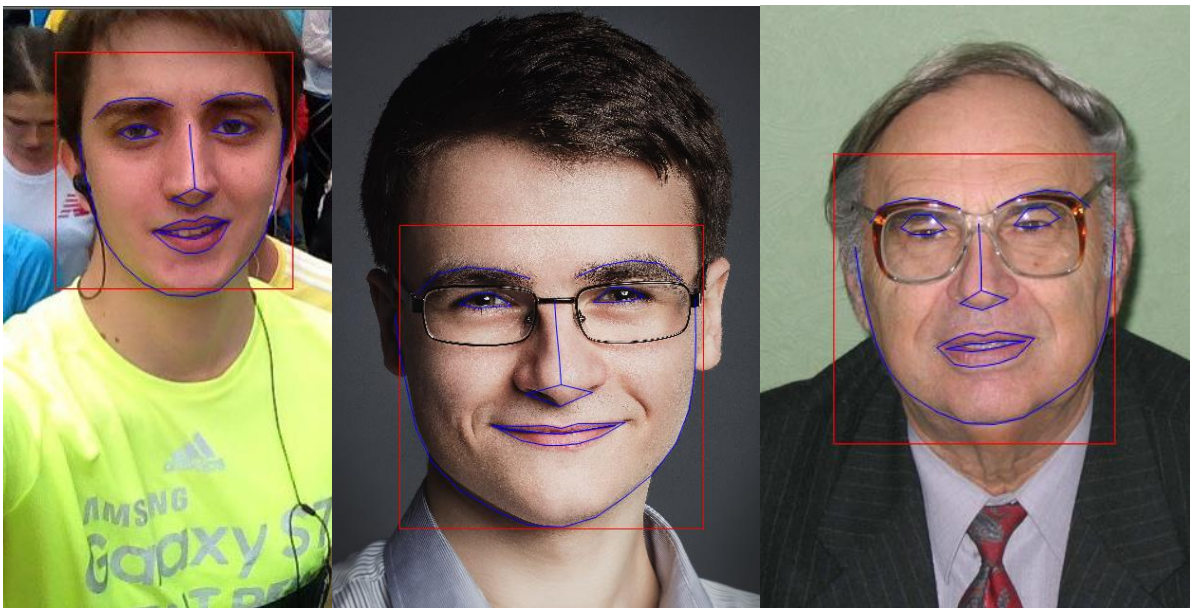


Рисунок 4.7 – Приклади роботи алгоритмів детекції та роботи FLE

Наступний крок полягає у аналізі отриманої розмітки. Для використання її для процесу розпізнавання людини ми маємо визначити які саме особливості нам слід порівнювати, оскільки пряме порівняння точок не є розумною ідеєю. Необхідний спосіб для визначення декількох базових метрик з кожного обличчя. З'ясовано, що метрики, такі як колір очей, не мають жодного змісту для комп'ютера. Тому для цих цілей ми можемо навчити нейронну мережу, яка б генерувала 128 вимірів.

Процес навчання полягає в порівнянні трьох зображень: два різних зображення однієї людини, та одне зображення абсолютно іншої.

Архітектура мережі:

```

nn.SpatialConvolutionMM: 1, 64, 48, 48
nn.SpatialBatchNormalization: 1, 64, 48, 48
nn.ReLU: 1, 64, 48, 48
nn.SpatialMaxPooling: 1, 64, 24, 24
nn.SpatialCrossMapLRN: 1, 64, 24, 24
nn.SpatialConvolutionMM: 1, 64, 24, 24
nn.SpatialBatchNormalization: 1, 64, 24, 24
nn.ReLU: 1, 64, 24, 24
nn.SpatialConvolutionMM: 1, 192, 24, 24
nn.SpatialBatchNormalization: 1, 192, 24, 24
nn.ReLU: 1, 192, 24, 24
nn.SpatialCrossMapLRN: 1, 192, 24, 24
nn.SpatialMaxPooling: 1, 192, 12, 12
nn.Inception: 1, 256, 12, 12
nn.Inception: 1, 320, 12, 12
nn.Inception: 1, 640, 6, 6
nn.Inception: 1, 640, 6, 6
nn.Inception: 1, 1024, 3, 3
nn.Inception: 1, 736, 3, 3
nn.Inception: 1, 736, 3, 3
nn.SpatialAveragePooling: 1, 736, 1, 1
nn.View: 736
nn.Linear: 128
nn.Normalize: 128

```

Процес тренування цієї мережі займає приблизно 24 години на відеокарті NVidia Titan X. Отже, що ми отримуємо в результаті? Як візуалізувати

ці зв'язки на обличчі? Не зрозуміло, графічно зобразити ці метрики є складною задачею.

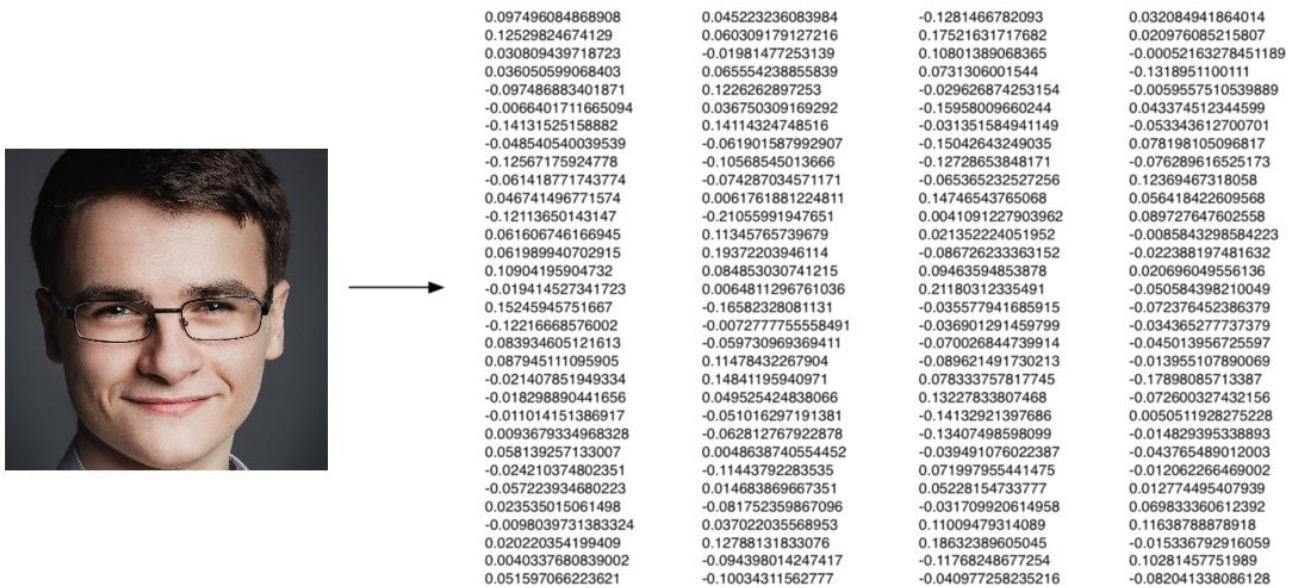


Рисунок 4.8 – Результат роботи мережі виділення особливостей

Насправді нам абсолютно не має різниці що саме генерує мережа. Важливо, щоб результати для однакових людей були приблизно однакові.

4.4 Пошук людини по базі даних

Останній крок є найлегшим в усьому процесі. Усе, що нам необхідно – це знайти людину серед відомих нам. Для цього можна використовувати будь-який алгоритм. Наприклад, для цієї ролі гарно підійде SVM класифікатор.

Зберемо дані, на яких будемо навчати класифікатор.



Рисунок 4.9 – фото Джеремі Кларксона

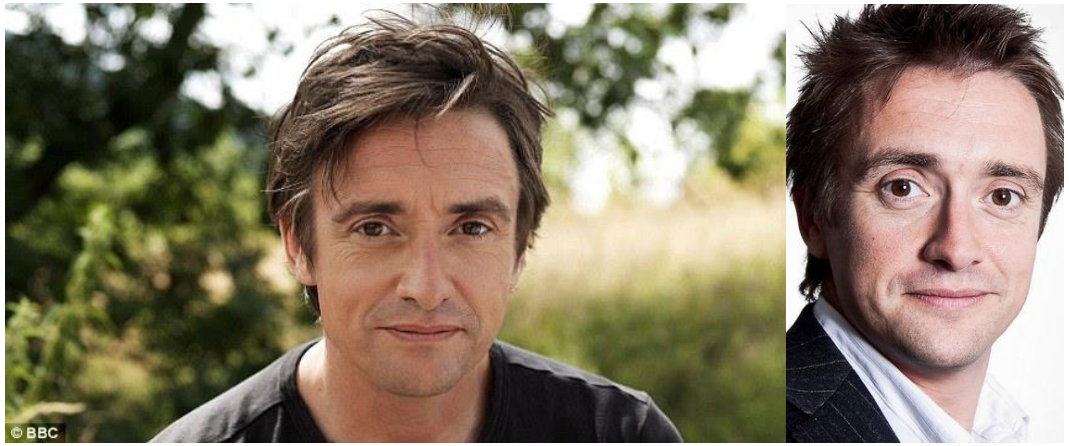


Рисунок 4.10 – Навчальна вибірка для Річарда Хаммонда

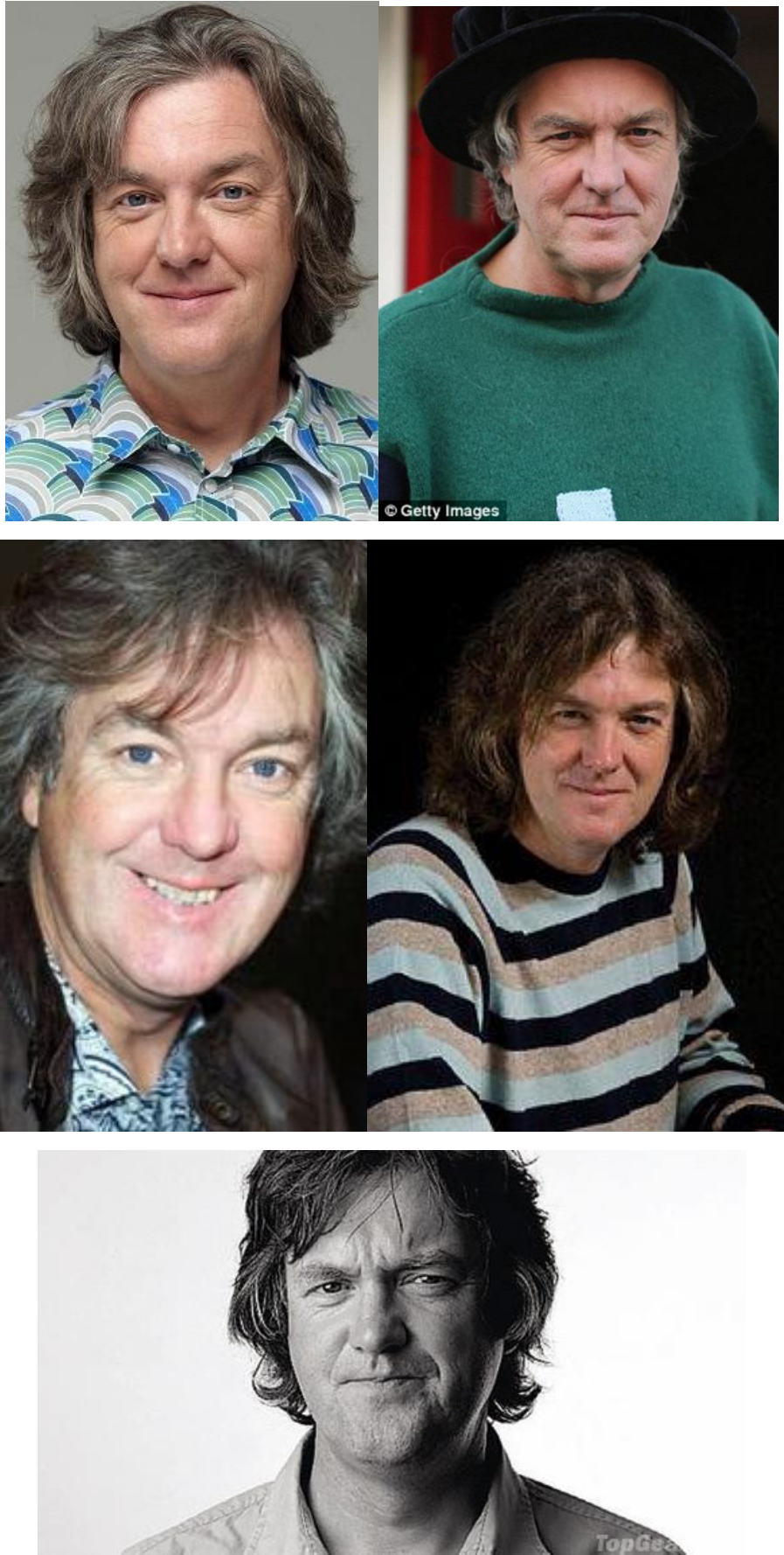


Рисунок 4.10 – Вибірка для навчання Джеймса Мея

Першим чином ми на даних фото запускаємо детектор облич. Потім проводимо нормування цих облич за допомогою створеної структурної моделі перед поданням на мережу, що буде виявляти ознаки кожного обличчя.



Рисунок 4.11 – приклади «нормованих» облич

```
0.014849368482828,0.10054879635572,0.10990730673075,-0.13388541340828,-0.017203044146299,-0.060
0.088297910988331,-0.004823740106076,0.099746376276016,-0.14157070219517,-0.060170289129019,0.0
0.09831478446722,0.070236377418041,0.068821899592876,-0.06771082431078,0.014927946962416,0.0332
0.094958402216434,0.045980162918568,0.081006750464439,-0.093230009078979,0.016965072602034,-0.0
0.030756305903196,0.12479083240032,0.1537029594183,-0.12454158067703,0.018650038167834,-0.07123
0.066001996397972,0.11692174524069,0.0035676613915712,-0.11703211069107,-0.073685929179192,-0.0
-0.010681649670005,0.015157928690314,-0.018830694258213,-0.016133764758706,-0.075115852057934,-
0.03742715716362,-0.034199804067612,-0.0365827716887,0.015391842462122,-0.089156538248062,0.026
0.014625946991146,-0.033607393503189,0.055713146924973,0.041101723909378,-0.049529023468494,0.1
0.052812650799751,-0.029624357819557,0.078864328563213,-0.018065735697746,-0.098214991390705,0.
0.020143695175648,-0.0006854044040665,-0.026975683867931,-0.0053625041618943,-0.062140122056007
0.011664037592709,-0.02269085496664,-0.019086403772235,0.012104541994631,-0.048881955444813,0.0
-0.047883406281471,0.19454275071621,0.15070277452469,-0.2516727745533,-0.014277837239206,-0.056
-0.1213983297348,0.11870682239532,0.065691709518433,-0.30478924512863,0.098245792090893,-0.0503
-0.050967533141375,-0.012118018232286,0.10637626051903,-0.21983744204044,0.10834065824747,0.006
-0.038127541542053,-0.010182742029428,0.096340551972389,-0.21043434739113,0.11051530390978,0.01
-0.0060945032164454,0.12262935936451,0.058394853025675,-0.24557285010815,0.0080501642078161,-0.
```

Рисунок 4.12 – Вихід нейронної мережі. Кожен рядок містить ознаки одного зображення

З отриманих ознак ми навчаємо нашу SVM. Тестуємо її роботу на зображенні, якого не було в навчальній вибірці.



Рисунок 4.13 – Тестове зображення

```

ml@gladosML:~/workspace/facefinder$ ./demos/classifier.py infer ./generated-embeddings/classifier.pkl 1.jpg
/home/ml/.local/lib/python2.7/site-packages/sklearn/lda.py:4: DeprecationWarning: lda.LDA has been moved to discrim
inant_analysis.LinearDiscriminantAnalysis in 0.17 and will be removed in 0.19
  "in 0.17 and will be removed in 0.19", DeprecationWarning)

=== 1.jpg ===
Predict hammond with 0.68 confidence.
ml@gladosML:~/workspace/facefinder$

```

Рисунок 4.14 – Робота на тестовому зображенні

Як бачимо, класифікатор відпрацював правильно. Обраних 17 фотографій для навчання вистачило для класифікації з впевненістю 68 відсотків.

Висновок

У практичній частині магістерської роботи було створено платформу для пошуку на ідентифікації людини. Ціллю платформи були продемонструвати, що дані про структурну модель обличчя можуть бути корисними для використання у практичних цілях.

У пункті розділі описується робота детектору облич. Він базується на провідній архітектурі Faster-R-CNN. Детектор демонструє відмінні результати на датасеті WIDER, а саме 95.2% на тестовій вибірці.

Далі на основі координат облич працює метод для генерації структурної моделі обличчя. В нашому даному випадку використано алгоритм FLE. Надалі при розвитку програми для цих цілей можна використати детектори на базі архітектур Stacked Hourglass Network та Convolutional Pose Machine. Дані про структурну модель обличчя в нашому випадку служать для нормалізації зображення. Після цього процесу використовується нейронна мережа, що навчена відрізняти одну особу від іншої за своїми ознаками, що ми не маємо змоги інтерпретувати осмислено.

Далі використовується SVM для ідентифікації людини, що відрізняє одну від іншої на основі ознак, отриманих від нейронної мережі.

Даний спосіб використання структурної моделі обличчя не є єдиним. Ці дані можуть стати основою для систем розпізнавання активності, стану, настрою і інших ознак.

5 РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ

5.1 Інформаційна карта проекту

1. Назва проекту	FaceFinder
2. Автори проекту	Савельєв Юрій
3. Коротка анотація	<p>Додаток, що буде аналізувати людей, їх середній вік, настрій, стан, активність. За наявності даних буде розпізнавати хто саме присутній на зображенні.</p> <p>В основі буде лежати алгоритм для визначення форми обличчя. Усі інші дані будуть базуватися саме на цих. Аналіз віку людей присутніх на зображенні допоможе розібратися в аудиторії певного місця, наприклад, магазину, конференції та інших масових зібрань. Аналіз настрою і стану допоможе також аналізувати що саме і як впливає на людей. Активність, тобто аналіз того, куди і на що вони сфокусовані допоможе у вирішенні задач безпеки.</p> <p>Безпосередньо ідентифікація потребуватиме спеціально зібраної бази даних, але це є основною і найбільш очевидною ціллю використання системи.</p>
4. Термін реалізації проекту	8 місяців
	Тривалість проекту (в місяцях)
5. Необхідні ресурси	Офіс, сервер, комп'ютери для кожного співробітника, гроші на зарплати.
	Перелік усіх необхідних ресурсів (фінансових, матеріальних, інтелектуальних тощо)

7. Головні цілі та завдання проекту	Розробити якісний та зручний продукт, який буде необхідний компаніям що потребують системи контролю людей.
-------------------------------------	--

8. Очікувані результати (Позитивні зміни, які відбудуться в результаті реалізації проекту після його завершення та в довгостроковій перспективі)
Універсальний сервіс, що матиме змогу слідкувати за обличчями, зберігати інформацію, що можна витягнути з форми людського обличчя.

5.2 Команда стартап-проекту

(Опис функціональних обов'язків та досвіду роботи головних виконавців проекту)	
Керівник проекту	Пошук інвесторів, керування компанією.
Технічний директор	Керування технічною стороною проекту, розробка архітектури, підбір необхідних технологій. Досвід у програмуванні та менеджменті.
Програміст (3)	Розробка додатку. Досвід у машинному навчанні. Досвід створення серверних додатків.
Веб-розробник (2)	Створення сайту-платформи для взаємодії з серверним додатком. Навички у дизайні та веб-розробці.

5.3 Маркетингова стратегія та маркетинговий план стартапу

5.3.1 Опис ідеї проекту (товару, послуги, технології)

В межах підпункту послідовно проаналізовані та подані у вигляді таблиць: зміст ідеї (що пропонується);

- можливі напрямки застосування;

- основні вигоди, що може отримати користувач товару (за кожним напрямком застосування);
- чим відрізняється від існуючих аналогів та замінників;

Перші три пункти подані у вигляді таблиці (табл. 4.1) і дають цілісне уявлення про зміст ідеї та можливі базові потенційні ринки, в межах яких потрібно шукати групи потенційних клієнтів.

Таблиця 4.1 - Опис ідеї стартап-проекту

Зміст ідеї	Напрямки застосування	Вигоди для користувача
Додаток, в основі якого йдуть алгоритми з машинного навчання, що допомагають знайти людей на зображенні та згенерувати структурну модель обличчя. На основі моделі користувач зможе отримати стан, аналіз емоцій та активності обличчя, настроїв, примірний вік людини.	1. Програма-ідентифікатор	Користувач, надавши базу даних у певній формі зможе шукати та ідентифікувати певних людей на зображеннях.
	2. Програма – аналізатор аудиторії	Програма зможе аналізувати стан людей, вік, стать, їх емоції та з цієї інформації генерувати звіт про аудиторію присутню на зображенні.

Аналіз потенційних техніко-економічних переваг ідеї (чим відрізняється від існуючих аналогів та замінників) порівняно із пропозиціями конкурентів передбачає:

- визначення переліку техніко-економічних властивостей та характеристик ідеї
- визначення попереднього кола конкурентів (проектів-конкурентів) або товарів-замінників чи товарів-аналогів, що вже існують на ринку, та

проводиться збір інформації щодо значень техніко-економічних показників для ідеї власного проекту та проектів-конкурентів відповідно до визначеного вище переліку;

- проводиться порівняльний аналіз показників: для власної ідеї визначаються показники, що мають а) гірші значення (W, слабкі); б) аналогічні (N, нейтральні) значення; в) кращі значення (S, сильні) (табл. 5.2).

Таблиця 5.2 - Визначення сильних, слабких та нейтральних характеристик ідеї проекту

№ п/п	Техніко- економічні характеристик и ідеї	(потенційні) товари/концепції конкурентів			W (слабка сторона)	N (нейтра- льна сторона)	S (сильна сторона)
		Мій проект	Конкурент 1	Конкурент 2			
1.	Ідентифікація	FaceFinder	Face++	Ivideon		+	
2.	Визначення віку	FaceFinder	Face++	Ivideon		+	
3.	Пошук людей на зображеннях низької якості	FaceFinder	Face++	Ivideon			+
4	Визначення емоцій людей	FaceFinder	Face++	Ivideon			+

Визначений перелік слабких, сильних та нейтральних характеристик та властивостей ідеї потенційного товару є підґрунтям для формування його конкурентоспроможності.

5.3.2 Технологічний аудит ідеї проекту

В межах даного підрозділу необхідно провести аудит технології, за допомогою якої можна реалізувати ідею проекту (технології створення товару). Визначення технологічної здійсненності ідеї проекту передбачає аналіз таких складових (табл. 5.3):

- за якою технологією буде виготовлено товар згідно ідеї проекту?
- чи існують такі технології, чи їх потрібно розробити/доробити?
- чи доступні такі технології авторам проекту?

Таблиця 5.3 - Технологічна здійсненність ідеї проекту

№ п/п	Ідея проекту	Технології її реалізації	Наявність технологій	Доступність технологій
		Технологія 1 (технологія виготовлення товару, надання послуги)	Чи вони наявні, або ж необхідно їх розробити/доробити?	Чи вони доступні авторам проекту?
	ML	Caffe	Є у наявності	Відкрите програмне забезпечення
	ML	Torch7	Є у наявності	Відкрите програмне забезпечення
	ML	Tensorflow	Є у наявності	Відкрите програмне забезпечення
	Web-server	Django	Є у наявності	Відкрите програмне забезпечення
	Web-server	PHP	Є у наявності	Відкрите програмне забезпечення
Обрана технологія реалізації ідеї проекту: TensorFlow + Django				

За результатами аналізу таблиці робиться висновок щодо можливості технологічної реалізації проекту: так чи ні, а також технологічного шляху, яким це доцільно зробити (з поміж названих технологій обираються такі, що доступні авторам проекту та є наявними на ринку).

5.3.3 Аналіз ринкових можливостей запуску стартап-проекту

Визначення ринкових можливостей, які можна використати під час ринкового впровадження проекту, та ринкових загроз, які можуть перешкодити реалізації проекту, дозволяє спланувати напрями розвитку проекту із урахуванням стану ринкового середовища, потреб потенційних клієнтів та пропозицій проектів-конкурентів.

Спочатку проводиться аналіз попиту: наявність попиту, обсяг, динаміка розвитку ринку (табл. 5.4).

Таблиця 5.4 - Попередня характеристика потенційного ринку стартап-проекту

№ п/п	Показники стану ринку (найменування)	Характеристика
1	Кількість головних гравців, од	7
2	Загальний обсяг продаж, грн./ум.од	1200 грн./ум.од
3	Динаміка ринку (якісна оцінка)	Зростає
4	Наявність обмежень для входу (вказати характер обмежень)	Немає
5	Специфічні вимоги до стандартизації та сертифікації	Немає
6	Середня норма рентабельності в галузі (або по ринку), %	34%

Рентабельність — поняття, що характеризує економічну ефективність виробництва, за якої за рахунок грошової виручки від реалізації продукції (робіт, послуг) повністю відшкодовує витрати на її виробництво й одержується прибуток як головне джерело розширеного відтворення [68].

Суть одного із найважливіших методів оцінки економічної ефективності інвестицій полягає у розрахунку їх середньої рентабельності за формулою [69]

$$R = \frac{P}{1*n} * 100,$$

де Р - прибуток за час експлуатації проекту; / - повна сума інвестиційних витрат; n - час експлуатації проекту.

Інвестувати грошові засоби доцільно тоді, коли від цього можна отримати більший прибуток, ніж від їх зберігання у банку. Порівнюючи середньорічну рентабельність інвестицій зі ставкою банківського відсотка, можна дійти висновку, що вигідніше [67].

Середня норма рентабельності в галузі (або по ринку) порівнюється із банківським відсотком на вкладення. За умови, що останній є вищим, можливо, має сенс вкласти кошти в інший проект.

За результатами аналізу таблиці робиться висновок щодо того, чи є ринок привабливим для входження за попереднім оцінюванням.

Надалі визначаються потенційні групи клієнтів, їх характеристики, та формується орієнтовний перелік вимог до товару для кожної групи (табл. 5.5).

Таблиця 5.5 - Характеристика потенційних клієнтів стартап-проекту

№ п/п	Потреба, що формує ринок	Цільова аудиторія (цільові сегменти ринку)	Відмінності у поведінці різних потенційних цільових груп клієнтів	Вимоги споживачів до товару
	Потреба у аналіз людей присутніх на системах фіксації зображень	Компанії, які потребують ідентифікації людей, аналізу аудиторій у певних місцях.	Різні потреби у аналізі. Наявність або відсутніх необхідних даних для ідентифікації	Кросс-платформенне рішення, доступне для усіх, для початку – веб-сервіс. Згодом можлива розробка веб-додатків.

Після визначення потенційних груп клієнтів проводиться аналіз ринкового середовища: складаються таблиці факторів, що сприяють ринковому впровадженню проекту, та факторів, що йому перешкоджають (табл. №№ 5.6-5.7). Фактори в таблиці подавати в порядку зменшення значущості.

Таблиця 5.6 - Фактори загроз

№ п/п	Фактор	Зміст загрози	Можлива реакція компанії
	Конкуренція	Наявність двох конкурентів та можливість появи нових.	Розробка унікального продукту, який буде більш якісним та мати цікаві нововведення.
	Зміна потреб користувачів	Користувачам необхідне програмне забезпечення з іншим функціоналом	Розширення функціональних можливостей додатку.

Таблиця 5.7 - Фактори можливостей

№ п/п	Фактор	Зміст можливості	Можлива реакція компанії
	Конкуренція	Відсутність аналогічного продукту для вітчизняного користувача.	Локалізація та адаптація сервісу для локальних груп. Адаптація до вітчизняних особливостей.
	Попит	Розвиток трендів правильного харчування	Постійна підтримка продукту

Надалі проводиться аналіз пропозиції: визначаються загальні риси конкуренції на ринку (табл. 5.8).

Таблиця 5.8 - Ступеневий аналіз конкуренції на ринку

Особливості конкурентного середовища	В чому проявляється дана характеристика	Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)
1. Вказати тип конкуренції - монополія/олігополія/ монополістична/чиста	Монополістична конкуренція	Підтримка якості продукту та постійні нововведення.
2. За рівнем конкурентної боротьби - локальний/національний/...	Міжнародний	Адаптація продукту як для вітчизняних так і для зарубіжних клієнтів.
3. За галузевою ознакою - міжгалузева/ внутрішньогалузева	Внутрішньогалузева	
4. Конкуренція за видами товарів: - товарно-родова - товарно-видова - між бажаннями	Товарно-видова	Підтримка якості продукту та постійні нововведення.
5. За характером конкурентних переваг - цінова / нецінова	Нецінова	Підтримка якості продукту та постійні нововведення.
6. За інтенсивністю - марочна/не марочна	Марочна	PR, реклама, просування бренду.

Таблиця 5.9 - Аналіз конкуренції в галузі за М. Портером

	Прямі конкуренти в галузі	Потенційні конкуренти	Постачальники	Клієнти	Товари-замінники
Складові аналізу	Ivideon , Face++	Наявність вже існуючих рішень	-	Контроль якості продукту	Наявність більш розповсюдженого рішення – традиційних журналів
Висновки :	Визначити інтенсивність конкурентної боротьби з боку прямих конкурентів	- чи є можливості входу в ринок? - чи є потенційні конкуренти? Строки виходу їх на ринок?	Чи постачальники диктують умови роботи на ринку? Які?	Чи клієнти диктують умови роботи на ринку? Які?	Обмеження для роботи на ринку через товари-замінники

За результатами аналізу таблиці робиться висновок щодо принципової можливості роботи на ринку з огляду на конкурентну ситуацію. Також робиться висновок щодо характеристик (сильних сторін), які повинен мати проект, щоб бути конкурентоспроможним на ринку. Другий висновок враховується при формулюванні переліку факторів конкурентоспроможності у п. 3.6.

На основі аналізу конкуренції, проведеного в п. 3.5 (табл. 5.9), а також із урахуванням характеристик ідеї проекту (табл. 5.2), вимог споживачів до товару (табл. 5.5) та факторів маркетингового середовища (табл. №№ 5.6-5.7) визначається та обґрунтовується перелік факторів конкурентоспроможності. Аналіз оформлюється за табл. 5.10

Таблиця 5.10 - Обґрунтування факторів конкурентоспроможності

№ п/п	Фактор конкурентоспроможності	Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)
	Якість продукту	Один із факторів для вибору продукту клієнтом.
	Кількість видів соціальної взаємодії	Більша зацікавленість клієнта продуктом.
	Ціна	Один із факторів для вибору продукту клієнтом.

За визначеними факторами конкурентоспроможності (табл. 5.10) проводиться аналіз сильних та слабких сторін стартап-проекту (табл. 5.11).

Таблиця 5.11 - Порівняльний аналіз сильних та слабких сторін «назва проекту»

№ п/ п	Фактор конкурентоспроможності	Бали 1-20	Рейтинг товарів-конкурентів у порівнянні						
			-3	-2	-1	0	+1	+2	+3
1	Якість продукту	20			+				
2	Різноманіття функціоналу	20		+					
3	Ціна	15					+		

Фінальним етапом ринкового аналізу можливостей впровадження проекту є складання SWOT-аналізу (матриці аналізу сильних (Strength) та слабких (Weak) сторін, загроз (Troubles) та можливостей (Opportunities) (табл.5. 12) на основі виділених ринкових загроз та можливостей, та сильних і слабких сторін (табл. 5.11).

Перелік ринкових загроз та ринкових можливостей складається на основі аналізу факторів загроз та факторів можливостей маркетингового середовища. Ринкові загрози та ринкові можливості є наслідками (прогнозованими

результатами) впливу факторів, і, на відміну від них, ще не є реалізованими на ринку та мають певну ймовірність здійснення. Наприклад: зниження доходів потенційних споживачів – фактор загрози, на основі якого можна зробити прогноз щодо посилення значущості цінового фактору при виборі товару та відповідно, – цінової конкуренції (а це вже – ринкова загроза).

Таблиця 5.12 - SWOT- аналіз стартап-проекту

Сильні сторони: Різноманіття функціоналу	Слабкі сторони: Ціна
Можливості: Перехід на розповсюдження по підписці, або наявність безкоштовного функціоналу	Загрози: видавлення з ринку конкурентами, не розуміння користувачем переваг

На основі SWOT-аналізу розробляються альтернативи ринкової поведінки (перелік заходів) для виведення стартап-проекту на ринок та орієнтовний оптимальний час їх ринкової реалізації з огляду на потенційні проекти конкурентів, що можуть бути виведені на ринок (див. табл. 9, аналіз потенційних конкурентів).

Визначені альтернативи аналізуються з точки зору строків та ймовірності отримання ресурсів (табл. 5.13).

Після аналізу зазначити обрану альтернативу.

З означених альтернатив обирається та, для якої: а) отримання ресурсів є більш простим та ймовірним; б) строки реалізації – більш стислими.

Таблиця 5.13 - Альтернативи ринкового впровадження стартап-проекту

№ п/п	Альтернатива (орієнтовний комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
	PR, просування бренду	70%	2
	Перехід на безкоштовне розповсюдження	60%	4

5.4 Розроблення ринкової стратегії проекту

Розроблення ринкової стратегії першим кроком передбачає визначення стратегії охоплення ринку: опис цільових груп потенційних споживачів (табл. 5.14).

Таблиця 5.14 - Вибір цільових груп потенційних споживачів

№ п/п	Опис профілю цільової групи потенційних клієнтів	Готовність споживачів сприйняти продукт	Орієнтовний попит в межах цільової групи (сегменту)	Інтенсивність конкуренції в сегменті	Простота входу у сегмент
1	Малі компанії, що займаються системами безпеки та аналізу аудиторії людей.	Висока	Високий	Висока	Середня
2	Крупні компанії, що мають системи безпеки	Невисока	Невисокий	Невисока	Середня
Які цільові групи обрано: 1					

За результатами аналізу потенційних груп споживачів (сегментів) автори ідеї обирають цільові групи, для яких вони пропонуватимуть свій товар, та визначають стратегію охоплення ринку:

- якщо компанія зосереджується на одному сегменті – вона обирає стратегію концентрованого маркетингу;
- якщо працює із кількома сегментами, розробляючи для них окремо програми ринкового впливу – вона використовує стратегію диференційованого маркетингу;
- якщо компанія працює із всім ринком, пропонуючи стандартизовану програму (включно із характеристиками товару/послуги) – вона використовує масовий маркетинг.

Для роботи в обраних сегментах ринку необхідно сформувати базову стратегію розвитку (табл. 5.15).

Стратегія диференціації передбачає надання товару важливих з точки зору споживача відмінних властивостей, які роблять товар відмінним від товарів конкурентів. Така відмінність може базуватися на об'єктивних або суб'єктивних, відчутних і невідчутних властивостях товару(у ширшому розумінні – комплексі маркетингу), бути реальною або уявною. Інструментом реалізації стратегії диференціації є ринкове позиціонування.

Переваги стратегії за Ж.-Ж. Ламбенем:

- по відношенню до прямих конкурентів диференціація знижує ступінь заміності товару, посилює прихильність марці, зменшує чутливість до ціни і тим самим підвищує рентабельність;
- прихильність клієнтів послабляє їх тиск на фірму і перешкоджає приходу на ринок нових конкурентів;
- підвищена рентабельність збільшує стійкість до можливого зростання витрат в результаті дій сильного постачальника;
- відмінні властивості товару і завойована прихильність клієнтів захищають фірму і від товарів-замінників.

Реалізація цієї стратегії вимагає, як правило, більш високих витрат. Проте успішна диференціація дозволяє компанії домогтись більшої рентабельності за рахунок того, що ринок готовий прийняти більш високу ціну (цінову премію бренду).

При веденні конкурентної боротьби з використанням цієї стратегії на ринку в першу чергу терплять фіаско фірми, що не здатні визначати потреби цільових ринків, оперативно реагувати на зміни в ринковому попиті, проводити ефективну політику маркетингових комунікацій, не мають необхідних навичок в області брендингу. Найважливішими здібностями, які повинна мати компанія, що приймає цю стратегію, є з генерування маркетингових ноу-хау, здійснення продуктових новацій.

Таблиця 5.15 - Визначення базової стратегії розвитку

№ п/п	Обрана альтернатива розвитку проекту	Стратегія охоплення ринку	Ключові конкурентоспроможні позиції відповідно до обраної альтернативи	Базова стратегія розвитку*
	Надання товару важливих з точки зору споживача відмітних властивостей	Стратегія диференціації	Більша кількість способів соціальної взаємодії, можливість переходу до безкоштовного типу розповсюдження	Стратегія диференціації

Наступним кроком є вибір стратегії конкурентної поведінки.

Стратегія наслідування лідеру

Компанії, що приймають слідування за лідером – це підприємства з невеликою часткою ринку, які вибирають адаптивну лінію поведінки на ринку, усвідомлюють своє місце на ній і йдуть у фарватері фірм-лідерів. Головна перевага такої стратегії – економія фінансових ресурсів, пов'язаних з необхідністю розширення товарного(галузевого) ринку, постійними інноваціями, витратами на утримання домінуючого положення.

Стратегія наслідування лідеру найчастіше має місце у випадку олігополії, коли кожен конкурент прагне уникнути боротьби, особливо цінової, а також у випадку, коли слабо виражений ефект масштабу, що не дозволяє отримати переваги від об'ємів продажів або ж він не грає істотної ролі. Стратегію наслідування лідеру приймають також фірми, які не змогли реалізувати стратегію виклику лідерові.

Компанії, що приймають таку стратегію, зазвичай випускають товари-імітатори, займаючи ринкову частку, яку з різних причин не можуть охопити фірми лідери. Вибір такої стратегії може також бути обумовлений також перевагою локалізації (краще знання ринку, налагоджені зв'язки з клієнтами тощо).

Для ефективною реалізації цієї стратегії компанії повинні задовольняти наступним основним умовам:

- систематичний аналіз сегментації ринку з метою виділення нових ринкових сегментів або таких, що незадовільно обслуговуються;
- ефективне використання НДДКР з метою вдосконалення технологічних процесів і незначних продуктових новацій;
- концентрація на прибутковості, а не на простому зростанні об'ємів продажів;
- постійний аналіз витрат на всіх стадіях виробництва і логістики;
- залишатися досить малим, щоб не бути досить цікавим для фірм-лідерів;
- сильний керівник, здатний не лише формулювати стратегію, але і тримати усю діяльність компанії під власним контролем.

Якщо врахувати, що лідерами ринку можуть бути лише декілька компаній, то ця стратегія є наймасовішою.

Таблиця 5.16 - Визначення базової стратегії конкурентної поведінки

№ п/п	Чи є проект «першопрохідцем» на ринку?	Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?	Чи буде компанія копіювати основні характеристики товару конкурента, і які?	Стратегія конкурентної поведінки*
	Ні	Забирати існуючих	Ні, він буде їх пов'язувати та розширювати, створюючи новий функціонал	Стратегія наслідування лідеру

На основі вимог споживачів з обраних сегментів до постачальника (стартап-компанії) та до продукту (див. табл. 5.5), а також в залежності від обраної базової стратегії розвитку (табл. 5.15) та стратегії конкурентної поведінки (табл. 16) розробляється стратегія позиціонування (табл. 5.17). що полягає у формуванні ринкової позиції (комплексу асоціацій), за яким споживачі мають ідентифікувати торгівельну марку/проект.

Таблиця 5.17 - Визначення стратегії позиціонування

№ п/п	Вимоги до товару цільової аудиторії	Базова стратегія розвитку	Ключові конкурентоспроможні позиції власного стартап-проекту	Вибір асоціацій, які мають сформувати комплексну позицію власного проекту (три ключових)
1	Якість	позиціонування за показниками якості	Тестування розробленого продукту та виправлення всіх багів	Стабільність роботи, якість роботи
2	Більша кількість функціональних можливостей	На основі специфічних відчутних характеристик	Розробка більшої кількості оригінальних можливостей	Велика можливостей

Результатом виконання підрозділу має стати узгоджена система рішень щодо ринкової поведінки стартап-компанії, яка визначатиме напрями роботи стартап-компанії на ринку.

5.5 Розроблення маркетингової програми стартап-проекту

Першим кроком є формування маркетингової концепції товару, який отримає споживач [70].

Таблиця 5.18 - Визначення ключових переваг концепції потенційного товару

№ п/п	Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)
	Велика кількість можливостей	Більша кількість можливостей	Більша кількість видів соціальної взаємодії
	Безкоштовне розповсюдження	Безкоштовність	Безкоштовне розповсюдження

Надалі розробляється трирівнева маркетингова модель товару: уточнюється ідея продукту та/або послуги, його фізичні складові, особливості процесу його надання (табл. 5.20).

Таблиця 5.19 - Техніко-економічні характеристики товару

№ п/п	Група показників	Склад показників
1	Економічні	Вартість обслуговування, експлуатації, утилізації, витратних матеріалів, ремонту, знижки
2	Призначення (технічні)	Показники, що визначають головний напрямок використання товару та можливу сферу його застосування: класифікаційні показники, складу і структури, технічної досконалості. Приклад: маса, розміри, кількість елементів, види та характеристики матеріалу (для меблів); густина, калорійність, відсоткове співвідношення цукру, жиру, солі (для харчових продуктів); ємність салону, кількість пасажирських місць, потужність двигуна (для автобуса); % вуглеводню (для нафти); структура плетіння, склад пряжі (для тканин) тощо)

№ п/п	Група показників	Склад показників
3	Надійності	<p>Характеризують здатність товару безвідмовно функціонувати: безвідмовність, довговічність, ремонтпридатність.</p> <p>Приклад: кількість разів використання (запуску в роботу) товару, строк безвідмовної праці, гарантійний термін.</p>
3	Технологічні	<p>Показники, що характеризують можливість оптимізації витрат матеріалів, праці, коштів, часу під час технологічної підготовки виробництва, виготовлення та використання товару.</p> <p>Приклад: трудомісткість виготовлення та технологічна собівартість товару.</p>
4	Ергономічні	<p>Показники ступеню адаптованості технічних та конструктивних рішень виробу до біологічних властивостей людини та середовища використання товару: гігієнічні, антропометричні, фізіологічні та психологічні.</p> <p>Приклад: рівень освітлення, температури, вологості, токсичності, шуму, вібрацій, достатність робочого простору, раціональність розміщення, зручність спостереження за сигнальними елементами, відповідність зросту людини, рівень статичного напруження м'язів робочого органу людини, відповідність виробу можливостям сприйняття інформації користувачем, зручність користування під час виконання основних та допоміжних операцій, зручність управління, простота набуття навичок.</p>
5	Органолептичні	<p>Показники, що визначають властивості товару, які людина може визначити за допомогою своїх органів чуття.</p> <p>Приклад: смак, присмак, запах, забарвлення, каламутність.</p>

№ п/п	Група показників	Склад показників
6	Естетичні	Показники, що оцінюють зовнішній вигляд товару. Приклад: інформаційна виразність, раціональність форми, цілісність композиції, досконалість виробничого виконання, стабільність товарного вигляду, відповідність стилю, відповідність моді.
7	Транспортабельності	Показники, що визначають пристосованість продукції до транспортування, підготовчих, початкових і кінцевих операцій перевезення. Приклад: Середню трудомісткість підготовки одиниці продукції до перевезень (із навантаженнями та закріпленнями включно), середню вартість пакування в тару для перевезення, середню тривалість завантаження/розвантаження партії товару з одиниці рухомого складу.
8	Екологічності	Показники, що характеризують рівень негативного впливу на довкілля. Приклад: вміст шкідливих домішок у викидах, ймовірність викидів шкідливих домішок під час транспортування, збереження, експлуатації.
9	Безпеки	Показники безпечності та нешкідливості споживання товару. Приклад: можливість безпечної праці протягом визначеного часу, час спрацювання захисних пристроїв, електрична міцність високовольтних мереж, наявність блокуючих пристроїв, ременів безпеки, ізоляції, аварійної сигналізації.

Таблиця 5.20 - Опис трьох рівнів моделі товару

Рівні товару	Сутність та складові		
I. Товар за задумом	Потреба у відслідковуванні людини, безпека, аналіз		
II. Товар у реальному виконанні	Властивості/характеристики	М/Нм	Вр/Тх /Тл/Е/Ор
	1. Якість		
	2. Кількість видів соціальної взаємодії		
	3. Ціна		
	Якість: тестування на предмет багів		
Пакування нема			
Марка: YS Studios + FaceFinder			
III. Товар із підкріпленням	До продажу: Безкоштовне розповсюдження		
	Після продажу: Постійне розвивання якості взаємодії, інтеграція з іншими системами.		
За рахунок чого потенційний товар буде захищено від копіювання: патент			

Після формування маркетингової моделі товару слід особливо відмітити – чим саме проект буде захищено від копіювання. Захист може бути організовано за рахунок захисту ідеї товару (захист інтелектуальної власності), або ноу-хау, чи комплексне поєднання властивостей і характеристик, закладене на другому та третьому рівнях товару.

Наступним кроком є визначення цінових меж, якими необхідно керуватись при встановленні ціни на потенційний товар (остаточне визначення ціни відбувається під час фінансово-економічного аналізу проекту), яке передбачає аналіз ціни на товари-аналоги або товари субститути, а також аналіз рівня доходів цільової групи споживачів (табл. 5.21). Аналіз проводиться експертним методом.

Таблиця 19 - Визначення меж встановлення ціни

№ п/п	Рівень цін на товари-замінники	Рівень цін на товари-аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
	0	12-15	3000	0-12

Наступним кроком є визначення оптимальної системи збуту, в межах якого приймається рішення (табл.5.22):

- проводити збут власними силами або залучати сторонніх посередників (власна або залучена система збуту);
- вибір та обґрунтування оптимальної глибини каналу збуту;
- вибір та обґрунтування виду посередників.

Таблиця 5.22 - Формування системи збуту

№ п/п	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
	Одна одиниця на особу	Роздрібна торгівля	Однорівневий	Платформи онлайн-дистрибуції

Функції каналу розподілу:

- Встановлення контактів із споживачами і підтримання їх;
- Формування попиту і стимулювання збуту;
- Організація руху товару;
- Розробка і реалізація програм з підтримки лояльності клієнтів;
- Дослідницька робота зі збору маркетингової інформації;
- Доробка товару, виходячи з потреб конкретного покупця;
- Фінансування (у тому числі лізинг, відстрочка платежів, платежі по частинах);
- Прийняття на себе ризику. Завдання капала розподілу:
- Знизити кількість необхідних контактів для збуту продукції;

- Зменшити витрати в ході збуту продукції;
- Максимізувати відчувається споживачем корисність від товару.

Канали розподілу характеризуються такими поняттями:

Рівень каналу дистрибуції визначається кількістю типів посередників, що беруть участь в розподілі товару від виробника до споживача.

Глибина каналу дистрибуції - кількість посередників, які безпосередньо беруть участь в розподілі товару від продавця до покупця.

Ширина каналу дистрибуції - кількість посередників на певному рівні. Дана характеристика показує ступінь охоплення ринку посередниками, що пропонують товар фірми [68].

Останньою складовою маркетингової програми є розроблення концепції маркетингових комунікацій, що спирається на попередньо обрану основу для позиціонування, визначену специфіку поведінки клієнтів (табл. 5.23).

Визначення ідеї та теми рекламного звернення зумовлює всі наступні етапи планування рекламної діяльності. Тому цей етап є основним для професіоналів — творчих працівників рекламних агенцій чи рекламних підрозділів підприємств. Один із класиків рекламного бізнесу Д. Огілві радить ретельно вивчити те, що необхідно рекламувати: «Отримавши завдання на рекламування автомобіля «Роллс Ройс», я витратив три тижні на вивчення цієї моделі. І тоді народилася ідея, яка втілилась у рекламному зверненні: на швидкості 60 миль за годину найголосніший звук, який можна почути в машині, — цокання годинника» [71].

Таблиця 5.23 - Концепція маркетингових комунікацій

№ п/п	Специфіка поведінки цільових клієнтів	Канали комунікацій, якими користуються цільові клієнти	Ключові позиції, обрані для позиціонування	Завдання рекламного повідомлення	Концепція рекламного звернення
	Клієнти обиратимуть якісний товар з функціями, які будуть їм подобатися	Соціальні мережі, електронна пошта, мобільні телефони	Якість, велика кількість видів аналізу, широкі можливості інтеграції	Показати якість, можливості товару і універсальність	Показ можливостей, які надає система, та переваг її використання.

Висновки

В даному розділі було проведено аналіз програмного продукту у якості стартап проекту. Можна зазначити що у проекті є можливість комерціалізації, адже ринок технологій розпізнання осіб динамічно розвивається, створюються нові сервіси які, в свою чергу, стимулюють попит на різноманітні системи для розпізнання осіб.

На локальному ринку наявна незначна конкуренція, тому вихід на нього не буде важким. Проект є доволі конкурентноспроможним завдяки, в першу чергу, своїй невисокій вартості, проте для реалізації проекту необхідно буде залучати певні інвестиції для розширення штату робітників, оренди серверів, оренди офісу, купівлі потужного обладнання, яке необхідне для коректної роботи сервісу .

Для впровадження ринкової реалізації проекту слід обрати альтернативу, яка передбачає розробку програмного продукту, а потім якісну рекламу та PR, сконцентровану навколо позитивних характеристиках даного програмного продукту, таких як низька ціна, кросплатформеність тощо.

ВИСНОВКИ

Задача автоматичного розпізнання облич на сьогоднішній день є досить актуальною, як через велику кількість наукових досліджень даної задачі так і через великий потенціал використання даної технології у комерційних проектах.

Задачу слід розглядати як комплексну проблему, наприклад розділюючи детекцію та ідентифікацію. Рішення в одній підзадачі матиме вплив на вирішення іншої, але тим не менш їх можна вирішувати окремо.

Процес вирішення задачі ідентифікації людини полягає у послідовній відповіді на наступні питання:

- Чи присутнє обличчя на зображенні?
- Де знаходиться обличчя?
- Кому належить це обличчя?

Нейронні мережі зарекомендували себе як потужний механізм для класифікації зображень, але протягом останнього часу вони активно використовуються і в інших областях комп'ютерного зору. Їм належать перші місця в змаганнях алгоритмів розпізнавання, таких як ImageNet, а якість розпізнавання близька, а іноді і перевищує людську.

Окрім того, конвуляційні мережі відносно легко вирішують проблеми високоточного розпізнавання, що визивають проблеми у людей, наприклад, визначення породи собаки чи марки автомобіля, та інші задачі, що потребують виділення вузькоспецифічних ознак.

Серед недоліків можна виділити проблеми при обробці елементів малого розміру, і відсутність можливості справлятися зі спотвореннями, такими як розмиваючий фільтр або шум. Також недоліком є проблема вибору набору даних, який задовольнить проблему, що вирішується та буде досить якісним і великим для навчання нейронної мережі.

Отже, дослідження використання нейронних мереж у цій проблемі є логічним рішенням.

Був проведений аналіз різних методів у розпізнаванні облич. У огляді методи на нейронних мережах суттєво краще демонструють себе в тестах і мають кращу точність, але тим не менш, важливо відмітити, що одні з найкращих методів є модифікаціями класичних підходів з використанням нейронних мереж.

Незважаючи на гарні результати на тестових даних, деякі алгоритми мають властивість погано демонструвати себе на інших наборах, або в реальних умовах. Тобто проблема універсальної системи все ще є відкритою.

У практичній частині магістерської роботи було створено платформу для пошуку на ідентифікації людини. Ціллю платформи були продемонструвати, що дані про структурну модель обличчя можуть бути корисними для використання у практичних цілях.

У пункті розділі описується робота детектору облич. Він базується на провідній архітектурі Faster-R-CNN. Детектор демонструє відмінні результати на датасеті WIDER, а саме 95.2% на тестовій вибірці.

Далі на основі координат облич працює метод для генерації структурної моделі обличчя. В нашому даному випадку використано алгоритм FLE. Надалі при розвитку програми для цих цілей можна використати детектори на базі архітектур Stacked Hourglass Network та Convolutional Pose Machine. Дані про структурну модель обличчя в нашому випадку служать для нормалізації зображення. Після цього процесу використовується нейронна мережа, що навчена відрізняти одну особу від іншої за своїми ознаками, що ми не маємо змоги інтерпретувати осмислено.

Далі використовується SVM для ідентифікації людини, що відрізняє одну від іншої на основі ознак, отриманих від нейронної мережі.

Даний спосіб використання структурної моделі обличчя не є єдиним. Ці дані можуть стати основою для систем розпізнавання активності, стану, настрою і інших ознак.

Було запропоновано та розроблено план використання програмного продукту у якості стартап продукту.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks //Advances in neural information processing systems. – 2012. – С. 1097-1105.
2. P. A. Viola and M. J. Jones. Rapid object detection using a boosted cascade of simple features. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2001.
3. Dalal N., Triggs B. Histograms of oriented gradients for human detection //Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. – IEEE, 2005. – Т. 1. – С. 886-893.
4. Olszewska J. Automated Face Recognition: Challenges and Solutions//Intech – Open science Open minds , [Електронний ресурс]. – Режим доступу: <https://cdn.intechopen.com/pdfs-wm/52911.pdf> - /. – Дата доступу : 31.05.2017
5. Sirovich L., Kirby M. Low-dimensional procedure for the characterization of human. Journal of the Optical Society of America// Optics, Image Science and Vision. 1987. 4(3):519–524.
6. Zhao W., Chellappa R., Rosenfeld A., Phillips P.J. Face recognition: A literature survey. ACM Computing Surveys. 2003. 35(4):399–458.
7. Iosifidis, A., Gabbouj, M. Scaling-up class-specific kernel discriminant analysis for large-scale face verification. IEEE Transactions on Information Forensics and Security. 2016.11(11):2453–2465
8. Almudhahka N., Nixon M., Hare J. Human face identification via comparative soft biometrics. In: Proceedings of the IEEE International Conference on Identity, Security and Behavior Analysis (ISBA).Sendai, JP, 29 Feb–02 Mar 2016. pp. 1–6.
9. Berg T.I., Berg A.C., Edwards J., Forsyth D.A. Who's in the Picture?. In: Proceedings of the Neural Information Processing Systems Conference (NIPS). Springer. 2004. pp. 137–144.

10. Torres L. Is there any hope for face recognition? In: Proceedings of the IEEE International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS). 2004.
11. Nash S., Rhodes M., Olszewska J. I. iFR: Interactively-pose-corrected face recognition. In: Proceedings of the INSTICC International Conference on Bio-Inspired Systems and Signal Processing (BIOSIGNALS). 2016. pp. 106–112.
12. Shreve M., Godavarthy S., Goldgof D., Sarkar, S. Macro- and micro-expression spotting in long videos using spatio-temporal strain. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition Workshops(AFGR). 2011. pp. 51–56.
13. Liu L., Xiong C., Zhang H., Niu Z., Wang M., Yan S. Deep aging face verification with large gaps. *IEEE Transactions on Multimedia*. 2016. 18 (1):64–75.
14. Julian P., Dehais C., Lauze F., Charvillat V., Bartoli A., Choukroum A. Automatic hair detection in the wild. In: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR). 2010. pp. 4617–4620.
15. Wood R., Olszewska J. I. Lighting-variable AdaBoost based-on system for robust face detection. In: Proceedings of the INSTICC International Conference on Bio-Inspired Systems and Signal (BIOSIGNALS). 2012. pp. 494–497.
16. Shan S., Gao W., Cao B., Zhao D. Illumination normalization for robust face recognition against varying lighting conditions. In: Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition (AFGR). 2003. pp. 157–164.
17. ISO/IEC 19794-5:2011. The Face Image Format Standards. In: Information Technology Biometric data interchange formats: Part 5: Face image data. 2nd Ed. 2011.

18. Prince S.J.D., Elder J., Hou Y., Sizinstev M., Olevskiy E. Towards face recognition at a distance. In: Proceedings of the IET Conference on Crime and Security. 2006. pp. 570–575.
19. Mudunuri S.P., Biswas S. Low resolution face recognition across variations in pose and illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2016. 38(5): 1034–1040.
20. Zou W.W.W., Yuen P.C. Very low resolution face recognition problem. *IEEE Transaction on Image Processing*. 2012. 21(1):327–340.
21. Huang H., He H. Super resolution method for face recognition using non-linear mappings on coherent features. *IEEE Transaction on Neural Networks*. 2011. 22(1):121–130.
22. Li H., Lam K.M. Guided iterative back-projection scheme for single image super-resolution. In: Proceedings of the IEEE Global High Tech Congress on Electronics (GHTCE). 2013. pp. 175–180.
23. Senior A.W, Pankanti S. Privacy protection and face recognition. In: Li S.Z., Jain A.K., editors. *Handbook of Face Recognition*. 2nd ed. Springer. 2011. pp. II.5–II.21.
24. Gross R. Face databases. In: Li S.Z., Jain A.K., editors. *Handbook of Face Recognition*. Springer Verlag. 2005. pp. 301–327.
25. Phillips P.J., Moon H., Rizvi S.A., Rauss P.J. The FERET evaluation methodology for face recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000. 22(10):1090–1104.
26. Jesorsky O., Kirchberg K., Frischholz R. Face detection using the Hausdorff distance. In Bigun J., Smeraldi F., editors. *Audio and Video based Person Authentication*. LNCS Springer. 2001. pp. 90–95.
27. Caltech 10000 Web Faces. Human Faces Collected from Google Image Search. CalTech University. Pasadena, California, USA [Internet]. 2005. [Электронный ресурс]. – Режим доступа: www.vision.caltech.edu/Image_Datasets/Caltech_10K_WebFaces. / . – Дата доступа : 21.05.2017.

28. Caltech 10000 Web Faces. Human Faces Collected from Google Image Search. CalTech University. Pasadena, California, USA [Электронный ресурс]. – Режим доступа: www.vision.caltech.edu/Image_Datasets/Caltech_10K_WebFaces. – Дата доступа: 02.05.2017
29. Phillips P.J., Flynn P.J., Scruggs T., Bowyer K.W., Chang J., Hoffman K., Marques J., Min, J., Worek W. Overview of the face recognition grand challenge. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2005. pp. I.947–I.954.
30. Yale Face Database. Yale University. Connecticut, USA [Электронный ресурс]. – Режим доступа: http://vision.ucsd.edu/yale_face_dataset_original/yalefaces.zip [Accessed: 2016-07-07] . – Дата доступа : 03.05.2017.
31. Matas J., Hamouz M., Jonsson K., Kittler J., Li Y., Kotroupolous C., Tefas A., Pitas I., Tan T., Yan H., Ameraldi F., Bigun J., Capdevielle N., Gerstner W., Ben-Yacoub S., Abd jaoued Y., Mayoraz E. Comparison of face verification results on the XM2VTS database. In: Proceedings of the IEEE International Conference on Pattern Recognition (ICPR). 2000. pp. 858–863.
32. Grgic M., Delac K., Grgic S., Klmpak B. SCface - Surveillance cameras face database. Multimedia Tools Applications. 2011. 51:863–879.
33. Fddb. A Benchmark for Face Detection in Unconstrained Settings Jain V., Learned - Miller E. Technical Report. UM-CS-2010-009. University of Massachusetts Amherst. USA [Электронный ресурс]. – Режим доступа: <http://vis-www.cs.umass.edu/fddb>. Дата доступа: 03.05.2017
34. Learned-Miller E., Huang G.B., Roy-Chowdhury A., Li H., Hua G. Labeled Faces in the Wild: A Survey. In: Kawulok M., Emre Celebi M., Smolka B., editors. Advances in Face Detection and Facial Image Analysis. Springer. 2016. pp. 189–248.
35. Li X., Pefister T., Huang X., Zhao G., Pietikainen M. A spontaneous micro-expression database: Inducement, collection and baseline. In: Proceedings

- of the IEEE International Conference on Automatic Face and Gesture Recognition (AFGR). 2013. pp. 1–6.
36. Panis G., Lanitis A., Tsapatsoulis N., Cootes T.F. Overview of research on facial ageing using the FG-NET ageing database. *IET Biometrics*. 2016. 5(2):37–46.
 37. Хуршудов, А.А. Разработка систем распознавания визуальных образов в потоке данных: дис.кан. тех. наук. 05.13.01: защищена 22.01.15: утв. 15.07.15 / Хуршудов Артем Александрович. – Краснодар, 2015. – 130 с.
 38. Hinton, G.E. Learning multiple layers of features from tiny images / A. Krizhevsky, G. Hinton // Computer Science Department, University of Toronto, Tech. Rep. — 2009. — no. 1.4. — Pp. 7–10.
 39. Deng, L. The MNIST database of handwritten digit images for machine learning research / L. Deng // *IEEE Signal Processing Magazine*. — 2012. — no. 29.6. — Pp. 141–142.
 40. Bengio, Y. Representation learning: A review and new perspectives / Y. Bengio, A. Courville, P. Vincent // *Pattern Analysis and Machine Intelligence*. — 2013. — no. 35(8). — Pp. 1798–1828.
 41. Fergus, R. A sparse object category model for efficient learning and exhaustive recognition / R. Fergus, P. Perona, A. Zisserman // *Computer Vision and Pattern Recognition*. — 2005. — no. 1. — Pp. 380–387.
 42. Murphy-Chutorian, E. Head pose estimation in computer vision: A survey. / E. Murphy Chutorian, M.M. Trivedi // *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31.4. — 2009. — Pp. 607–626.
 43. Hubel, D. H. Receptive fields and functional architecture of monkey striate cortex / D. H. Hubel, T. N. Wiesel // *The Journal of physiology*. — 1968. — no. 195(1). — Pp. 215–243.
 44. Kreutz-Delgado, K. Dictionary learning algorithms for sparse representation / K. Kreutz-Delgado // *Neural computation*. — 2003. — no. 15.2. — Pp. 349–396.

45. Fukushima, K. Neocognitron: A hierarchical neural network capable of visual pattern recognition / K. Fukushima // *Neural networks*. — 1988. — no. 2. — Pp. 119–130.
46. Serre, T. Object recognition with features inspired by visual cortex / T. Serre, L. Wolf, T. Poggio // *Computer Vision and Pattern Recognition*. — 2005. — no. 2. — Pp. 994–1000.
47. Krizhevsky, A. Imagenet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // *Advances in neural information processing systems*. — 2012. — Pp. 1097–1105.
48. Ciresan, D. Multi-column deep neural networks for image classification / D. Ciresan, U. Meier, J. Schmidhuber // *Computer Vision and Pattern Recognition (CVPR)*. — 2012. — Pp. 3642–3649.
49. Matsugu, M. Subject independent facial expression recognition with robust face detection using a convolutional neural network / M. Matsugu // *Neural Networks*. — 2003. — no. 16(5). — Pp. 555–559.
50. Yahia, S. Human detection based on integral Histograms of Oriented Gradients and SVM / S. Yahia, M. Atri, R. Tourki // *Communications, Computing and Control Applications*. — 2011. — Pp. 1 – 5.
51. Deng, L. Recent advances in deep learning for speech research at Microsoft /L. Deng // *Acoustics, Speech and Signal Processing (ICASSP)*. — 2013. — Pp. 8604–8608.
52. Konolige, K. Projected texture stereo / K. Konolige // *Robotics and Automation (ICRA)*. — 2010. — Pp. 23–28.
53. Keyzers, D. Comparison and combination of state-of-the-art techniques for handwritten character recognition: topping the mnist benchmark / D. Keyzers // *arXiv*. — 2007. — no. 0710.2231. — Pp. 21–27.
54. Convolutional Neural Networks for Visual Recognition: [Электронный ресурс]. – Режим доступа: <http://cs231n.github.io>. – Дата доступа : 01.05.2017.

55. Выбор функции активации обучения нейронной сети [Электронный ресурс]. – Режим доступа: <https://monographies.ru/ru/book/section?id=2465>. – Дата доступа : 03.05.2017.
56. Manisha M. Face Recognition Using Neural Network: A Review// Manisha M.Kasar¹ , Debnath Bhattacharyya¹ and Tai-ho Kim². [Электронный ресурс]. – Режим доступа: http://www.sersc.org/journals/IJSIA/vol10_no3_2016/8.pdf - – Дата доступа : 03.06.2017.
57. Хемант Синг Миттал, Гарприт Каур Распознавание с использованием метода главных компонент и нейросети. – 2013.[Электронный ресурс]. – Режим доступа: <http://www.ijese.org/attachments/File/v1i6/F0266041613.pdf> – Дата доступа :04.06.2017.
58. N Jindal, V Kumar, ” Enhanced Face Recognition Algorithm using PCA with Artificial Neural Networks”, International Journal of Advanced Research in Computer Science and Software Engineering, vol 3, no. 6, (2013), pp. 864-872.
59. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Image net classification with deep convolutional neural networks”, Edited F. Pereira and C.J.C. Burges and L. Bottou and K.Q. Weinberger, publish by Curran Associates, Inc. , (2012), pp. 1097-1105.
60. S-H Yooa, S-K Oha, Witold Pedrycz, ” Optimized face recognition algorithm using radial basis function neural networks and its practical applications”, International journal on Neural Networks, volume 69, (2015), pp. 111-125.
61. H Liy, Z Linz, X Shenz, J Brandtz, GHua, ” A Convolutional Neural Network Cascade for Face Detection”, IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, (2015), pp. 5325 – 5334.
62. A Roy Chowdhury Tsung-Yu Lin Subhranshu Maji Erik Learned-Miller, ” Face Identification with Bilinear CNNs”, Computer vision and pattern recognition, (2015).

63. M.Nandini, P.Bhargavi, G.Raja Sekhar, "Face Recognition Using Neural Network", International Journal of Scientific and Research Publications, vol. 3, no. 3, (2013), pp. 1-5.
64. H A. Rowley, Student Member, "Neural Network-Based Face Detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA , (1996), pp. 203 – 208.
65. H A. Rowley, S Baluja, T Kanade, "Rotation Invariant Neural Network-Based Face Detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA, (1998), pp. 38 – 44.
66. One Millisecond Face Alignment with an Ensemble of Regression Trees. [Електронний ресурс]. – Режим доступу: <http://www.csc.kth.se/~vahidk/papers/KazemiCVPR14.pdf>. – Дата доступу: 01.06.2017.
67. Cmusatyalab/openface [Електронний ресурс]. – Режим доступу: <https://github.com/cmusatyalab/openface>. – Дата доступу: 09.06.2017.
68. Рентабельність виробництва і методика визначення її показників. [Електронний ресурс]. – Режим доступу: <http://buklib.net/books/29473/> – Дата доступу: 04.02.2016.
69. Прогнозування ефективності інвестиційного проекту. [Електронний ресурс]. – Режим доступу: http://pidruchniki.com/1566072162240/turizm/prognozuvannya_efektivnost_i_investitsiyного_proektu – Дата доступу: 04.02.2017.
70. Розробка та перевірка концепції товару. [Електронний ресурс]. – Режим доступу: <http://westudents.com.ua/glavy/35910-3-rozrobka-ta-perevrka-kontsepts-tovaru.html>. – Дата доступу: 04.02.2017.
71. Концепції маркетингової діяльності. [Електронний ресурс]. – Режим доступу: <http://referat-ok.com.ua/marketing/marketingova-diyalnist-2> – Дата доступу: 04.02.2017.