

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)

Кафедра Інформатики
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

ДОСЛІДЖЕННЯ ТА РЕАЛІЗАЦІЯ МЕТОДУ ІДЕНТИФІКАЦІЇ
ОБЛИЧЧЯ ДЛЯ ПРОХОДЖЕННЯ РЕЄСТРАЦІЇ В SMS СИСТЕМАХ
(тема)

Виконав:
студент 2 курсу, групи ІНФМ-19-1

Лапченко В.О.
(прізвище, ініціали)

Спеціальності 122 Комп'ютерні науки
(код і повна назва спеціальності)

Освітня програма Інформатика
(повна назва освітньої програми)

Керівник доц. Шафроненко А.Ю
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

Кобилін О.А.
(прізвище, ініціали)

2020 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту
(повна назва)Кафедра Інформатики
(повна назва)Рівень вищої освіти другий (магістерський)Спеціальність 122 Комп'ютерні науки
(код і повна назва)Освітня програма Інформатика
(повна назва освітньої програми)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«_____» _____ 20__ р.

ЗАВДАННЯ
НА АТЕСТАЦІЙНУ РОБОТУстудентові Лапченку Владиславу Олеговичу

(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження та реалізація методу ідентифікації обличчя для проходження реєстрації в CMS системах

затверджена наказом по університету від «23» жовтня _____ 2020 року № 1428Ст.

2. Термін подання студентом роботи до екзаменаційної комісії 30 листопада 2020 р.

3. Вихідні дані до роботи Математичні моделі розпізнавання образів, перелік використовуваних програмних засобів: Java Script, Node JS, Tensor Flow
теоретичні відомості про методи розпізнавання облич на зображеннях

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Аналіз проблеми розпізнавання образів

2. Аналіз предметної області

3. Особливості задачі розпізнавання облич

4. Огляд та класифікація нейронних мереж

5. Побудова математичної моделі методу розпізнавання облич

6. Програмна реалізація методу ідентифікації людей на зображеннях за допомогою мови програмування Java Script та бібліотек NodeJs і TensorFlow

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) Актуальність проблеми ідентифікації людей для проходження реєстрації в SMS системах, постановка задачі, математична модель методу розпізнавання облич, архітектура проекту, програмна реалізація, аналіз результатів.

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на атестаційну роботу	23.10.2020	
2	Аналіз завдання, підбір літератури	24.10.20-25.10.20	
3	Аналіз літератури з досліджуваної проблеми	26.10.20-27.10.20	
4	Аналіз технічних засобів	28.10.20-29.10.20	
5	Розробка методу	30.10.20-05.11.20	
6	Програмна реалізація	06.11.20-16.11.20	
7	Оформлення пояснювальної записки	17.11.20-25.11.20	
8	Перевірка на плагіат	26.11.20	
9	Рецензування	28.11.20	
10	Підготовка презентації та доповіді	03.12.20	
11	Занесення роботи в електронний архів	04.12.20	
12	Попередній захист атестаційної роботи	07.12.20	

Дата видачі завдання 23 жовтня 2020 р.

Студент _____
(підпис)

Керівник роботи _____ доц. Шафроненко А.Ю.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ/ABSTRACT

Пояснювальна записка до атестаційної роботи: 81 с., 33 рис., 41 джерело.

КОМП'ЮТЕРНИЙ ЗІР, РОЗПІЗНАВАННЯ ОБРАЗІВ, РОЗПІЗНАВАННЯ ОБЛИЧ, НЕЙРОННІ МЕРЕЖІ, ЗГОРТКОВА НЕЙРОННА МЕРЕЖА, CMS, JAVA SCRIPT, NODE JS, TENSOR FLOW.

Метою даного дослідження є огляд та реалізація методу згорткової нейронної мережі для ідентифікації облич для проходження реєстрації в CMS системах.

Об'єктом дослідження є послідовність різноракурсних зображень облич людей.

Досліджено предметну область та виявлено актуальні проблеми у сфері ідентифікації людей в CMS системах фармацевтичних компаній. Проведено аналіз методів розпізнавання образів. Проаналізовано типи нейронних мереж та обрано оптимальну для задачі розпізнавання облич. Використано метод згорткової нейронної мережі. Проведено дослідження методу та побудовано його математичну модель. Проведено аналіз засобів програмної реалізації методу.

У результаті роботи здійснена програмна реалізація системи для ідентифікації обличчя для проходження реєстрації в CMS системах.

COMPUTER VISION, PATTERN RECOGNITION, FACE RECOGNITION, NEURAL NETWORKS, CONVOLUTIONAL NEURAL NETWORK, CMS, JAVA SCRIPT, NODE JS, TENSOR FLOW.

The aim of the research is to develop a method of face recognition based on the a convolutional neural network for the authentication in CMS.

The object of the research is the sequence of different images of people's faces from different angles.

The subject area is researched and actual problems in the field of identification of people in CMS systems of pharmaceutical companies are revealed. The analysis of methods of pattern recognition is carried out. The types of neural networks are analyzed and the optimal one for the task of face recognition is chosen. The method of convolutional neural network is used. The research of the method is carried out and its mathematical model is built. The analysis of means of software realization of a method is carried out.

As a result of the work the software implementation of the system for face identification for registration in CMS systems is carried out.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	7
Вступ.....	9
1 Огляд методів комп'ютерного зору та аналіз предметної області	10
1.1 Комп'ютерний зір	10
1.1.1 Визначення комп'ютерного зору.....	10
1.1.2 Основні методи комп'ютерного зору.....	12
1.1.3 Актуальність методів комп'ютерного зору.....	19
1.2 Ідентифікація людей.....	21
1.2.1 Особливості розпізнавання облич	24
1.2.2 Загальні етапи алгоритмів ідентифікації облич	26
1.2.3 Методи розпізнавання облич	27
1.2.4 Існуючі проблеми ідентифікації облич.....	28
1.3 Аналіз предметної області	29
1.3.1 Ідентифікації людини за обличчям, як засіб систем безпеки.....	29
1.3.2 Присутні на ринку системи ідентифікації облич.....	31
1.3.3 CMS системи	32
1.4 Постановка задачі дослідження.....	34
2 Математична модель згорткової нейронної мережі	36
2.1 Визначення згорткової нейронної мережі.....	36
2.2 Конструкція згорткової нейронної мережі.....	36
2.2.1 Вихідний шар.....	38
2.2.2 Згортковий шар	39
2.2.3 Підвибірковий шар.....	42
2.2.4 Повнозв'язний шар	43
2.2.5 Вихідний шар.....	45
2.3 Архітектура класичної згорткової нейронної мережі	45
2.4 Архітектура узагальненої згорткової нейронної мережі	47
2.5 Багатозадачна каскадна згорткова нейронна мережа	49

	6
2.6 Детектор MultiBox	52
3 Комп'ютерна модель фільтрації зображень	58
3.1 Обґрунтування вибору середовища програмної реалізації	58
3.1.1 Java Script	58
3.1.2 Node.JS	60
3.1.3 Tensorflow	61
3.2 Вимоги до програми	64
3.3 Програмна реалізація.....	64
3.4 Тестування розробленої моделі.....	66
Висновки	73
Перелік джерел посилання	76

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ЗНМ – згорткова нейронна мережа

НМ – нейронна мережа

ХР – характерні риси

2D – two dimensional

3D – three dimensional

AGI – artificial general intelligence

AI – artificial intelligence

API – application programming interface

CMS – content management systems

CNN – convolutional neural network

CV – computer vision

DHTML – dynamic hypertext markup language

DL – deep learning

DOM – document object model

FCN – fully convolutional network

FRCNN – fast regions with convolutional neural network

GWN – gabor wavelet network

IT – informational technologies

JS – java script

ML – machine learning

MTCNN – multi-task cascaded convolutional network

MV – machine vision

NMS – non max suppression

RCNN – regions with convolutional neural network

REST – representational state transfer

RFID – radio-frequency identification

SLAM – simultaneous localization and mapping

SSD – single shot detector

TCNN – tiled convolutional neural network

VGG – visual geometry group

W3C – world wide web consortium

ВСТУП

Багато напрямів науки, техніки і виробництва в значній мірі орієнтовані на розвиток систем, в яких інформація має характер поля. При обробці такої інформації виникає ряд складних наукових, технічних і технологічних проблем, однією з яких є обробка та розпізнавання зображень. Про важливість цієї проблеми свідчить той факт, що дослідження з розпізнавання образів, аналізу зображень та мови включені в перелік пріоритетних напрямів розвитку науки і техніки і критичних технологій федерального рівня.

Розпізнавання зображень знаходить застосування в різних додатках: контроль топології друкованих плат, текстури тканини, контроль доступу до інформації щодо ідентифікації особи (біометрична ідентифікація), доступ до об'єктів обмеженого доступу, оперативний пошук в картотеці зображень, дактилоскопія тощо.

Широке поширення отримують біометричні системи ідентифікації людини.

На відміну від традиційних систем ідентифікації, що вимагають знання пароля, наявності ключа, ідентифікаційної картки або іншого ідентифікує предмета, який можна забути або втратити, біометричні системи ґрунтуються на унікальних біологічних характеристиках людини, які важко підробити і які дозволяють однозначно визначити конкретну людину. До числа таких характеристик відносяться відбитки пальців, форма долоні, візерунок райдужної оболонки, зображення сітківки ока, індивідуальні характеристики особи.

У даному дослідженні розглядається саме ця задача. Основною метою роботи є дослідження та реалізація методу ідентифікації обличчя для проходження реєстрації в CMS системах.

1 ОГЛЯД МЕТОДІВ КОМП'ЮТЕРНОГО ЗОРУ ТА АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Комп'ютерний зір

Комп'ютерний зір – це прикордонна область знань в сфері цифрових технологій, штучного інтелекту та анатомії. Як будь-яка прикордонна область вона цікава для вивчення і непередбачувана; багато корисні ідеї не мають під собою теоретичних основ, а деякі теорії марні на практиці; вивчені області різноманітні і, здається, що будь-який зв'язок між ними відсутній [1].

На думку Девіда Форсайт і Жан Понс, розробки в області комп'ютерного зору є одними з найбільш затребуваних в сфері сучасних комп'ютерних технологій. Комп'ютерний зір потрібний на виробництві, при управлінні роботами, при автоматизації процесів, в медичних і військових додатках, при спостереженні з супутника і при роботі з персональним комп'ютером, зокрема, при пошуку цифрового зображення, або виявленні об'єкта на знімку.

1.1.1 Визначення комп'ютерного зору

Комп'ютерний зір – це така дисципліна (розділ штучного інтелекту), яка займається вилученням інформації з зображень, причому зображення можуть бути різного типу. Це можуть бути фотографії, можуть бути відео, можуть бути набори фотографій або медичний знімок з магнітно-резонансного томографа [2].

Комп'ютерний зір (Computer Vision, CV), в тому числі машинний зір (Machine Vision, MV) – це автоматична фіксація і обробка зображень, як нерухомих, так і рухомих об'єктів за допомогою комп'ютерних засобів.

Перші спроби змусити комп'ютер «бачити» були ще до початку 60-х років 20 століття. Однак лише в останні роки в зв'язку з підвищенням

обчислювальних потужностей і швидкодії процесорів, обсягів пам'яті, підвищенням роздільної здатності та інших параметрів камер, розвитком смуги пропускання каналів зв'язку, а також з появою таких технологій, як машинне глибоке навчання (Machine / Deep Learning), штучний інтелект AI (Artificial Intelligence) технології CV / MV стали знаходити все більше застосувань в різних галузях промисловості і повсякденному житті людей.

В останні роки CV стало активно використовуватися в промисловості, в т.ч. в таких галузях, як автомобілебудування, харчова промисловість, фармацевтика, виробництво мікроелектронних виробів і багатьох інших.

Завдання CV полягають, головним чином, в отриманні корисної інформації (insight) з фото- або відеозображень. Найбільш вживаними завданнями CV можуть бути:

- завдання калібрування камер і оптичних систем, як складаються з однієї камери, так і набору камер;
- завдання визначення руху по зображеннях;
- завдання визначення перешкод по ходу руху;
- у 3D-хмарі по стерео-камерами або набору камер;
- за однією камері за рахунок руху;
- завдання розпізнавання об'єктів на сцені;
- завдання просторової реконструкції сцени;
- завдання локалізації зображення в задалегідь відомій сцені;
- завдання аналізу відмінності в наборі зображень.

Будь-які завдання, пов'язані з комп'ютерним зором, в тому числі і завдання ідентифікації особистості, вимагають істотних технічних ресурсів. В силу цих обставин алгоритми, вирішальні це завдання, повинні відповідати наступним критеріям: швидкість ідентифікації, мінімальний розмір еталонної вибірки, кількість помилкових спрацьовувань, підтримка всіх найбільш популярних програмно-апаратних засобів.

Під апаратними засобами розуміються центральні процесори різної архітектури і підтримуваним набором інструкцій, під еталонної вибіркою

мається на увазі колекція фотознімків ідентифікованого об'єкта. Кожний фотознімок повинен бути унікальний в рамках даної вибірки, отже, створення колекцій одного об'єкта з відносно великого числа фотознімків є трудовитрати процедурою.

1.1.2 Основні методи комп'ютерного зору

Основні підходи до вирішення завдань комп'ютерного зору:

- контурний аналіз;
- пошук за шаблоном (template matching);
- пошук поза шаблонів, зіставлення по ключових точках (feature detection, description matching);
- поєднання даних (Data Fusion).

Комп'ютерний зір не обмежується тільки цими основними методами, наприклад, можна виділити так звані генетичні алгоритми, що застосовуються, зокрема, для розпізнавання осіб [3].

1.1.2.1 Контурний аналіз

Контур об'єкта – це крива, відповідна кордоні об'єкта на зображенні. У цьому методі аналізується не повне зображення об'єкта, а тільки його контур, що істотно знижує складність алгоритмів і обчислень при обробці.

Обмеження методу контурного аналізу:

- при однаковій яскравості з фоном об'єкт може не мати чіткої межі на зображенні або воно може бути «зашумлено» перешкодами, що призводить до неможливості виділення контуру;
- перекриття об'єктів або їх угруповання призводять до того, що контур виділяється неправильно і не відповідає кордоні об'єкта;

– слабка стійкість до перешкод, що призводить до того, що будь-яке порушення цілісності контуру або погана видимість об'єкта призводять або до неможливості детектування, або до помилкових спрацьовувань [4].

1.1.2.2 Пошук за шаблоном

Найпоширеніший метод розпізнавання об'єктів в комп'ютерному зорі – пошук відповідності шаблонам зображень, щоб визначити, чи є заданий об'єкт на зображенні, і, якщо є, де він знаходиться на зображенні. Додатки методу: розпізнавання транспортних засобів, прокладка маршрутів для мобільних роботів, виробництво і додатки в медицині, тощо.

Проста відповідність – один з основних методів знаходження потрібного об'єкту на зображенні при пошуку за шаблоном. Метод полягає в покроковому скануванні шаблоном вихідного зображення, при кожному кроці якого вимірюється або розраховується ступінь відповідності ділянки зображення шаблоном. В кінці сканування на зображенні виділяється область, найбільш відповідна шаблоном.

Відповідність на базі характерних рис (Feature-based matching). Метод відповідності з характерних рис (ХР), застосуємо, коли як зображення, так і шаблон містять більше відповідностей по ХР і контрольних точках, ніж по цілісному образу. В цьому випадку, ХР можуть включати точки, криві або моделі поверхні, які перевіряються на відповідність шаблону. Мета такої перевірки – знайти парні зв'язки між метою («референсом») і частиною зображення з використанням просторових співвідношень або ХР.

1.1.2.3 Відповідність на базі областей

Методи знаходження відповідності на базі областей (Area-based), які також називаються кореляційними методами, засновані на комбінованому алгоритмі знаходження характерних рис і відповідності шаблоном (template matching). Такий метод добре працює, якщо шаблони не мають помітних загальних ХР з зображенням, оскільки порівняння відбувається на піксельному рівні. Відповідності вимірюються за показниками інтенсивності шаблону і зображення.

У деяких випадках знаходження прямого відповідності між шаблоном і зображенням неможливо (рис. 1.1). Тому, при знаходженні відповідності використовуються власне значення (eigenvalue) і власний простір (eigenspace). Ці величини містять інформацію, необхідну для порівняння образів при різних умовах освітленості, контрастності контурів або збігу по положенню об'єктів [5].

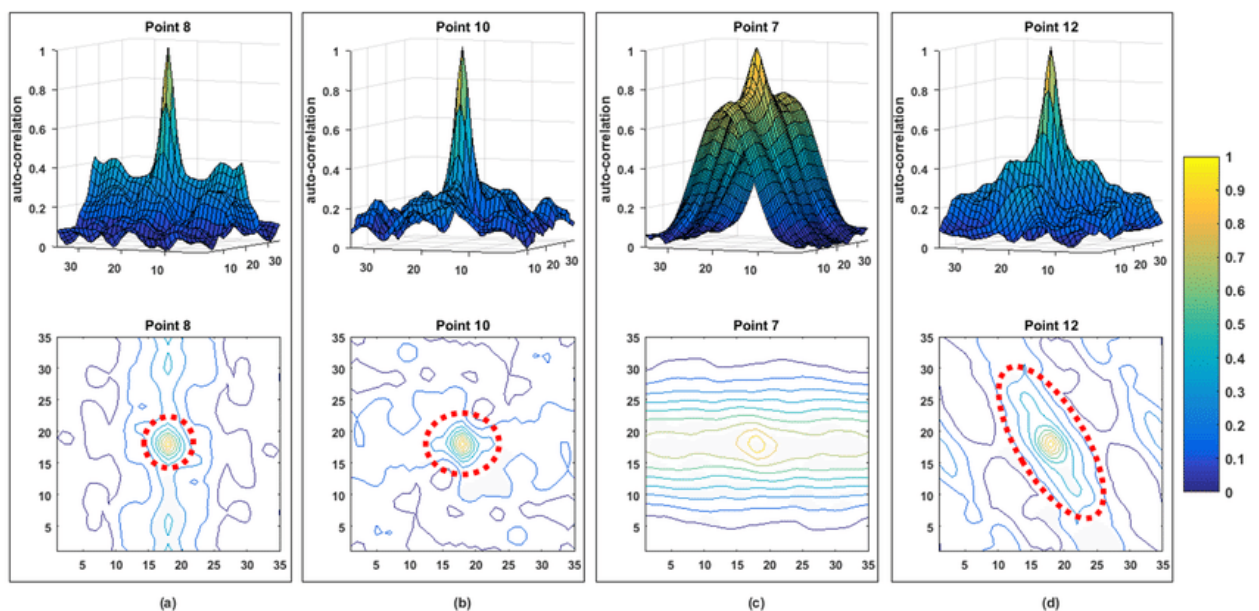


Рисунок 1.1 – Застосування Area-based метода в геодезії

1.1.2.4 Кореляція зображень

У цьому методі вимірюються метрики подібності (similarity metric) між вихідним зображенням і шаблоном. На відміну від методу простої відповідності, вихідне зображення і шаблон можуть мати різні інтенсивності зображення або рівні шуму. У цьому випадку порівняння проводиться по метриці подібності на основі кореляцій між шаблоном і оригіналом.

Комп'ютерний зір значно розширює можливості контролю якості продукції (фактично переводячи контроль на новий рівень) безпосередньо в виробничому процесі, а не після виготовлення деталі або продукту.

Автоматична візуальна інспекція (визначення дефектів) за допомогою комп'ютерного зору сьогодні значно перевершує ручні методи інспекції по точності, швидкості, легкості виконання і вартості.

Термін «неймережі» (Neural networks) був дуже популярний в кінці 1980-х – початку 1990-х років. Неймережі (рис. 1.2) складаються з шарів, «нейронів», які являють собою обчислювальні вузли, що імітують роботу нейронних клітин живого організму. Ці мережі можуть передавати інформацію тільки в одному напрямку і можуть навчатися на прикладах (для класифікації об'єктів або регресивного аналізу) [6].

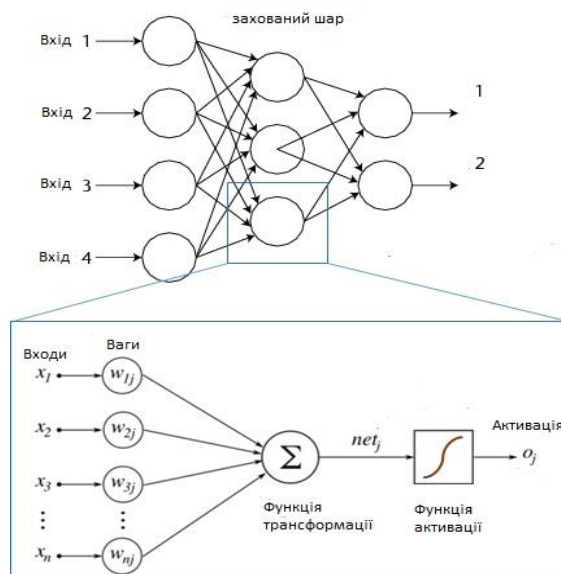


Рисунок 1.2 – Приклад нейронної мережі

1.1.2.5 Глибоке навчання в системах комп'ютерного зору

Глибоке навчання (Deep learning) може бути корисно в задачах, коли базовий елемент (окремий піксель зображення, одна частота сигналу, одне слово або буква) не несе великого смислового значення, проте, комбінація таких елементів має корисне значення.

Системи глибокого навчання можуть отримувати такі корисні комбінації без втручання людини (Unsupervised Feature Learning) [7].

Глибокі нейромережі (deep neural network), з більш ніж 1-2 шарами, раніше здавалася або нереалізованим, або непрактичними у використанні. До 2006 року, зовнішні шари нейромережі були нездатні до вилучення характерних рис (features) вхідних зображень, оскільки алгоритми навчання нейромереж залишалися недосконалими.

На рисунку 1.3 показаний приклад системи CV з машинним навчанням на прикладі завдання сегментування сцени за трьома типами: «горизонтальний», «вертикальний» і «небо».

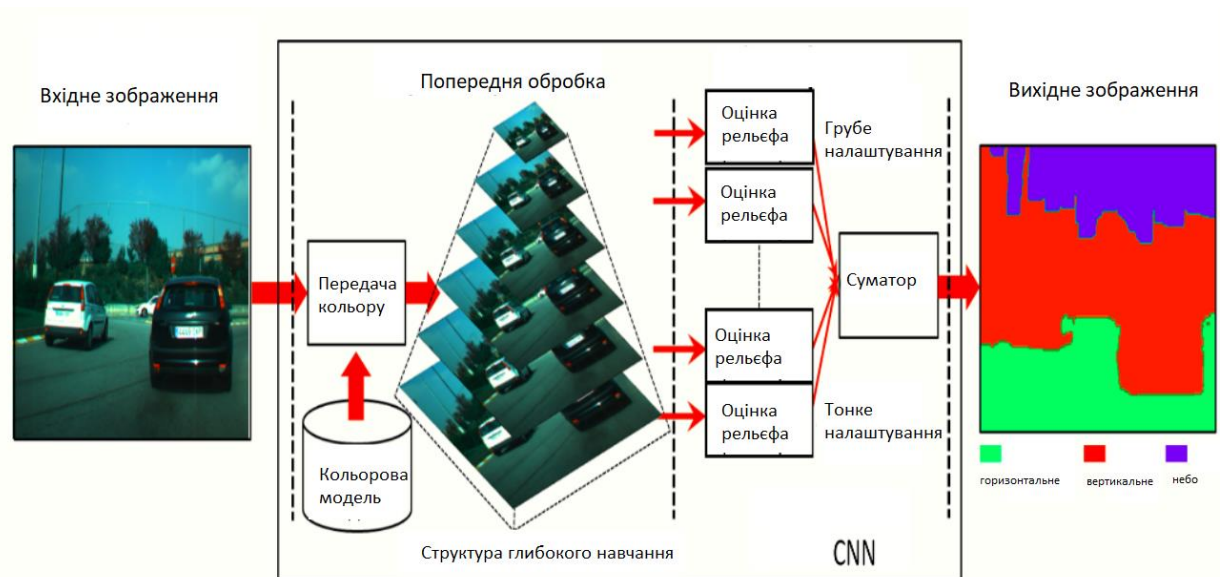


Рисунок 1.3 – Приклад системи глибокого навчання

Піксельні дані від вихідного зображення з корекцією кольору надходять на нейромережу глибокого навчання, де проводиться попередня обробка

зображення і розпізнається, до якого типу рельєфу належить кожен піксель з певним ступенем ймовірності [8].

Приклад розпізнавання рукописних цифр з використанням простої однорівневої нейромережі наведено на рисунку 1.4.

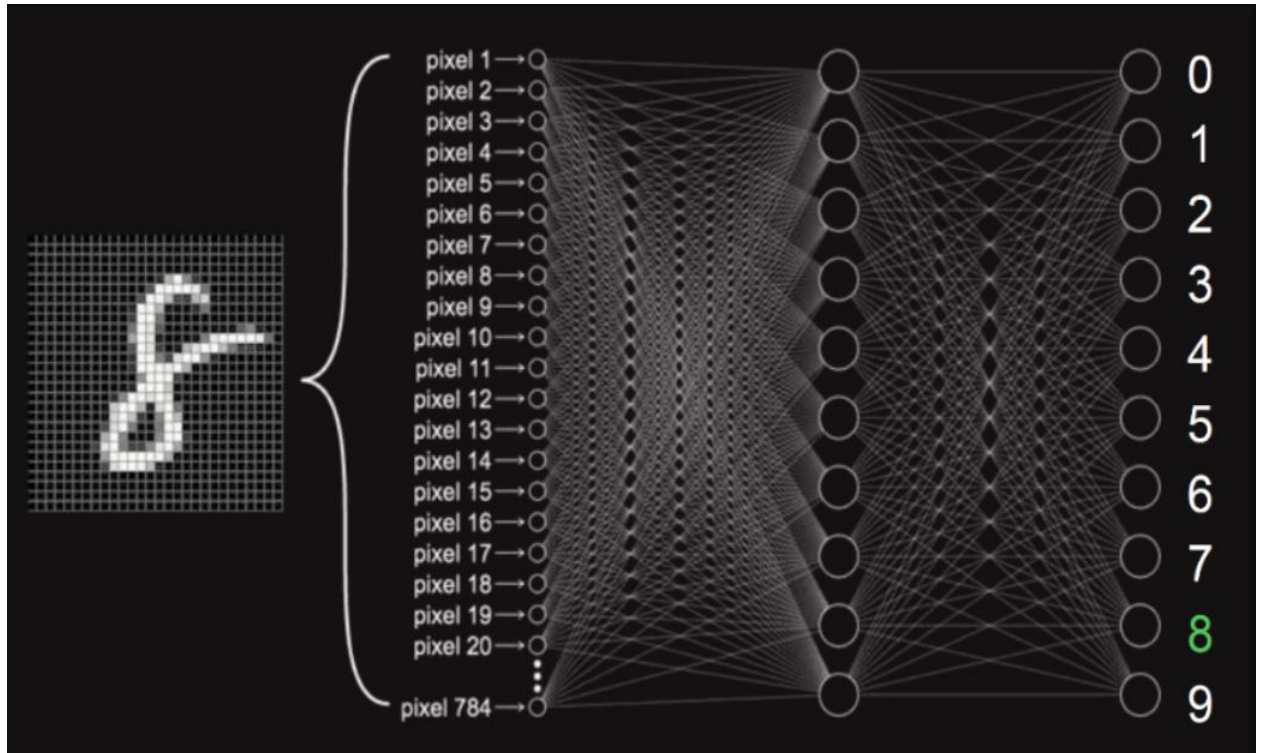


Рисунок 1.4 – Приклад нейронної мережі для розпізнавання числових символів

1.1.2.6 Детектування і розпізнавання об'єктів

Набори характерних рис для обробки зображень в комп'ютерному зорі можуть, наприклад, являти собою елементи зображення, такі як точки, краю, лінії або межі об'єктів. Інші приклади характерних рис відносяться до руху в послідовності зображень, до форм, представлених у вигляді кривих між областями зображення, або до властивостей цих областей [9].

Детектування об'єктів – це знаходження примірників об'єктів на зображенні. При розпізнаванні об'єктів не тільки встановлюється факт

наявності об'єкта на зображенні, але також і визначається його розташування на зображенні. На рисунку 1.5 показані приклади детектування (зліва) і розпізнавання об'єктів (праворуч).

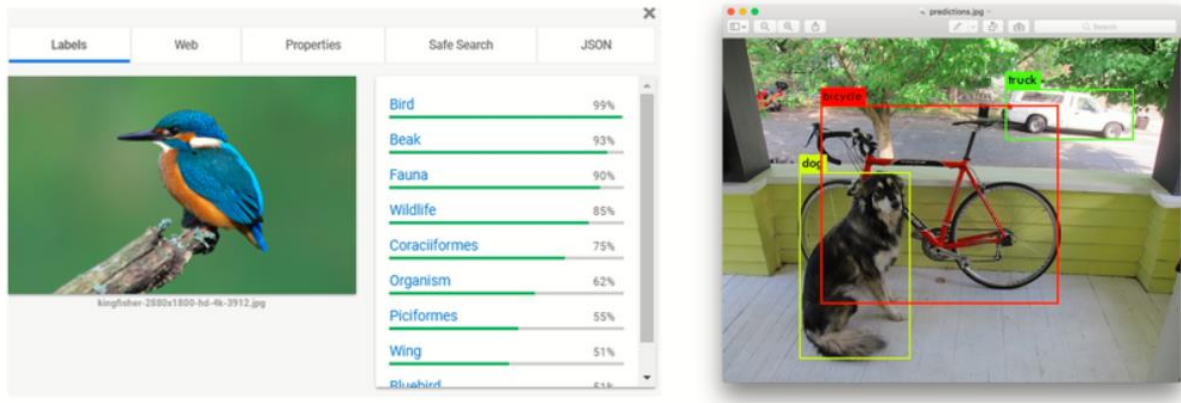


Рисунок 1.5 – Розпізнавання об'єктів на фото

Детектування об'єктів передбачає зіставлення двох і більше зображень при пошуку зображень унікальних об'єктів, наприклад, архітектурних споруд, скульптур, картин і т.д., виявлення на зображеннях класів об'єктів різного ступеня спільності (автомобілів, тварин, меблів, осіб людей і т.д., а також їх підкласів), категоризація сцен (місто, ліс, гори, узбережжя і т.д.) [10].

Додатки для детектування об'єктів також досить різноманітні: сортування зображень в домашніх цифрових фотоальбомах, пошук товарів за їхніми зображеннями в інтернет-магазинах, витяг зображень в геоінформаційних системах, біометрична ідентифікація особи, цільовий пошук зображень в соціальних мережах і багато іншого.

Розпізнавання такого розмаїття об'єктів і додатків обумовлює необхідність використання методів машинного і глибокого навчання.

Деякі інші приклади застосування методу розпізнавання поза шаблонів: фотограмметрія, виявлення перешкод, одночасна локалізації об'єктів і побудова карти в невідомому просторі (SLAM), дефектоскопія.

Крім термінів «детектування» (detection) і «розпізнавання» (recognition) в технологіях комп'ютерного зору використовуються також терміни

«класифікація» (classification) і «локалізація» (localization), а також «сегментація» (segmentation) об'єктів [11].

1.1.3 Актуальність методів комп'ютерного зору

Комп'ютерний зір – швидко зростаюча область цифрових технологій, яка зачіпає багато сторін повсякденного життя.

Компанія Apple впровадила функцію розпізнавання осіб в нові моделі iPhone, придбавши такі компанії, як PrimeSense, RealFace і Faceshift. Американський портал AngelList, який об'єднує стартапи та інвесторів, склав список з 529 нових компаній, які працюють в області комп'ютерного зору. Середня капіталізація таких стартапів становить 5,2 млн. дол. Багато стартапів залучають капітал від 5 до 10 млн. дол. Портал відзначає, що потік інвестицій в комп'ютерному зорі наростає. Заміна людського зору на комп'ютерне у багатьох областях – дуже вигідне вкладення капіталу.

Точність аналізу відеоінформації комп'ютером весь час зростає і застосування комп'ютерного зору може дати велику економію коштів поряд з поліпшенням якості [12].

Можна виділити п'ять основних тенденцій розвитку комп'ютерного зору:

1. Зростання промислових систем комп'ютерного зору. CV для медичних пристроїв, фармацевтики, виробництва харчових продуктів, автомобільної промисловості надає більш високий рівень контролю якості, а CV для промисловості, як очікується, в 2019 році стане основним трендом в області комп'ютерного зору.

2. Хмарні системи глибокого навчання. Алгоритми глибокого навчання і класифікатори нейромереж дозволять більш швидко і точно проводити класифікацію та розпізнавання зображень від систем CV. У найближчі роки число таких розробок значно зросте.

3. Робототехніка. Використання промислових роботів стрімко збільшується. Тому попит на системи CV для роботів буде рости.

4. Зростання вимог до параметрів оптики для CV, який викликаний зростанням вимог до чіткості і роздільної здатності зображень CV. Розробляються і виробляються сенсори для CV-камер з більшою роздільною здатністю і з великою кількістю пікселів, проте, без якісної оптики ці удосконалення будуть малокорисні. Тому розробляються такі інноваційні рішення, як мікролінзи на кожен піксель і ін., які кардинально можуть підвищити параметри роботи оптичних систем, які вже підійшли до своїх технологічних меж в традиційних рішеннях.

5. Штучний інтелект загального застосування (AGI – Artificial General Intelligence). Термін Artificial General Intelligence, що з'явився відносно недавно, означає здатність комп'ютера здійснювати абстраговані заключення або, принаймні, імітувати цей процес, тим самим наближаючи його до мислення людини. Однак AGI поки знаходиться на самій ранній стадії розвитку. Абстрактне мислення залишається поки нерозв'язною проблемою для штучного інтелекту. Саме з цієї причини технології AGI знаходяться на кривій Гартнера на самому початку підйому «тригера інновацій». Комп'ютерний зір є однією з важливих складових технологій для штучного інтелекту AI (рис. 1.6) [13].

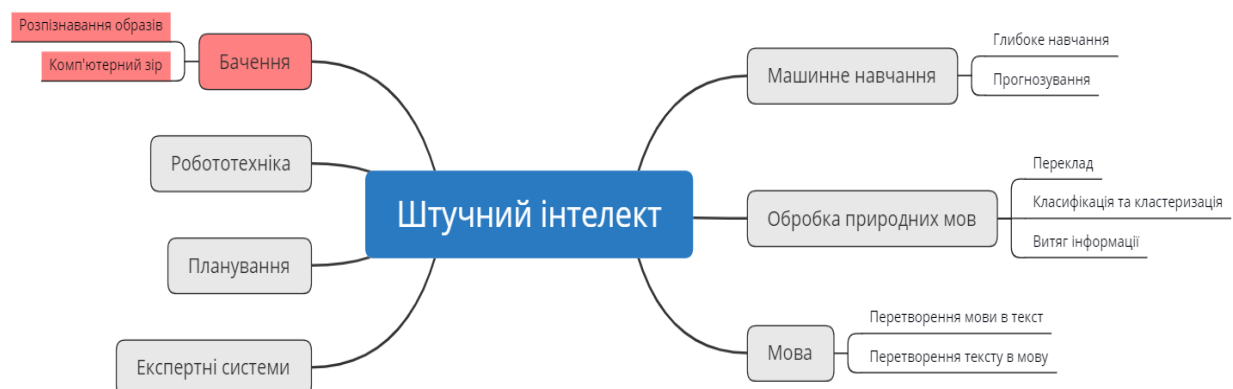


Рисунок 1.6 – Штучний інтелект загального застосування та його складові

Діапазон застосувань CV в останні 10-15 років значно розширився. Компанія Tractica в своєму звіті про ринок комп'ютерного зору в 2014 році вказує шість областей CV. В 2016 році в новій версії звіту Tractica вказує вже вісім областей застосування комп'ютерного зору: додані Retail (роздрібна торгівля) і Agriculture (сільське господарство).

Компанія Market Research Future оцінює світовий ринок CV в 2017 році в 9,2 млрд. Доларів США і очікує, що до 2023 року він перевищить 48,3 млрд. дол. При стійкому зростанні, який збільшується після 2020 року (рис. 1.7) [14].

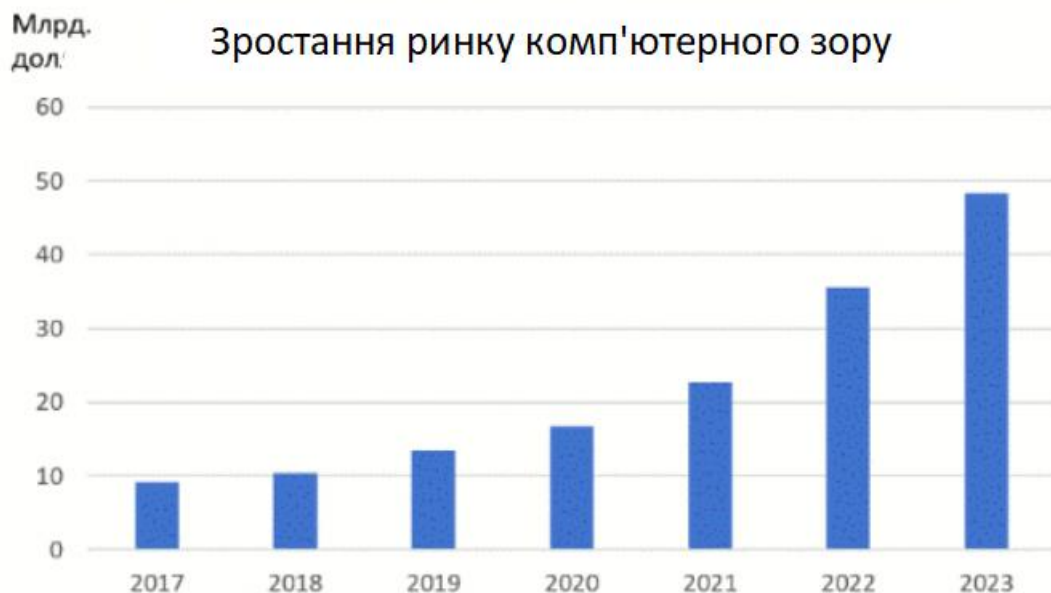


Рисунок 1.7 – Графік росту ринку комп'ютерного зору

1.2 Ідентифікація людей

Ідея використовувати індивідуальні характеристики людини для його ідентифікації не нова. На сьогоднішній день відомий ряд технологій, які можуть бути задіяні в системах безпеки для ідентифікації особистості по:

- відбитками пальців (як окремих, так і руки в цілому);
- рисами обличчя (на основі оптичного та інфрачервоного зображень);

- райдужній оболонці очей;
- голосу;
- іншим характеристикам.

У всіх біометричних технологій існують загальні підходи до вирішення завдання ідентифікації, хоча всі методи відрізняються зручністю застосування, точністю результатів. Будь-яка біометрична технологія застосовується поетапно:

- сканування об'єкта;
- витяг індивідуальної інформації;
- формування шаблону;
- порівняння поточного шаблону з базою даних.

Біометрична система розпізнавання встановлює відповідність конкретних фізіологічних або поведінкових характеристик користувача деякому заданому шаблону. Зазвичай біометрична система складається з двох модулів: модуль реєстрації та модуль ідентифікації.

Модуль реєстрації «навчає» систему ідентифікувати конкретну людину. На етапі реєстрації відеокамера або інші датчики сканують людини для того, щоб створити цифрове представлення його вигляду. Сканування особи триває близько 20 – 30 с, в результаті чого формуються кілька зображень. В ідеальному випадку, ці зображення будуть мати злегка різні ракурси і виразу обличчя, що дозволить отримати більш точні дані. Спеціальний програмний модуль обробляє це уявлення і визначає характерні особливості особистості, потім створює шаблон. Існують деякі частини обличчя, які практично не змінюються з плином часу, це, наприклад, верхні обриси очниць, області навколишні вилиці, і краю рота. Більшість алгоритмів, розроблених для біометричних технологій, дозволяють враховувати можливі зміни в зачісці людини, так як вони не використовують для аналізу ділянку обличчя вище межі росту волосся. Шаблон зображення кожного користувача зберігається в базі даних біометричної системи [15].

Модуль ідентифікації отримує від відеокамери зображення людини і перетворює його в той же цифровий формат, в якому зберігається шаблон. Отримані дані порівнюються з збереженим в базі даних шаблоном для того, щоб визначити, чи відповідають ці зображення один одному. Ступінь подібності, необхідна для перевірки, є певний поріг, який може бути відрегульований для різного типу персоналу, потужності РС, часу доби і ряду інших чинників.

Ідентифікація може виконуватися у вигляді верифікації, аутентифікації або розпізнавання. При верифікації підтверджується ідентичність отриманих даних і шаблону, що зберігається в базі даних. Аутентифікація – підтверджує відповідність зображення, одержуваного від відеокамери одному з шаблонів, що зберігаються в базі даних. При розпізнаванні, якщо отримані характеристики і один зі збережених шаблонів виявляються однаковими, то система ідентифікує людину з відповідним шаблоном [16].

При використанні біометричних систем, особливо системи розпізнавання по обличчю, навіть при введенні коректних біометричних характеристик не завжди рішення про аутентифікації вірно. Це пов'язано з рядом особливостей і, в першу чергу, з тим, що багато біометричні характеристики можуть змінюватися. Існує певна ступінь ймовірності помилки системи. Причому при використанні різних технологій помилка може мати відчутні відмінності. Для систем контролю доступу при використанні біометричних технологій необхідно визначити, що важливіше не пропустити «чужого» або пропустити всіх «своїх».

Важливим фактором для користувачів біометричних технологій в системах безпеки є простота використання. Людина, характеристики якого скануються, не повинен при цьому відчувати жодних незручностей. У цьому плані найбільш цікавим методом є, безумовно, технологія розпізнавання по обличчю. Правда, в цьому випадку виникають інші проблеми, пов'язані в першу чергу, з точністю роботи системи.

Незважаючи на очевидні переваги, існує ряд негативних упереджень проти біометрії, які часто викликають питання про те, чи не будуть біометричні дані використовуватися для стеження за людьми і порушення їх права на приватне життя. Через сенсаційних заяв і необгрунтованою галасу сприйняття біометричних технологій різко відрізняється від реального стану [17].

І все ж, використання біометричних методів ідентифікації набуло особливої актуальності в останні роки. Особливо гостро дана проблема проявилася після подій 11 вересня 2001 року в США. Світова спільнота усвідомила ступінь зростання загрози тероризму в усьому світі і складність організації надійного захисту традиційними методами. Саме ці трагічні події послужили відправною точкою для посилення уваги до сучасних інтегрованих систем безпеки. Загальновідомо думку, що якби контроль в аеропортах був суворіше, то нещастя можна було б уникнути. Та й сьогодні пошук винних в ряді інших випадків міг би бути істотно полегшений при використанні сучасних систем відеоспостереження в інтеграції з системами розпізнавання осіб [18].

1.2.1 Особливості розпізнавання облич

Залежно від конкретних умов структура і реалізація окремих кроків алгоритму можуть відрізнятися. У найбільш складному випадку, при використанні системи виявлення та ідентифікації людини по зображенню його особи в сильно змінюються, з великим потоком вхідних даних (робота на міських вулицях з інтенсивним рухом, в метро, аеропортах і т.д.), Потрібне використання максимально доступною інформації для досягнення задовільних результатів роботи алгоритму.

Алгоритм повинен вміти ефективно відсікати статичні і повільно змінюються елементи сцени, працювати в різних умовах освітленості,

впізнавати фігуру людини під різними ракурсами, відстежувати пересування великої кількості людей і автоматично вибирати момент, відповідний для виконання ідентифікації даної людини (наприклад, коли можна отримати фронтальне зображення особи з достатнім дозволом).

Для забезпечення таких можливостей алгоритму необхідна певна апаратна насиченість системи, що включає багато інсценує огляд і аналіз сцени з можливістю виділення 3D-структури сцени, швидкісне введення відеопотоку для фільтрації елементів сцени за параметрами руху, використання кольору для виділення елементів сцени. Крім того потрібні камери з високою роздільною здатністю і хорошою оптикою для забезпечення можливо більшої дальності достовірної ідентифікації.

У простіших випадках, при статичній сцені і обмеженому потоці подій (появ людей), можливе використання більш простої структури апаратного забезпечення і алгоритму, наприклад, стереопари або однієї камери і заздалегідь підготовленої моделі сцени може бути досить для достовірного визначення факту знаходження людини в зоні контролю, виділення його фігури і ідентифікації. Завдання визначення факту присутності людини на сцені, вимагає від алгоритму певного рівня інтелекту. Це не повинна бути система, що реагує просто на факт зміни сцени. Алгоритм виявлення людини не повинен давати неправдиві тривоги при змінах освітленості, русі тіней від статичних об'єктів, появи в зоні контролю тварин і т. д.

У випадку, коли це необхідно, з'являється проблема створення адекватного опису сцени. Це опис може представляти тривимірну модель сцени, вірогідну модель розподілу кольорів або яскравостей елементів сцени або систему ознак, що відрізняє елементи сцени від об'єктів розпізнавання (в нашому випадку – людських фігур). Відносини між елементами сцени, які вважаються фоном, або елементів переднього плану можуть змінюватися. Та ж фігура людини, якщо її зображення менше деякого порогового значення, що визначається здатністю оптичної системи, може бути віднесена до елементів

фону, так як її аналіз є непродуктивним для виконання основного завдання – ідентифікації людини [19].

1.2.2 Загальні етапи алгоритмів ідентифікації облич

У найзагальнішому випадку алгоритм вирішення задачі виявлення та ідентифікації людини по зображенню його особи складається з наступних очевидних кроків:

- виявлення факту присутності людини на аналізованій сцені;
- виділення фігури людини;
- виділення голови;
- визначення ракурсу спостереження голови (анфас, профіль);
- виділення особи;
- порівняння з еталонами і ідентифікація.

Вибір алгоритму, який використовується для ідентифікації людини по зображенню його особи, також залежить від конкретних умов його застосування. Наприклад, із завданням розпізнавання в строго обмеженому колективі легко справляється багат шарова нейронна мережа. У той же час завдання виявлення конкретної людини в натовпі (з невизначеним складом) вимагає застосування витончених методів для зниження рівня помилок. Швидше за все, в цьому випадку буде потрібно багаторівнева система, що містить безліч аналізаторів, що працюють в різних просторах ознак, з прийняттям рішення методом голосування. На початкових етапах роботи система ідентифікації повинна відсікати свідомо невідповідних кандидатів і використовувати час, що залишився безліч кандидатів для прийняття остаточного рішення про ідентифікацію.

Нижче представлений огляд по існуючим методам виявлення та ідентифікації людини по зображенню його особи. В огляд включалися тільки ті методи, які, на думку авторів, найбільш широко використовуються в

сучасних алгоритмах виявлення людини і його персональної ідентифікації по зображеннях особи в природніх умовах [20].

При всьому різноманітті різних алгоритмів і методів розпізнавання зображень, типовий метод розпізнавання складається з трьох основних компонентів:

- перетворення вихідного зображення в стандартний вигляд;
- виділення ключових характеристик;
- механізм класифікації (моделювання): кластерна модель, метрика, нейронна мережа і т. п.

1.2.3 Методи розпізнавання облич

Найпопулярнішими методам, які використовуються для розпізнавання образів, у тому числі й облич є наступні:

- метод головних компонент;
- лінійний дискримінантний аналіз;
- синтез об'єктів лінійних класів;
- гнучкі контурні моделі особи;
- порівняння еластичних графів;
- методи, засновані на геометричних характеристиках обличчя;
- порівняння шаблонів;
- приховані Марковские моделі;
- багат шарові нейронні мережі;
- мережі Габорових вейвлетов (GWN) ;
- нейронні мережі Хопфілда;
- згорткова нейронна мережа.

У даному дослідженні більш детально розглянемо методи, пов'язані з нейронними мережами, у тому числі й згорткові нейронні мережі. Детальний огляд яких описано у розділі 2.

1.2.4 Існуючі проблеми ідентифікації облич

Серйозною проблемою, що стоїть перед системами комп'ютерного зору, є велика мінливість візуальних образів, пов'язана зі змінами освітленості, забарвлення, масштабів, ракурсів спостереження. Крім того, люди мають звичку ходити по вулицях і в приміщенні одягненими, що призводить до суттєвої мінливості зображень одного і того ж людини. Однак найбільш складним завданням комп'ютерного зору є проблема усунення неоднозначності, що виникає при проектуванні тривимірних об'єктів реального світу на плоскі зображення. Колір і яскравість окремих пікселів на зображенні також залежить від великої кількості важко прогнозованих факторів. У число цих факторів входять:

- число і розташування джерел світла;
- колір і інтенсивність випромінювання;
- тіні або віддзеркалення від навколишніх об'єктів.

Завдання виявлення об'єктів на зображенні ускладнюється також більший обсяг даних, що містяться в зображенні. Зображення може містити тисячі пікселів, кожен з яких може мати важливе значення. Повне використання інформації, що міститься в зображенні, вимагає аналізу кожного пікселя на приналежність його об'єкту або фону з урахуванням можливої мінливості об'єктів. Такий аналіз може зажадати високих витрат в необхідній пам'яті і продуктивності комп'ютера [21].

Вирішення цієї проблеми лежить в правильному виборі опису об'єктів, для виявлення і розпізнавання яких створюється система. Опис об'єкта має його враховувати найбільш характерні особливості і бути досить

представницьким, щоб відрізнити даний об'єкт від інших елементів навколишнього сцени. Щоб уникнути суб'єктивності при виборі потрібного опису, можна використовувати методи автоматичного вибору відповідних характеристик об'єкта, які реалізуються в генетичних алгоритмах і при навчанні штучних нейронних мереж. У той же час існує ряд параметрів в описі об'єкта, які в даний час повинен вибрати дослідник, який розробляє систему виявлення і розпізнавання. До такого вибору відносяться:

- вибір між 2D і 3D-виставою сцени і об'єкта. Алгоритми, що використовують 2D-уявлення, зазвичай простіші, ніж 3D-алгоритми, але в той же час вимагають великого числа різних описів, відповідних поданням об'єкта в різних умовах спостереження;

- вибір між описом об'єкта як єдиного цілого або як системи, що складається з певної кількості взаємопов'язаних елементів;

- вибір між системою ознак, що ґрунтуються на геометричних чи інших описують специфіку об'єкта характеристиках.

1.3 Аналіз предметної області

Для реалізації дослідження у даній роботі було обрано тему ідентифікації облич співробітників для входу в CMS систему.

1.3.1 Ідентифікації людини за обличчям, як засіб систем безпеки

Традиційні методи ідентифікації особистості, в основі яких знаходяться різні ідентифікаційні карти, ключі або унікальні дані, такі як, наприклад, пароль не є надійними в тій мірі, яка потрібна на сьогоднішній день. Природним кроком в підвищенні надійності ідентифікаторів стали спроби використання біометричних технологій для систем безпеки [22].

Діапазон проблем, вирішення яких може бути знайдено з використанням нових технологій, надзвичайно широкий:

- запобігти проникненню зловмисників на охоронювані території і в приміщення за рахунок підробки, крадіжки документів, карт, паролів;
- обмежити доступ до інформації і забезпечити персональну відповідальність за її збереження;
- забезпечити допуск до відповідальних об'єктах тільки сертифікованих фахівців;
- уникнути накладних витрат, пов'язаних з експлуатацією систем контролю доступу (карти, ключі);
- виключити незручності, пов'язані з втратою, псуванням чи елементарним забування ключів, карт, паролів;
- організувати облік доступу і відвідуваності співробітників [23].

На практиці, при використанні систем розпізнавання осіб в складі стандартних електронних охоронних систем, передбачається, що людина, яку слід ідентифікувати, дивиться прямо в камеру. Таким чином, система працює з відносно простим двовимірним зображенням, що помітно спрощує алгоритми і знижує інтенсивність обчислень. Але навіть в цьому випадку завдання розпізнавання все ж не тривіальна, оскільки алгоритми повинні враховувати можливість зміни рівня освітлення, зміна виразу обличчя, наявність або відсутність макіяжу або очок.

Надійність роботи системи розпізнавання осіб дуже сильно залежить від декількох факторів:

- якість зображення. Помітно знижується ймовірність безпомилкової роботи системи, якщо людина, яку ми намагаємося ідентифікувати, дивиться не прямо в камеру або знятий при поганому освітленні;
- актуальність фотографії, занесеної до бази даних;
- величина бази даних.

Технології розпізнавання особи добре працюють зі стандартними відеокамерами, які передають дані і управляються персональним

комп'ютером, і вимагають дозволу 320×240 пікселів на дюйм при швидкості відео потоку, принаймні, 3 – 5 кадрів в секунду. Для порівняння – прийнятну якість для відео конференції вимагає швидкості відеопотоку вже від 15 кадрів в секунду. Більш висока швидкість відеопотоку при більш високому дозволі веде до поліпшення якості ідентифікації. При розпізнаванні осіб з великої відстані існує сильна залежність між якістю відеокамери і результатом ідентифікації [24].

1.3.2 Присутні на ринку системи ідентифікації облич

На сьогоднішній день розроблено ряд комерційних продуктів, призначених для розпізнавання осіб. Алгоритми, що використовуються в цих продуктах, різні і поки ще складно дати оцінку, яка з технологій має переваги. Лідерами зараз є такі системи: Visionic, Viisage і Miros.

В основі програми FaceIt компанії Visionic лежить алгоритм аналізу локальних ознак, розроблений в Університеті Рокфеллера. Одна комерційна компанія в Великобританії інтегрувала FaceIt в телевізійну антикримінальну систему під назвою Mandrake. Ця система шукає злочинців по відеоданих, які надходять з 144 камер, об'єднаних в замкнуту мережу. Коли встановлюється ідентичність, система повідомляє про це офіцерові безпеки. У Росії представником компанії Visionic є компанія «Данко».

Ще один лідер в цій області, компанія Viisage, використовує алгоритм, розроблений в Массачусетському технологічному інституті. Комерційні компанії і державні структури в багатьох американських штатах і в ряді інших країн використовують систему компанії Viisage разом з ідентифікаційними посвідченнями, наприклад, водійськими правами [25].

ZN Vision Technologies AG (Німеччина) пропонує на ринку ряд продуктів, в яких застосовується технологія розпізнавання осіб. Ці системи представляються на російському ринку компанією «Солинг».

В системі розпізнавання осіб TrueFace компанії Miros використовується технологія нейронних мереж, а сама система застосовується в комплексі видачі готівки корпорації Mr.Payroll і встановлена в казино та інших розважальних закладах багатьох штатів США.

1.3.3 CMS системи

Система управління контентом (CMS) – це програмне забезпечення, яке працює в вашому браузері. Вона дозволяє створювати, управляти і змінювати веб-сайт і його вміст, не маючи ніяких знань в області програмування. Система управління контентом надає вам графічний інтерфейс користувача. У ньому ви можете керувати всіма аспектами вашого сайту. Можливо створювати і редагувати контент, додавати зображення і відео, а також налаштовувати загальний дизайн сайту.

CMS – це програмне забезпечення на базі скриптів, яке дозволяє керувати вмістом ресурсу, змінювати його, переглядати і контролювати. Сьогоднішні системи мають широкую функціональність і складаються з величезної кількості модулів, кожен з яких відповідає за свої елементи. Програмне забезпечення допомагає складати типові сайти з блоків подібно до конструктора. Для цього практично не потрібно навіть навичок програмування. Умовно CMS можна розділити на кілька частин:

- сховище баз даних, де знаходиться інформація про користувачів, наповненні сайту і інших важливих сутності;
- сховище елементів інтерфейсу, з якими безпосередньо взаємодіє користувач при перегляді сайту;
- візуальний редактор, що допомагає з легкістю створювати сторінки.

Крім того, множинні модулі дозволяють додати до сайту ті або інші додаткові функції [26].

Сучасні CMS використовуються вкрай широко: без них складно обійтися будь-якої компанії, яка виходить на інтернет-майданчики і потребує власному сайті. На відміну від спеціалізованих ІТ-фірм, що володіють професійними командами фахівців, більшість непрофільних організацій не може забезпечити собі створення ресурсу з нуля і тому застосовує поширені CMS для розробки типового сайту. Це відмінне рішення для тих, хто потребує ресурсі зі стандартним набором функцій, будь то візитка або інтернет-магазин. CMS дозволяє:

- наповнювати сайт контентом, змінювати і адмініструвати ресурс, при цьому не будучи ІТ-фахівцем і не маючи серйозних навичок програмування;
- створювати нові сторінки в короткі терміни без зайвих витрат;
- оптимізувати зовнішній вигляд сайту і покращувати якість його наповнення.

Від CMS багато в чому залежать функціональність ресурсу, його можливості і зручність для користувача. Правильно обрана система дозволить успішно створити і розкрутити сайт, зробивши його привабливим для клієнта, надійним і працюючим рівно так, як потрібно.

Так як в CMS системах зберігається велика кількість даних, до багатьох з яких доступ повинен бути обмежено, то стає питання безпеки. Для цього використовуються системи контролю доступу [27].

Одне з найкращих застосувань систем розпізнавання осіб на даний момент саме в системах контролю доступу. По-перше співробітник сам зацікавлений у наданні йому доступу і не буде свідомо саботувати роботу системи розпізнавання осіб. По-друге ви контролюєте, все зовнішні чинники впливають на якість розпізнавання – освітлення, фон, схема руху співробітників. Використовую все це ви можете створити ідеальні умови.

Системи розпізнавання осіб можуть використовуватися в системах контролю доступу в двох режимах – ідентифікації та верифікації.

Режим ідентифікації – рішення про допуск приймається на основі тільки даних від системи розпізнавання осіб. Тобто, наприклад база даних з ваших співробітників 100 осіб, і завдання системи розпізнавання порівняти особа поточного людини з базою даних в 100 чоловік. Тобто порівняння відбувається 100:1. Якщо людина буде ідентифікований як співробітник, то йому буде надано доступ. Даний режим, найефективніше використовувати в задачах виявлення сторонніх на контрольованій території. Як правило є сенс використовувати в особливо охоронюваних зонах підприємства, куди доступ дозволено обмеженому колу осіб. До системи розпізнавання підключаються всі камери встановлені на даній території в разі виявлення будь-якої особи, яка не міститься в базі даних відбувається інформування служби безпеки.

Режим верифікації – ідентифікація в даному випадку проводиться за допомогою іншої технології, наприклад RFID, або може використовуватися мобільні ідентифікатори [28].

Людина підносить карту до зчитувача система його ідентифікує, тобто встановлює що це Іванов, і Іванову дозволений доступ в даний час. Система розпізнавання осіб в даному випадку вже знає, що це Іванов, і використовуючи тільки фото Іванова з бази даних порівнює, пред'явника RFID карти з фотографією Іванова в базі даних. Тобто порівняння відбувається 1:1.

У даній роботі буде розглянуто саме режим ідентифікації.

1.4 Постановка задачі дослідження

Таким чином, ідентифікація облич є актуальною задачею. Тому ставиться завдання розробки методу розпізнавання облич CMS системах.

Об'єктом дослідження є послідовність різноракурсних зображень облич людей.

Метою даного дослідження є огляд та реалізація методу згорткової нейронної мережі для ідентифікації облич для проходження реєстрації в CMS системах.

Для цього необхідно вирішити такі завдання:

- провести аналіз існуючих методів розпізнавання облич;
- скласти математичну модель обраного методу;
- реалізувати комп'ютерну модель обраного методу;
- провести тестування розробленої програми.

2 МАТЕМАТИЧНА МОДЕЛЬ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ

2.1 Визначення згорткової нейронної мережі

Для розпізнавання облич людей широко використовується архітектура нейронної мережі, яка отримала назву згорткової нейронної мережі (ЗНМ). Дана нейронна мережа була вперше описана в і зараз успішно застосовується для вирішення широкого класу задач пов'язаних з розпізнаванням патернів, таких як розпізнавання тривимірних об'єктів, передбачення погоди, автоматичне керування тощо [29].

2.2 Конструкція згорткової нейронної мережі

ЗНМ є суттєвим з різних видів шарів: згорткові, агрегувальні шари та шари «звичайної» нейронної мережі – перцептрона, відповідно до рисунка 2.1.

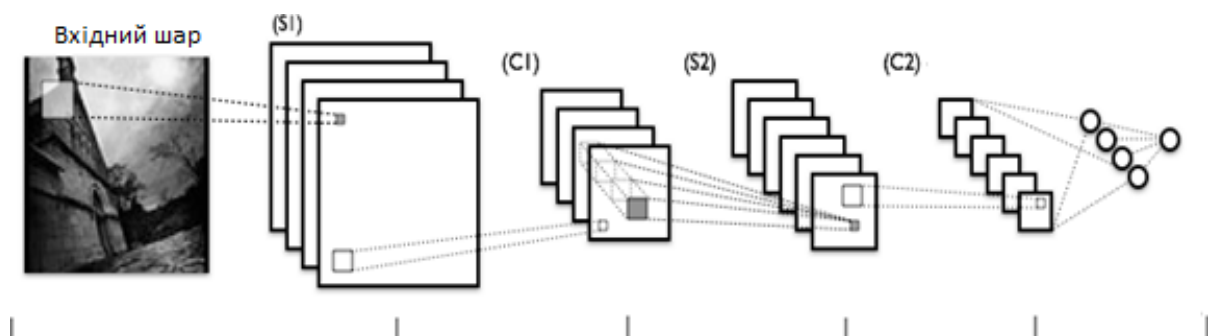


Рисунок 2.1 – Топологія згорткової нейронної мережі

Перші два типи шарів (згорткові, агрегувальні), чергуються між собою, формують вхідну векторну ознаку для багатослівного перцептрона.

Свою назву згорткові мережі отримали по назві операції – згортання, суть якої буде описана далі.

Згорткові мережі є середнім між біологічними мережами та звичайним багатословним перцептроном. На сьогоднішній день найкращі результати в

розпорядженні зображень отримують з їх допомогою. У середній точності розподілу таких мереж переходить звичайні НМ на 10-15%. ЗНМ – це ключова технологія Deep Learning [30].

Основною причиною успіху ЗНМ стала концепція інших ваг. Не дивлячись на великий розмір, ці мережі мають невелику кількість змінних параметрів. Існують варіанти TCNN (Tiled Convolutional Neural Network), аналогічно неокогнітрону, у таких мережах відбуваються, частковий показ від пов'язаних вагів, але алгоритм навчання залишається тим же і заснований на основному розповсюдженні помилок. ЗНМ може швидко працювати на наступній машині та швидко навчатись за рахунок чистого оптимізації процесів згортання на кожній карті, а також зворотніх згорток при розповсюдженні помилок через мережі (рис. 2.2).

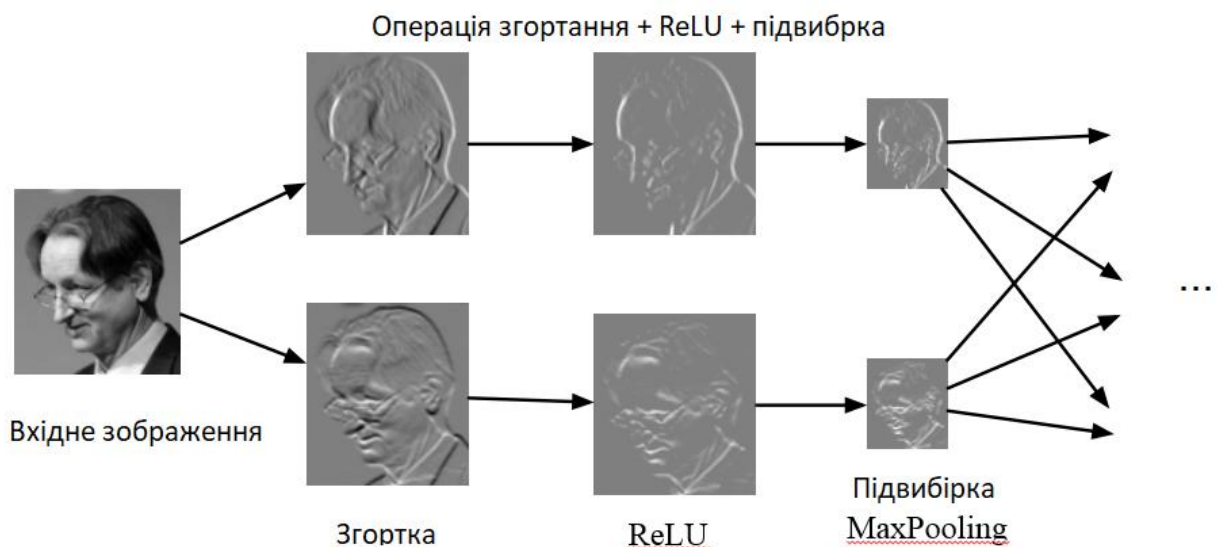


Рисунок 2.2 – Операція згортки

Визначення топології мереж орієнтується на найменшу задачу, дані з наукових статей та власний експериментальний досвід [31].

Можливо видати наступні етапи впливу на вибір топології:

– визначити вирішальну задачу нейромережі (класифікація, прогнозування, модифікація);

- визначити обмеження у вирішальному завданні;
- визначити вхідні та вихідні дані (кількість класів).

Завдання, що вирішує – класифікація зображень, конкретно осіб. Накладаються обмеження на мережу – це швидкість відтворення – не більше 1 секунди і точність розподілу не менше 70%. Загальна топологічна мережа відповідно до рисунка 2.3.

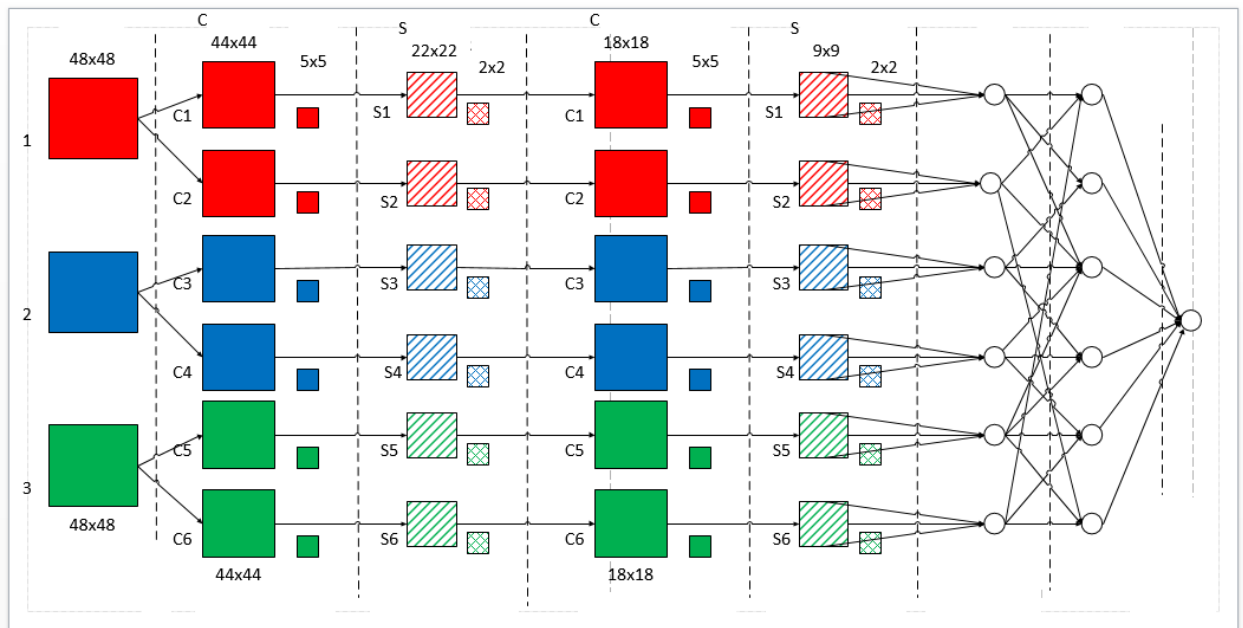


Рисунок 2.3 – Повна схема згорткової нейронної мережі

2.2.1 Вихідний шар

Попередні дані представляють із себе кольорові зображення JPEG розміром 48×48 пікселів. Якщо розмір буде занадто великим, то обчислювальна складність підвищується, відповідно обмеження на швидкість відправи будуть порушені, визначення розміру в даному завданні вирішується методом підбору. Якщо вибрати розмір занадто маленького, щоб не побачити ключових призначених осіб. Кожне зображення розбивається на 3 канали: червоний, синій, зелений. Таким чином отримується 3 зображення розміром 48×48 пікселів.

Вхідний шар виконує двосторонню топологію зображень і знаходиться на кількох картках (матрицях), карта може бути одна, у тому випадку, якщо зображення, представлене у відтинках сірого, інакше їх 3, де кожна карта відповідає зображенню з конкретним каналом (красним, синім та зеленим) .

Вхідні дані кожного конкретного значення пікселя нормалізуються в діапазоні від 0 до 1 за формулою:

$$f(p, min, max) = \frac{p - min}{max - min}, \quad (2.1)$$

де f – функція нормалізації;

p – значення конкретного кольору пікселя;

min – мінімальне значення 0, max – максимальне значення 255.

2.2.2 Згортковий шар

Згортковий шар представлений із набору карт (інше назва – карти ознак, в звичайних матрицях), у кожній карті є синаптичне ядро (в різних джерелах його називають по-різному: скануюче ядро або фільтр) [32].

Кількість карт визначає вимоги до завдань, якщо взяти велику кількість карт, щоб підвищити якість розподілу, а також збільшити обчислювальна складність. Що відбувається з аналізу наукових статей, у більшості випадків пропонується брати співвідношення один до двома, якщо є карта попереднього слова (наприклад, у першому згортковому шарі, що передуює вхідному), зв'язане з двома картами згорткового слова, відповідно до рисунка 2.4.

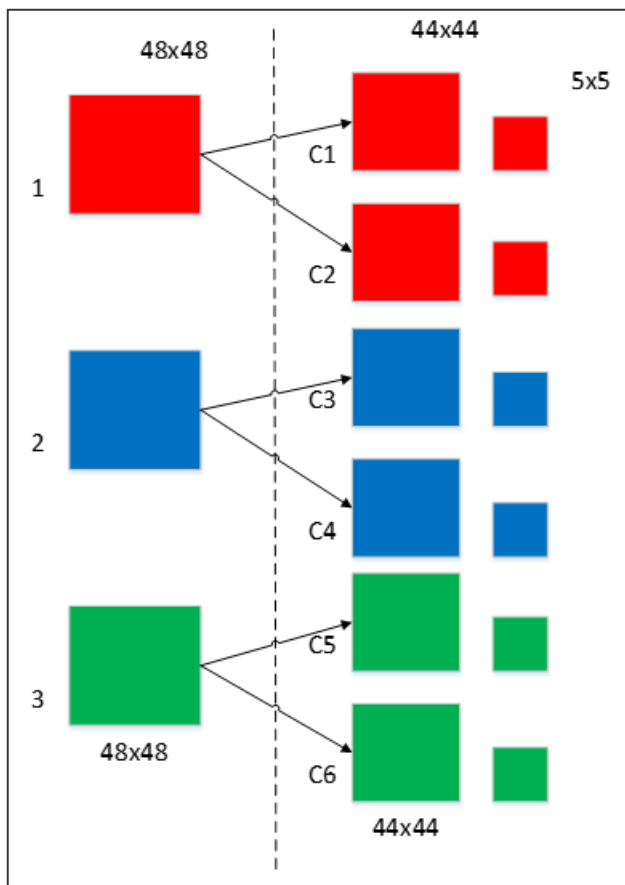


Рисунок 2.4 – Організація зв'язків між картами згорнутого і попереднього шарів

Розмір у всіх картках згорнутого шару – однакові та обчислюються:

$$(w, h) = (mW - kW + 1, mH - kH + 1), \quad (2.2)$$

де (w, h) – розмір згорткової карти;

mW – ширина попередньої карти;

mH – висота попередньої карти;

kW – ширина ядра;

kH – висота ядра.

Ядро це фільтр або вікно, яке проходить по всій області попередньої карти та знаходить визначені ознаки об'єктів. Наприклад, якщо ви навчаєтеся на багатьох особах, то одно з ядерних можливостей у процесі навчання видасте найбільший сигнал в області очей, рота, брів або носа, а інше ядро могло

виявити інші ознаки. Розмір ядра зазвичай становить у межах від 3×3 до 7×7 . Якщо розмір ядра маленький, то він не зможе видати які-небудь признаки, якщо занадто велике, збільшить кількість зв'язків між нейронами. Також розмір ядра вибирається таким чином, щоб розмір картки верхнього шару був одним, що дозволяє не передавати інформацію при зменшенні розміру в підвиборчому шару, описаному нижче.

Ядро представляє систему, розділену вагами. Це одна з головних особливостей згорткової нейромережі. У звичайній багатошаровій мережі дуже багато зв'язків між нейронами, тому є синапси, що в цілому замінює процес виявлення. У верхній мережі – навпаки, загальна інформація дозволяє скоротити число зв'язків та дозволити знайти один і той же признак по всій області зображення [33].

Значущі значення кожної карти згорткового слова дорівнює 0. Значення вагів ядер задаються випадковим чином в областях від -0,5 до 0,5. Ядро проходить за попередньою картою та виробляє операцію згортки, яка часто використовується для обробки зображень:

$$(f * g)[m, n] = \sum_{k, l} f[m - k, n - l] * g[k, l], \quad (2.3)$$

де f – вихідна матриця зображення;

g – ядро згортки.

У виправленому вигляді цього шару можна описати формулу:

$$x^l = f(x^{l-1} * k^l + b^l), \quad (2.4)$$

де x^l – вихід шару l ;

$f()$ – функція активації;

b^l – коефіцієнт зсуву шару l ;

$*$ – операція згортки входу x з рядом k .

При цьому за рахунок краєвих ефектів розмір вихідних матриць зменшується, формула:

$$x_j^l = f(\sum_i x_i^l * k_j^l + b_j^l), \quad (2.5)$$

де x_j^l – карта ознак j (вихід шару l);

$f()$ – функція активації;

b^l – коефіцієнт зсуву шару l для карти ознак;

k_j^l – ядро згортки j карти, шару l ;

* – операція згортки входу x з ядром k .

2.2.3 Підвибірковий шар

Підвибірковий шар також має карти, але їх кількість відповідає попередньому шару, їх 6. Мета шару – зменшення розміру карт попереднього слова. Якщо за попередньою операцією звернення вже були виявлені деякі визнання, для подальшої обробки настільного детального зображення вже не потрібно, а оновлене до менш детального. До того часу фільтрація вже неоднорідних деталей допомагає не перенавчання.

У процесі сканування ядром підвибіркового шару (фільтром) карт попереднього шару, скануюче ядро не перетинається зі згортковим шаром. Як правило, кожна карта має розмір 2×2 , що дозволяє зменшити попередні карти згорткового слова в 2 рази. Кожна карта ознак розділена на чарунки 2×2 елемента, з яких вибираються максимальні значення.

Зазвичай в підвибірковому шарі застосовується функція активації ReLU. Операція підвиборки (або MaxPooling – вибір максимального) відповідно до рисунка 2.5.

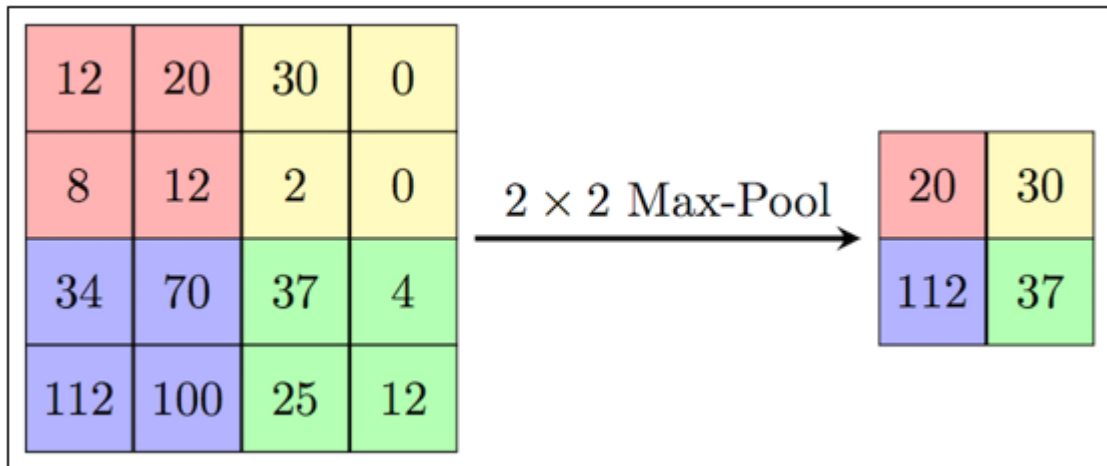


Рисунок 2.5 – Формування нової карти підвибіркового шару на основі попередньої карти згорткового шару

Формальний шар може бути описаний формулою:

$$x^l = f(a^l * \text{subsample}(x^{l-1}) + b^l), \quad (2.6)$$

де x^l – вихід шару;

$f()$ – функція активації;

$a^l b^l$ – коефіцієнти зсуву шару l ;

$\text{subsample}()$ – операція вибірки локальних максимальних значень.

2.2.4 Повнозв'язний шар

Останній з типових шарів це шар звичайного багатошарового персептрона. Мета шару – класифікація, що моделює складну нелінійну функцію, оптимізує та покращує якість розподілу (рис. 2.6) [34].

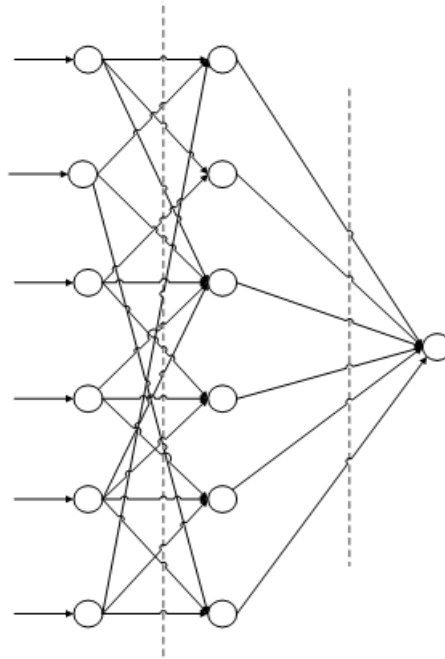


Рисунок 2.6 – Шари нейронної мережі

Нейрони кожної карти попереднього підвибіркового шару пов'язані з одним нейроном прихованого слова. Таким чином, число нейронів прихованого слова рівно числу карт підвибіркового слова, але зв'язок може бути не обов'язково таким, наприклад, лише частина нейронів як-небудь із картки підвибіркового слова пов'язана з першим нейронним скритим словом, а встановлена частина з боку, будь-які нейрони першої карти пов'язані з нейронами 1 і 2 прихованого слова. Вирахування знайомого нейрона можливо описати формулою:

$$x_j^l = f(\sum_i x_i^{l-1} * w_{i,j}^{l-1} + b_j^{l-1}), \quad (2.7)$$

де x_j^l – карта ознак j (вихід шару l);

$f()$ – функція активації;

b^l – коефіцієнт зсуву шару l ;

$w_{i,j}^l$ – матриця вагових коефіцієнтів шару l .

2.2.5 Вихідний шар

Вихідний шар пов'язаний з усіма нейронами попереднього слова. Кількість нейронів відповідає кількості розподілених класів, тобто 2 – лице і не лице. Для зменшення кількості зв'язків і обчислення для бінарного випадку можна використовувати один нейрон і при використанні в якості функцій активації гіперболічних тангенсів, вихід нейрона з значенням -1 означає належність до класу «не обличчя», попередній вихід нейрона за значенням 1 – означає відповідність класу обличчя.

2.3 Архітектура класичної згорткової нейронної мережі

Розглянемо архітектуру звичайної згорнутої нейронної мережі (рис. 2.7).

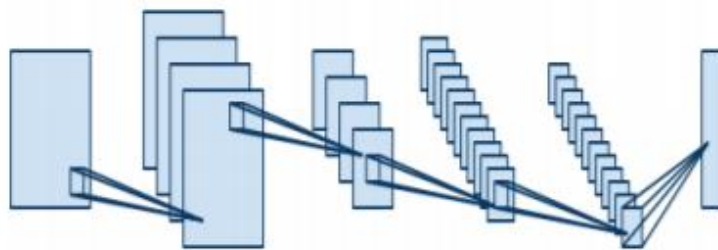


Рисунок 2.7 – Структура ЗНМ

Нейронна мережа складається з пар шарів – шарів підвибірки і шарів згортки, кожен з яких в свою чергу складається з карт ознак.

Неважко переконатися в тому, що кожна карта ознак в ідеалі фільтрує зображення, знаходячи якийсь один певний, специфічний для даної карти, ознака. наприклад, перша карта ознак навчена шукати кружечки, друга – квадратики і т.д.

Початкове зображення подається на вхідний шар. У першому шарі підвибірки кожна карта ознак здійснює пошук певного, закріпленого тільки за

даною карткою, ознаки. Досягається це за рахунок використання загальної для всієї карти ознак матриці ваг і особливою організацією локального рецептивного поля для кожного нейрона такої карти. Кожен нейрон карти ознак отримує вхідні дані від прямокутної області розміру $n \times t$ вхідного зображення. Така область досить мала і багато таких областей на вхідному зображенні перетинаються і накладаються за принципом черепиці.

Суміжні нейрони карти ознак отримують в якості вхідного впливу суміжні прямокутні області, причому вагові коефіцієнти для всіх нейронів карти ознак будуть однаковими.

Для простоти викладу будемо називати область, яка формує локальне рецептивне поле нейрона шару підвибірки, вікном. Відповідно, площа вікна – кількістю нейронів в такій області.

Таким чином, можна говорити про те, що карта ознак в цілому здійснює операцію пошуку ознаки у вхідних даних. Інші карти ознак мають інший набір вагових коефіцієнтів і, відповідно, здійснюють пошук інших ознак у вхідних даних [35].

Конкретні ознаки, які добуваються тієї або іншої картою ознак, визначаються в процесі навчання нейронної мережі з учителем.

N, t – досить малі числа, які визначають роздільну здатність нейронної мережі – мінімальний розмір ознаки, що ця мережа може реєструвати.

Для введення інваріантності нейронної мережі до зсувів і невеликим деформацій, використовується шар згортки. Для кожної карти ознак існує відповідна їй карта згортки, яка зменшує розмірність карти ознак з $n \times t$ до $n / 2 \times t / 2$ шляхом усереднення значень по квадрату 2×2 нейронів.

Після виконання згортки мережу втрачає частину інформації про точне положенні знайденого ознаки, але зберігає інформацію щодо взаємного розташування різних ознак.

Наступний шар підвибірки здійснює аналогічну першому шару сегментацію вхідних даних на прямокутні області $n \times t$, тільки вхідними даними другого шару служить вихід першого шару. Кожна карта ознак

другого шару здійснює пошук ознак другого порядку одночасно у всіх картах ознак першого шару.

Очевидно, що з ростом кількості шарів зменшується розмірність кожної карти ознак, хоча, в цілому, кількість нейронів в шарі сильно зростає за рахунок використання більшої кількості карт ознак у верхніх шарах мережі.

Згорткової нейронної мережі з трьома парами шарів підвибірки-згортки цілком достатньо для точного розпізнавання облич людей.

Така нейронна мережа добре себе зарекомендувала в задачах розпізнавання, але її використання в деяких випадках досить проблематично.

Однією з проблем класичної згорткової нейронної мережі є підбір оптимального значення розміру локального рецептивного поля (вікна) нейрона в шарі підвибірки. Малі значення n і m дозволяють підвищити роздільну здатність мережі і дають можливість знаходити досить малі ознаки, але в той же час аналогічний ознака більшого масштабу буде пропущений і прийнятий за сукупність інших ознак. Таким чином, класична згорткова нейронна мережа погано працює з зображеннями, на яких можуть бути присутні однакові ознаки різного масштабу (наприклад, на ненормалізованих по масштабу зображеннях)[36].

2.4 Архітектура узагальненої згорткової нейронної мережі

Припустимо, на зображенні є як маленькі кружечки, так і великі, і необхідно знайти максимум кружечків в першому ж шарі підвибірки. Змінюючи площу локального рецептивного поля кожного нейрона в кожній карті ознак можна домогтися знаходження кружечків різного розміру за один раз, але тоді шматочки інших розмірів залишаться непоміченими.

У першому шарі підвибірки зробимо кілька карт ознак з одним розміром рецептивного поля, кілька – з трохи більшим розміром поля, кілька – з ще більшим.

Для вирішення цієї проблеми пропонується використовувати в одному шарі карти ознак хоча і одного розміру, але з різними розмірами вікна для різних карт ознак.

Наприклад, замість 10 карт ознак з розміром вікна кожного нейрона 3×3 пропонується використовувати 5 карт ознак з розміром поля 3×3 , 3 карти з полем 5×5 і 2 карти ознак з розміром вікна 7×7 .

Така конфігурація дозволить знаходити ознаки різного розміру одночасно відразу в першому шарі підвибірки, що має підвищити загальну якість розпізнавання об'єктів, прискорити навчання нейронної мережі і сприятиме зменшенню кількості карт ознак в вищих шарах нейронної мережі, що призведе до зменшення кількості зв'язків нейронної мережі, і як наслідок – зменшення споживання пам'яті і прискорення навчання [37].

При такій організації першого шару підвибірки закономірно виникає проблема з перетином вікон суміжних нейронів однієї карти ознак – для отримання однакового розміру карт ознак з різними розмірами вікон необхідно розмістити однакову кількість таких полів на вхідному шарі, що веде до збільшення площі перетину суміжних полів з ростом розміру поля.

На жаль, боротися з таким різким збільшенням кількості зв'язків вкрай проблематично, але наступне поліпшення дозволить зменшити загальну кількість зв'язків в згортальних нейронних мережах, що особливо актуально для описаної в даній статті архітектури згорткової нейронної мережі.

Розглянемо докладно один нейрон з карти ознак першого шару підвибірки. такий нейрон отримує інформацію з прямокутної області вхідного зображення, яка формує його локальне рецептивної полі (вікно). Загальна кількість вхідних зв'язків для нейрона одно $n \times m$, де $(n; m)$ – розмір вікна. Зі збільшенням n і m пропорційно зростає кількість зв'язків, але в той же час внесок, внесений одним зв'язком в сумарний вхід нейрона зменшується.

Це означає, що при великій кількості зв'язків можна без значної шкоди для якості роботи мережі переходити до розрідженій формі сполук, видаляючи деяку малу їх частину випадковим чином. Наприклад, для вікна 5×5 нейронів

можна видалити 5 з'єднань, незначно зменшивши точність, але при цьому на 20% скоротивши загальна кількість з'єднань.

На відміну від сполук, загальна кількість ваг за рахунок використання методики поділюваних ваг, зростає незначно і немає сенсу їх скорочувати.

Для даної карти ознак, застосовуючи розрідження матриці зв'язків вхідного шару з картою ознак, можна отримати значне скорочення зв'язків, зменшивши, тим самим, витрата пам'яті і прискоривши процес навчання мережі.

Слід зазначити, що в разі розрідження матриці зв'язків кожен нейрон карти ознак отримує індивідуальну, відмінну від сусідніх нейронів в даній карті ознак, матрицю зв'язків [38].

Застосування вищевикладених рекомендацій щодо поліпшення архітектури згорткової нейронної мережі ставить задачу оптимального розподілу карт ознак з різним розміром поля, а також завдання вибору кількості відсікаються зв'язків в разі розрідження матриць зв'язків. На жаль, кількісні параметри згорткової нейронної мережі повинні підбиратися суворо під поставлену задачу після проведення ряду експериментів для доказу допустимості тієї чи іншої конфігурації. Слід також зазначити, що наведені вище викладки, незважаючи на свою орієнтованість на задачу розпізнавання образів, можуть бути цілком використані для побудови згортальних нейронних мереж, які вирішують інші подібні завдання, до числа яких відносяться прогнозування і управління.

2.5 Багатозадачна каскадна згорткова нейронна мережа

Багатозадачні каскадні згорткові мережі (MTCNN) – це структура, розроблена як рішення для виявлення обличчя та вирівнювання обличчя. Процес складається з трьох етапів згорткових мереж, які здатні розпізнавати обличчя та орієнтири, такі як очі, ніс та рот.

У статті пропонується MTCNN як спосіб інтеграції обох завдань (розпізнавання та вирівнювання) за допомогою багатозадачного навчання. На першому етапі він використовує неглибоку CNN для швидкого створення вікон-кандидатів. На другому етапі він вдосконалює запропоновані вікна-кандидати через більш складний CNN. І нарешті, на третьому етапі він використовує третій CNN, більш складний, ніж інші, для подальшого уточнення результату та виведення позицій на орієнтирі обличчя.

Розглянемо три етапи MTCNN.

Перший крок – зробити зображення та змінити його розмір до різних масштабів, щоб побудувати піраміду зображення, яка є вхідною інформацією наступної триступеневої каскадної мережі (рис. 2.8).

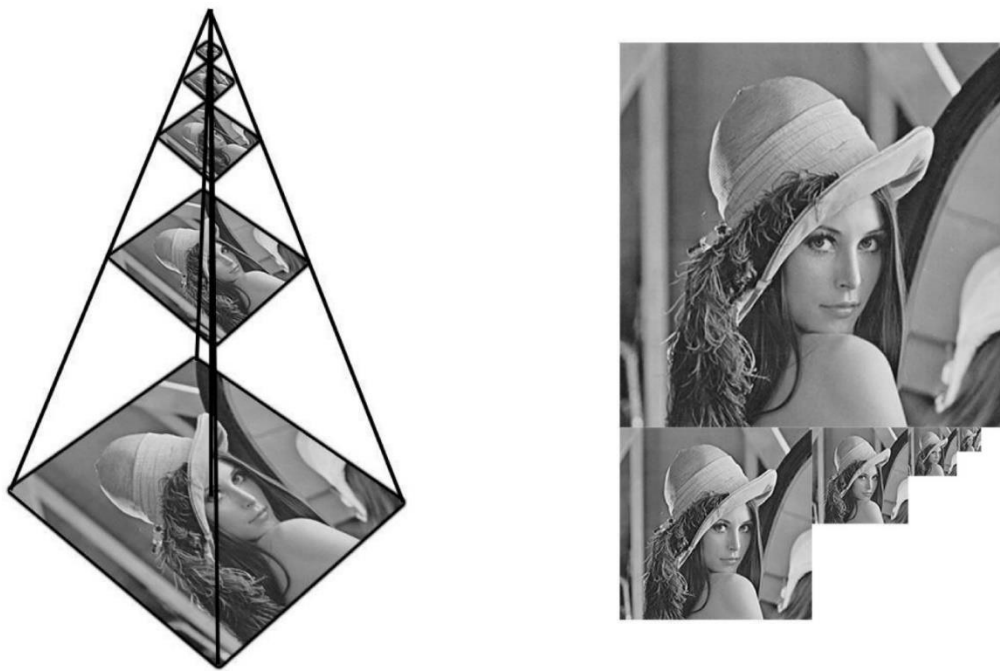


Рисунок 2.8 – Приклад згортання зображення

Етап 1: Мережа пропозицій (P-Net)

Цей перший етап являє собою повністю згорнуту мережу (FCN). Різниця між CNN та FCN полягає в тому, що повністю згорнута мережа не використовує щільний шар як частину архітектури. Ця мережа пропозицій

використовується для отримання вікон-кандидатів та їх регресійних векторів (рис. 2.9).

Регресія обмежувальних ящиків – це популярна техніка для прогнозування локалізації ящиків, коли метою є виявлення об'єкта якогось заздалегідь визначеного класу, в даному випадку обличчя. Після отримання векторів обмежувальної коробки проводиться деяке вдосконалення, щоб поєднати перекриваються області. Кінцевим результатом цього етапу є всі вікна кандидатів після уточнення, щоб зменшити кількість кандидатів [38].

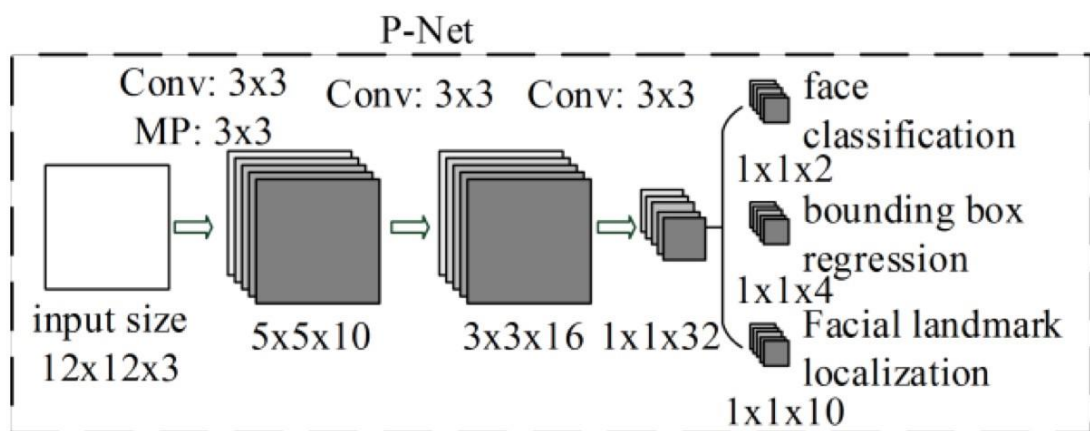


Рисунок 2.9 – Мережа пропозиції

Етап 2: Уточнення мережі (R-Net)

Усі кандидати від мережі P-Net потрапляють до мережі Refine. Зверніть увагу, що ця мережа є CNN, а не FCN, як та, що була раніше, оскільки на останньому етапі архітектури мережі є щільний шар. R-Net додатково зменшує кількість кандидатів, виконує калібрування з регресією обмежувальної рамки та застосовує не максимальне придушення (NMS) для об'єднання накладених кандидатів (рис. 2.10).

R-Net виводить, чи є вводом грань чи ні, 4-елементний вектор, який є обмежувальним полем для обличчя, і 10-елементний вектор для локалізації особового орієнтира.

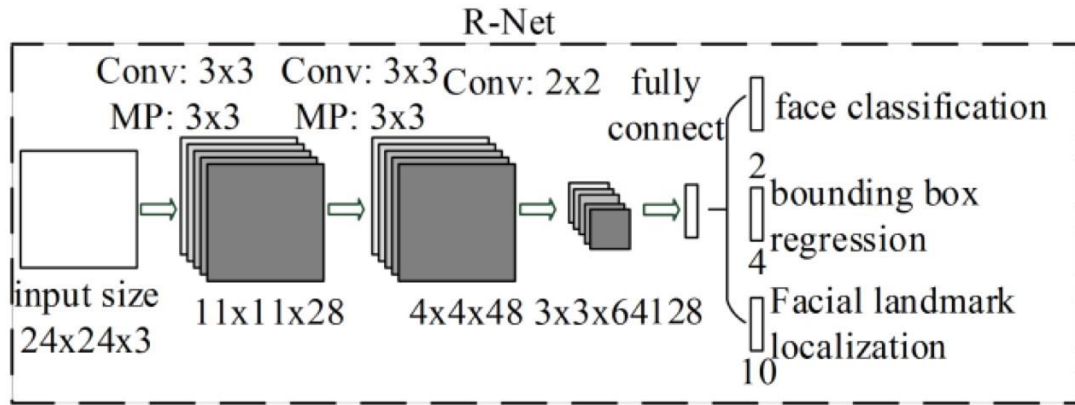


Рисунок 2.10 – Уточнення мережі

Етап 3: Вихідна мережа (O-Net)

Цей етап схожий на R-Net, але ця Мережа вихідних даних має на меті більш детально описати обличчя та вивести п'ять позицій на обличчі для очей, носа та рота (рис. 2.11).

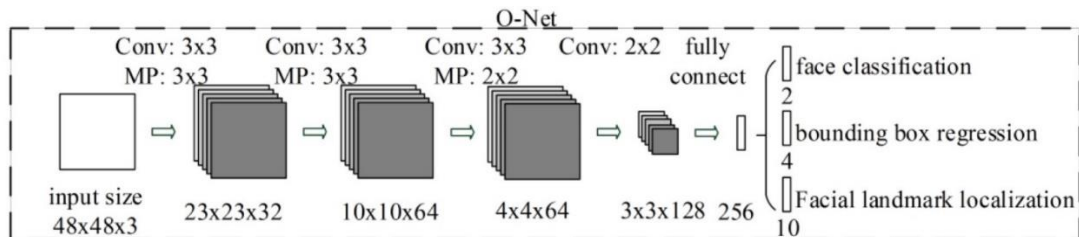


Рисунок 2.11 – Вихідна мережа

Завдання Мережі полягає у виведенні трьох речей: класифікація обличчя, нерегресія, регресія обмежувального вікна та локалізація орієнтиру обличчя [39].

2.6 Детектор MultiBox

Розглянемо детектор MultiBox SSD, який застосовується для розпізнавання облич у багатозадачній згортковій НМ.

Щоб краще зрозуміти SSD, для початку розглянемо, звідки походить назва цієї архітектури:

- Single Shot – означає, що завдання локалізації та класифікації об'єктів виконуються за один прямий прохід мережі;
- MultiBox – назва методики регресії обмежувального контуру;
- детектор – виявляє та класифікує ці виявлені об'єкти.

Як видно, з рисунку 2.12, архітектура SSD базується на архітектурі, що має назву VGG-16, але відкидає повністю зв'язані шари. Причиною того, що VGG-16 був використаний як базова мережа, є його висока ефективність у виконанні висококласних завдань класифікації зображень та його популярність серед проблем, де навчання за допомогою трансферу допомагає поліпшити результати.

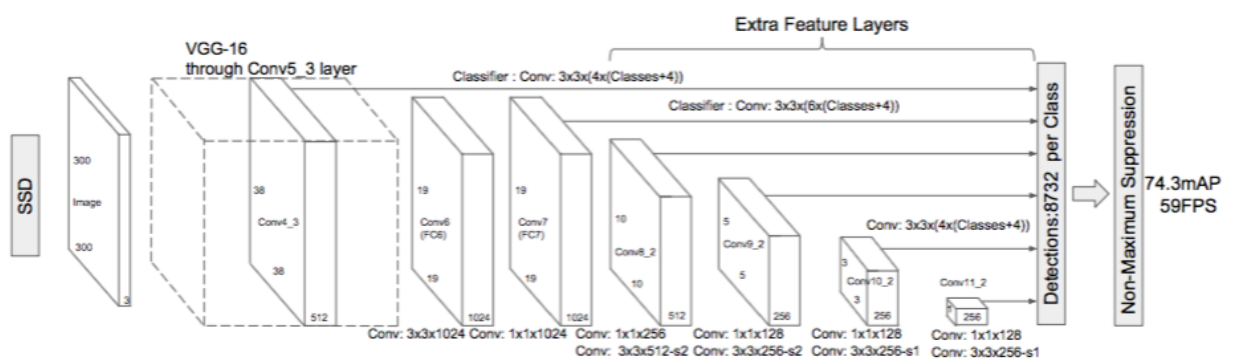


Рисунок 2.12 – Архітектура MultiBox SSD

Замість оригінальних повністю зв'язаних шарів VGG (рис. 2.13) було додано набір допоміжних згорткових шарів, що дозволяє отримувати функції в декількох масштабах і поступово зменшувати розмір вхідних даних для кожного наступного шару.

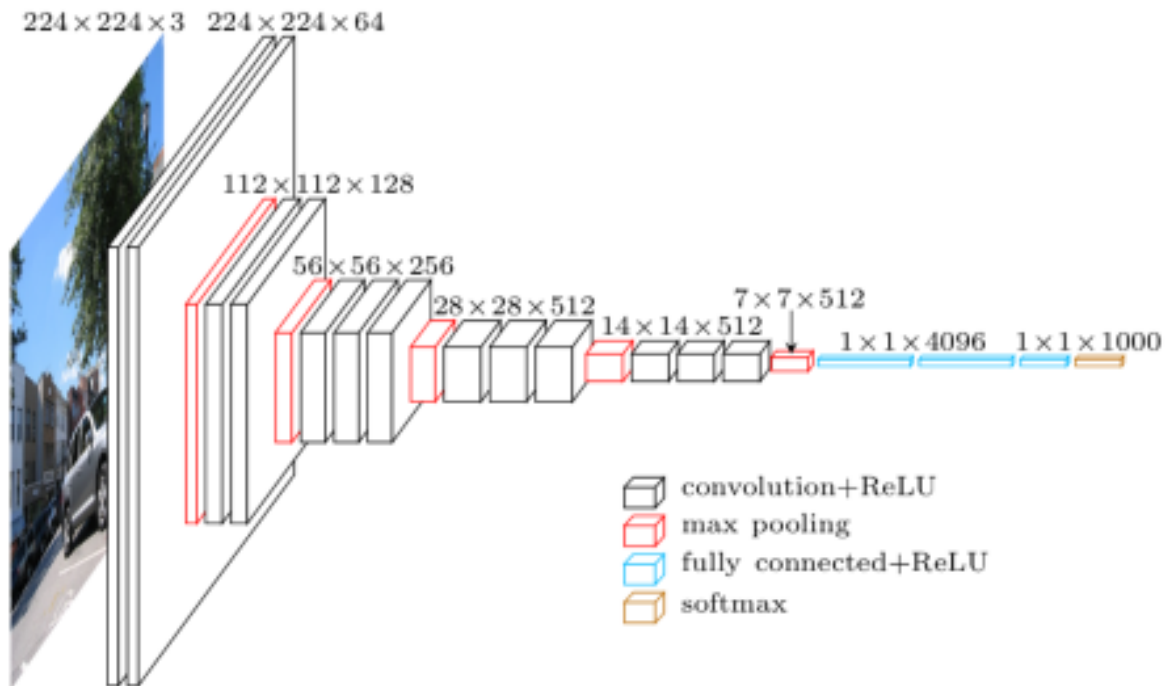


Рисунок 2.13 – Архітектура VGG

Техніка регресії обмежувального контуру твердотільного накопичувача натхнена роботою над MultiBox, методом швидких пропозицій щодо координації обмежувальних контурів, що агностикуються. Цікаво, що в роботі над MultiBox використовується згорткова мережа в стилі Inception. Витки 1×1 допомагають зменшити розмірність, але ширина та висота залишаються незмінними.

Функція втрат MultiBox також поєднує два найважливіші компоненти, які увійшли в SSD:

- втрата впевненості – вимірює, наскільки мережа впевнена в об'єктивності обчислюваного обмежувального поля. Для обчислення цієї втрати використовується категорична перехресна ентропія;

- втрата місцезнаходження – вимірює відстань передбачуваних обмежувальних контурів мережі від основних істинних до навчального набору. Для цього використовується L2-Norm.

Формула для втрати, яка вимірює, наскільки далеко потрапили дані від нашого прогнозу:

$$multibox_loss = confidence_loss + alpha * location_loss. \quad (2.8)$$

Альфа-термін допомагає збалансувати внесок втрати місцезнаходження. Як зазвичай, при глибокому навчанні, метою є знайти значення параметрів, які найбільш оптимально зменшують функцію втрат, тим самим наближаючи наші прогнози до основної істини.

Логіка, яка обертається навколо генерації обмежувальних контурів, насправді є більш складною, ніж те, що було вказано раніше.

У MultiBox реалізовано те, що називаються пріоритетами (або якорями за термінологією Faster-R-CNN), які є попередньо обчисленими обмежувальними рамками фіксованого розміру, які точно відповідають розподілу оригінальних основних істинних контурів. Насправді ці пріоритети вибрані таким чином, що їх коефіцієнт перетину (він же *IoU*, а іноді його називають індексом Жакарда) перевищує 0,5. Зробимо висновок, що *IoU* зі значенням 0,5 все ще недостатньо добрий (рис. 2.14), але він, однак, забезпечує потужну вихідну точку для алгоритму регресії обмежувального контуру. Це набагато краща стратегія, ніж запуск прогнозів із випадкових координат. Тому MultiBox починається з пріоритетів як прогнозів і намагається регресувати ближче до основи, що обмежує істину (рис. 2.15).

Сама формула для підрахунку індексу має вигляд:

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}, \quad (2.9)$$

де Area of Overlap – область перетину;

Area of Union – область об'єднання.

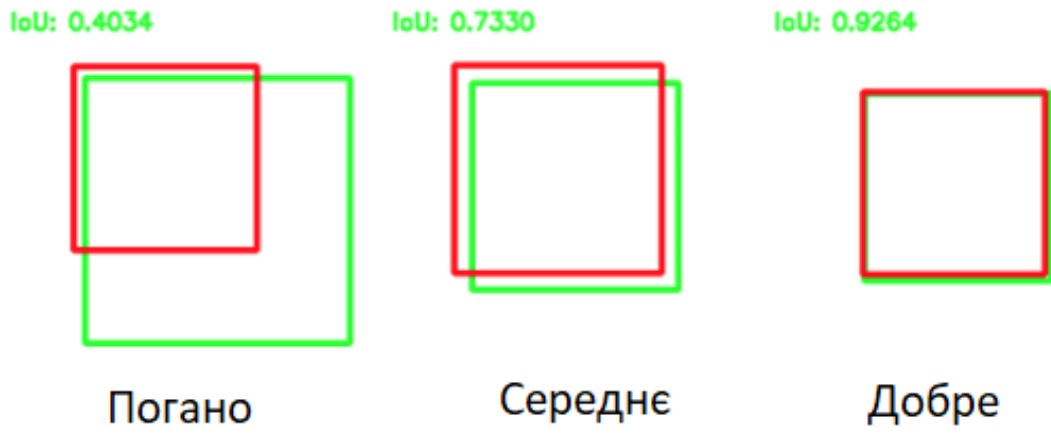


Рисунок 2.14 – Значення індексу та формування контурів



Рисунок 2.15 – Формування обмежувальних контурів для задачі розпізнавання обличчя

Отримана архітектура містить 11 пріоритетів на кожну частину карти об'єктів (8×8 , 6×6 , 4×4 , 3×3 , 2×2) і лише одну на карті об'єктів 1×1 , в результаті отримуючи 1420 пріоритетів на зображення, таким чином забезпечуючи надійне покриття вхідних зображень у різних масштабах, для виявлення об'єктів різного розміру. До того ж, MultiBox зберігає лише найвищі прогнози, які мінімізували як втрати місця, так і впевненість.

3 КОМП'ЮТЕРНА МОДЕЛЬ ФІЛЬТРАЦІЇ ЗОБРАЖЕНЬ

3.1 Обґрунтування вибору середовища програмної реалізації

У рамках атестаційної роботи був розроблений метод ідентифікації обличчя за допомогою багатозадачної згорткової нейронної мережі. Для реалізації були обрані Java Script, Node JS, Tensorflow.

3.1.1 Java Script

Мова сценаріїв – це мова програмування, яка інтерпретується під час виконання, а не компілюється в машинний код. Мова може виконуватися різними способами – інтерпретуватися або компілюватися (в байт-код або машинний код). Традиційно, коли говорять про різницю між сценаріями та програмуванням, мови сценаріїв інтерпретуються та компілюються мови програмування. В основному всі мови сценаріїв є мовами програмування. Теоретична відмінність між ними полягає в тому, що мови сценаріїв не вимагають етапу компіляції, а скоріше їх інтерпретують.

Спектр мов сценаріїв варіюється від дуже малих і дуже специфічних для домену мов до мов програмування загального призначення, що використовуються для сценаріїв. Найбільш базовим інтерфейсом сценаріїв, розробленим на W3C, є DOM, об'єктна модель документа, що дозволяє програмам та сценаріям динамічно отримувати доступ та оновлювати вміст, структуру та стиль документів. Специфікації DOM складають ядро DHTML. Сценарії можуть зробити веб-сторінки більш динамічними. Наприклад, без перезавантаження нової версії сторінки це може дозволити зміни вмісту цієї сторінки або дозволити додавати вміст на цю сторінку або надсилати її.

Мови сценаріїв стають все більш популярними завдяки появі веб-додатків. Хоча сценарії широко використовуються у світі програмування, останнім часом вони стали більш пов'язаними із Всесвітньою павутиною, де

їх широко використовували для створення динамічних веб-сторінок. Хоча технічно існує багато мов сценаріїв на стороні клієнта, які можна використовувати в Інтернеті, на практиці це означає використання JavaScript. Додатки для JavaScript охоплюють від інтерактивних веб-сайтів до Інтернету речей, що робить його чудовим вибором для початківців та досвідчених розробників, які хочуть вивчити нову мову програмування.

Як і всі мови комп'ютера, JavaScript має певні переваги та недоліки.

Розглянемо переваги JavaScript.

Швидкий для кінцевого користувача: сценарій Java, написаний для клієнтської сторони, для підтримки веб-сервера не вимагає підтримки. Він також не потребує компіляції на стороні клієнта, що дає йому певні переваги швидкості. Після сценарію виконується на комп'ютері користувача, залежно від завдань, результати виконуються майже миттєво. Наприклад, ви можете перевірити будь-який користувацький ввід перед відправкою запрошення на сервер. Це знижує навантаження на сервер.

Простота: JavaScript щодо прост у освоєнні та реалізації. Він використовує модель DOM, яка забезпечує безліч попередньо встановлених функцій для різних об'єктів на сторінках, що робить його легким для розробки сценарію для рішення користувацької цілі.

Універсальність: JavaScript відмінно працює з іншими мовами та може використовуватися в самих різних додатках. На даний момент існує безліч способів використання JavaScript через сервери Node.js. Якщо ви завантажите node.js за допомогою Express, використовуйте базу даних документів, використовуйте як mongodb, використовуйте JavaScript в інтерфейсі для клієнтів, ви можете створити додаток JavaScript повністю з одного вікна вперед, використовуючи лише JavaScript.

Недоліки JavaScript:

– безпечність: JavaScript явно додає на веб-сторінку та клієнтський браузер, він може використовувати систему користувача, тому тимчасовий код може бути запущений на клієнтській машині;

- підтримка браузера: JavaScript іногда інтерпретується по-різному різними браузерами. Різні механізми компоновки можуть відображати JavaScript по-різному, що призводить до негласованності з точками зрілості функціональності та інтерфейсу. Більша частина JavaScript залежить від маніпуляції елементами браузерів DOM. І різні браузери пропонують різні типи доступу до об'єктів, зокрема Internet Explorer;
- багато конкурентів: JavaScript – це дуже старий мовний сценарій, що працює на машинах, а також інші технології, які роблять це саме замість нього (наприклад, JQuery) найкращим і легким способом;
- можливість вимкнути JavaScript: якщо ви вимкнете JavaScript у браузері, весь код JavaScript не запущений;
- завантаження файлу: файл JavaScript завантажується на клієнтській машині, щоб кожен міг прочитати код і повторно використовувати його.

3.1.2 Node.JS

У наші дні платформа Node.js є однією з найпопулярніших платформ для побудови ефективних та масштабованих REST API. Вона так само підходить для побудови гібридних мобільних програм, десктопних програм і навіть для IoT.

Node.js представляє середу виконання коду на JavaScript, який побудований на основі движка JavaScript Chrome V8, що дозволяє транслювати виклики на мові JavaScript у машинному коді. Node.js попередньо призначений для створення серверних додатків на мові JavaScript. Хоча також існують проекти за написанням десктопних програм (Electron) і навіть за створенням коду для мікроконтролерів. Попередньо про все ми говоримо про Node.js, як про платформу для створення веб-додатків.

Node.js (або просто Node) – це серверна платформа для роботи з JavaScript через движок V8. JavaScript виконує дію на стороні клієнта, а Вузол

– на серверах. За допомогою Node можна писати повноцінні додатки. Вузол вміє працювати з зовнішніми бібліотеками, визивати команди з коду на JavaScript і виконувати роль веб-сервера.

С Вузол проще масштабуватися. При одночасному підключенні до сервера тисяч користувачів Node працює бездротово, щоб отримати пріоритети та розподілити ресурси грамотне. Java, наприклад, виходить на кожне підключення окремого потоку.

Вузол з'явився в 2009 році завдяки Раяну Далу. До цього в серверах доступний підхід «один потік на кожному з'єднанні», а Дал придумав використовувати систему, яка орієнтована на події. Що реагує на дію або бездіяльність і виділяється під цим ресурсом. Головна мета Node – побудова великомасштабних мережевих серверів.

3.1.3 Tensorflow

Запропонована технологія Google глибинного навчання з відкритим кодом вже успіху завоювати популярність у корпоративному секторі.

Представлений в листопаді 2015 року програмне забезпечення з відкритим кодом TensorFlow, компанія Google об'єднала зацікавлених потенційних користувачів, а на сьогоднішній день ця технологія розміщується на все більшій популярності.

Бібліотека програмного забезпечення для машинного навчання представляє нове покоління внутрішніх технологій DistBelief, яка була розроблена командою Google Brain для багатьох завдань, включаючи пошук зображень та покращення алгоритмів розподілу речей.

TensorFlow – це нейронна мережа, яка виконує відповідні завдання за допомогою позитивного посилення та обробки даних на різних рівнях (вузлах), що допомагає знайти коректний результат.

Відкрийте вихідний код бібліотеки машинного навчання TensorFlow, в Google упростили процес побудови та розвертання важких нейронних мереж. TensorFlow не пропонує кожному розробнику можливостей використовувати плоди машинного навчання, а також не пропонує інтерфейси API для мов Python та C / C ++, дозволяючи підключити програму розробників.

Машинне навчання такого роду призначене виключно для дослідницької мети, завдяки програмному забезпеченню з відкритим кодом наповнення підприємство TensorFlow отримує потужні кошти для використання власних даних та їх обробки в дешевій хмарній середі.

Бібліотеки TensorFlow замінюють вбудовування в додатки самонавчальних елементів та функцій мистецького інтелекту, призначених для розподілу речей, організації комп'ютерного бачення або обробки естетичної мови.

Звісно, TensorFlow не є єдиною бібліотекою глибинного навчання, але, як і пошуковий механізм Google, вона вважає найкращою у своєму класі. Альтернативами є програмне забезпечення Torch, створене швейцарськими дослідниками та розробка Каліфорнійського університету в Берклі Caffe, остання версія якого, Caffe2, спроектована за участю Facebook.

Відповідно до інформації, опублікованої на сайті TensorFlow, бібліотека використовує цілий ряд найбільших компаній, у тому числі Airbnb, Airbus, Dropbox, Snapchat та Uber (правда, можливо, і не найкращий для себе спосіб).

Між тим продовжуються пошуки нових сценаріїв. Одною з компаній, яка не пов'язується з TensorFlow, є британський електронний супермаркет Ocado.

Команда команди дослідників даних застосовує бібліотеку для найрізноманітніших товарів, навчання від алгоритмів маршрутизації роботів, переміщення по складах та завершення покращення прогнозування програм (що базується на алгоритмі лінійної регресії десятилітньої давності) та пропозицію покупців товарів відповідно до їх останніх переваг.

Початок TensorFlow в Ocado використовувався для управління електронною поштою, що працює в контактному центрі. Спочатку всі листи обробляються послідовно, незалежно від їх пріоритетів. Ослаблення возика в плоху погоду, коли кількість листів збільшується в три-чотири рази, а їх доставка переважає в серйозній проблемі

TensorFlow – це вільно розповсюджена програмна бібліотека для додатків з областей машинного навчання. TensorFlow можна застосовувати для створення нейронних мереж та реалізації інших алгоритмів глибинного навчання [40].

Це бібліотека, випущена Google у листопаді 2015 р., написана на Python. Для навчання та оцінки створених моделей у них застосовуються розрахунки на CPU або GPU. Етап бібліотеки створюється для роботи на високопродуктивних серверах із використанням ресурсозатратних графічних процесорів.

Останнє оновлення дозволило оптимізувати цю бібліотеку та використовувати в середах із більш обмеженими ресурсами – наприклад, на мобільних пристроях та у веб-браузерах.

Tensorflow Lite, полегшена версія цієї бібліотеки для мобільних пристроїв та вбудованих системи, випущена в травні 2017 року. Разом з нею пропонується новий набір передбачених глибинних моделей для завдань, пов'язаних з розпізнавання образів; ця колекція називається MobileNet. Моделі MobileNet розроблені спеціально для ефективної роботи в округах з обмеженою кількістю ресурсів, наприклад, на мобільних пристроях.

Слідом за Tensorflow Lite в березні 2018 року була анонсована TensorFlow.js. Ця версія бібліотеки, призначена для роботи в браузері, базується на більш ранніх проектах під назвою deeplearn.js. WebGL забезпечує GPU-доступ до бібліотек. Розробники використовують API на JavaScript для навчання, завантаження та запуску моделей.

Пізніше TensorFlow.js було розширено для роботи з Node.js, для цього застосовується бібліотечне доповнення tfjs-node.

3.2 Вимоги до програми

Алгоритм, який необхідно розробити буде використано для CMS систем у сфері фармацевтики. Тобто основне направлення використання даного алгоритму це B2B та B2C проекти.

Фармацевтичні компанії регулярно збільшують обсяг онлайн-продажів. Для продажу багатьох ліків необхідно мати рецепт або певну кількість років. Тому необхідно проводити ідентифікацію людей, що намагаються купити певні ліки.

Даний алгоритм можливо покращувати, додавати нові функції та згодом інтегрувати до великих CMS систем.

Основними вимогами до такого розширення можуть бути:

- розпізнавання облич;
- ідентифікація людини;
- порівняння людини з фото з документів, таких як паспорт, у режимі реального часу;
- визначення гендеру;
- визначення кількості років.

На даному етапі необхідно реалізувати розпізнавання облич та ідентифікацію людей шляхом порівняння завантаженого фото з базовим зображенням для жінок та чоловіків.

Результати розпізнавання не повинні залежати від ракурсу обличчя людини та світла на фото.

3.3 Програмна реалізація

Можливо виділити наступні етапи розробки програми ідентифікації людей у CMS системах:

1. Формування набору даних для тренування згорткової нейронної мережі:

- а) пошук та обрізка зображень з обличчями людей;
- б) виділення контурів та призначення ярликів.

2. Створення стандартних контурів різних розмірів та з різними відношеннями сторін для підвищення універсальності алгоритму.

3. Формування декількох карт особливостей, за якими буде проходити навчання. Це необхідно, бо якість роботи MultiBox детектора підвищується, якщо він працює на декількох картах.

4. Видалення жорстких негативів. Під час навчання, оскільки більшість обмежувальних контурів матимуть низький рівень IoU , то мережа їх трактуватиме як негативні приклади тренувань, ми можемо мати в результаті непропорційно велику кількість негативних прикладів у нашому навчальному наборі.

5. Обробити різні ракурси. Для цього ми збільшуємо кількість даних шляхом створення нових зображень зі зміною кута відображення.

6. Зменшення кількості обмежувальних контурів. Даний процес відбувається шляхом видалення контурів із порогом втрати достовірності менше встановленого наближення (наприклад, 0,01) і IoU менше, наприклад, 0,45 відкидаються, і зберігаються лише верхні прогнози. Це гарантує, що мережа зберігає лише найбільш вірогідні прогнози, тоді як більш шумні видаляються.

Для навчання згорткової нейронної мережі було створено набір даних, який складається з 1081 файлів. Дані файли є зображеннями формату .jpg. Кожне зображення містить обличчя різних людей великим планом, де чітко можливо визначити базові точки та контури. Приклад зображень, що використовувались для навчання згорткової нейронної мережі зображено на рисунку 3.1. Розміри усіх зображень повинні бути 300×300 пікселів.

< training_real (1081 files)

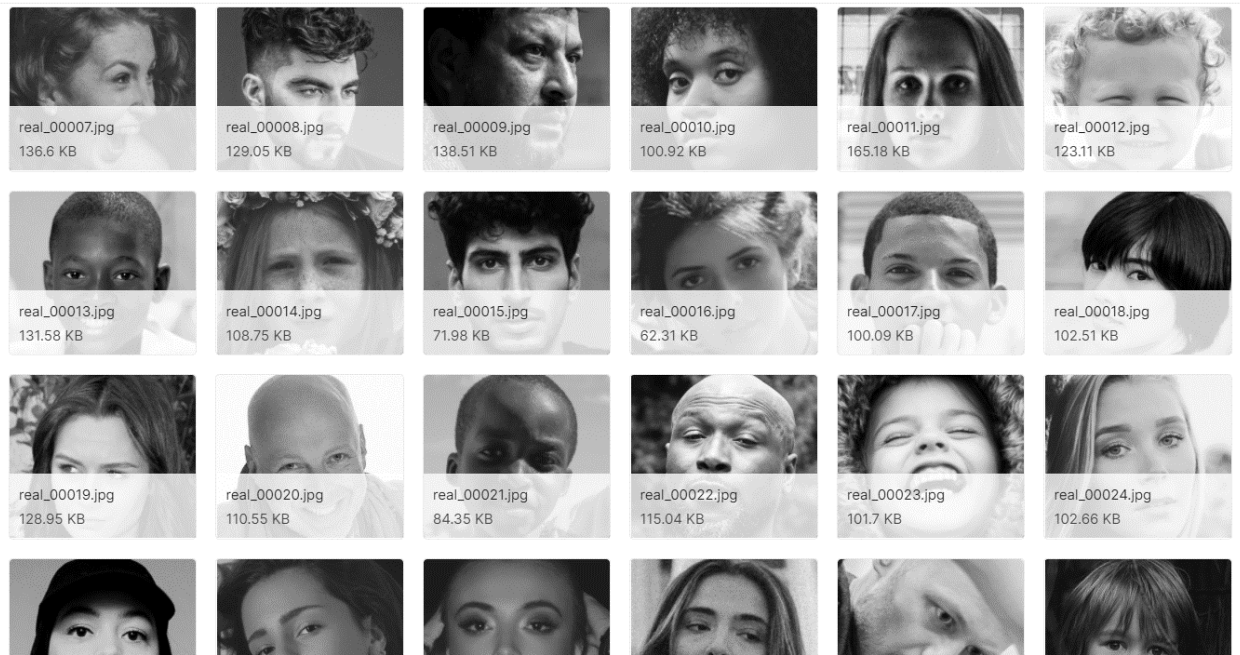


Рисунок 3.1 – Набір даних з обличчями різних людей для тренування нейронної мережі

3.4 Тестування розробленої моделі

Для проведення тестування розробленої системи було додано 2 зображення чоловіка та жінки, з якими надалі повинні порівнюватись завантажені зображення. Дані обличчя людей зображено на рисунках 3.2 і 3.3.



Рисунок 3.2 – Базове зображення жінки



Рисунок 3.3 – Базове зображення чоловіка

Наступним кроком відібрано 4 зображення для тестування схожості облич. Дані обличчя зображені на рисунках 3.4 – 3.7.



Рисунок 3.4 – Зображення жінки з базового зображення з іншого ракурсу для тестування



Рисунок 3.5 – Зображення іншої жінки для тестування

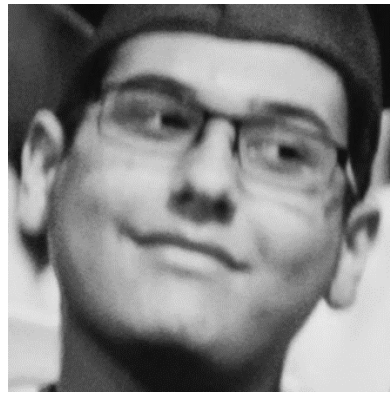


Рисунок 3.6 – Зображення чоловіка, схожого на людину з базового зображення для тестування

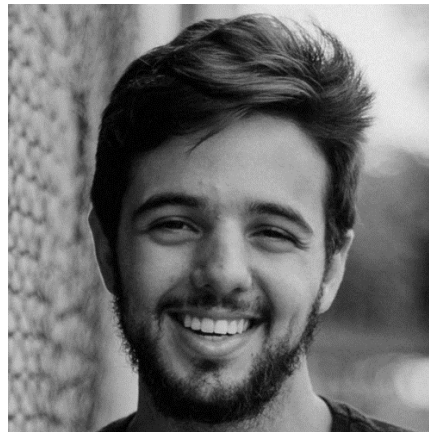


Рисунок 3.7 – Зображення іншого чоловіка для тестування

Після вибору зображень для тестування у систему необхідно завантажити базові зображення. Наразі це 2 фото чоловіка та жінки. На наступних етапах розробки це можуть бути фото з документів людини, яку необхідно буде ідентифікувати.

Проведемо декілька експериментів для тестування розробленого додатку.

В першу чергу порівняємо чоловіка та жінку (рис. 3.8).

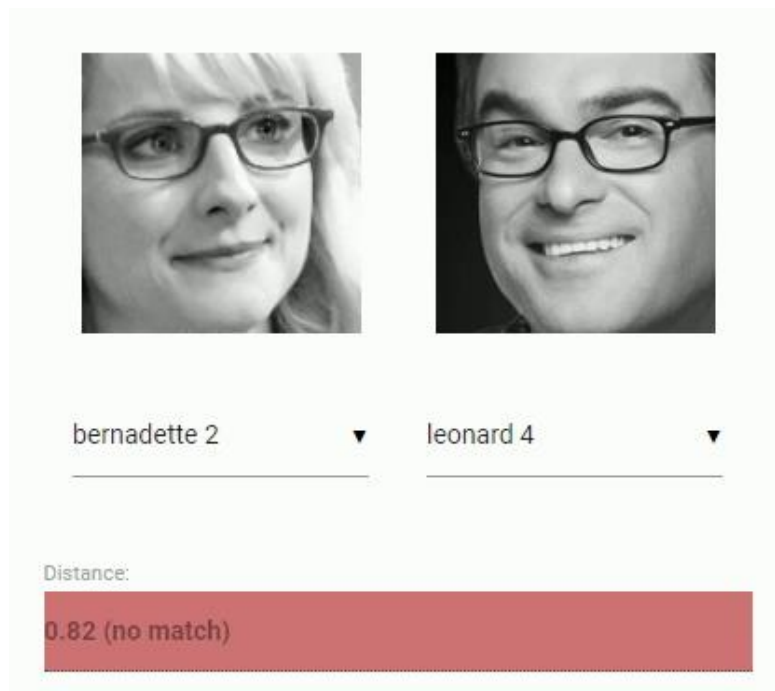


Рисунок 3.8 – Тестування порівняння чоловіка та жінки

Отриманий результат є коректним, адже індекс дорівнює значенню 0,82, тобто на зображення обличчя знайдені та вони не є однією людиною.

Наступний тест полягає у порівнянні двох зображень однієї жінки але з різних ракурсів. Експеримент зображено на рисунку 3.9.



Рисунок 3.9 – Тестування позитивного кейсу з жінкою

Результати є коректними. Алгоритм визначив обличчя однієї людини. Індекс є середнім, а саме 0,42, що вказує, що зображення не ідентичні, тобто ракурс обличчя різний на обох зображеннях.

Наступним кроком порівняно фото різних жінок (рис. 3.10). За результатами обличчя не належать одній людині, адже значення індексу 0,77.

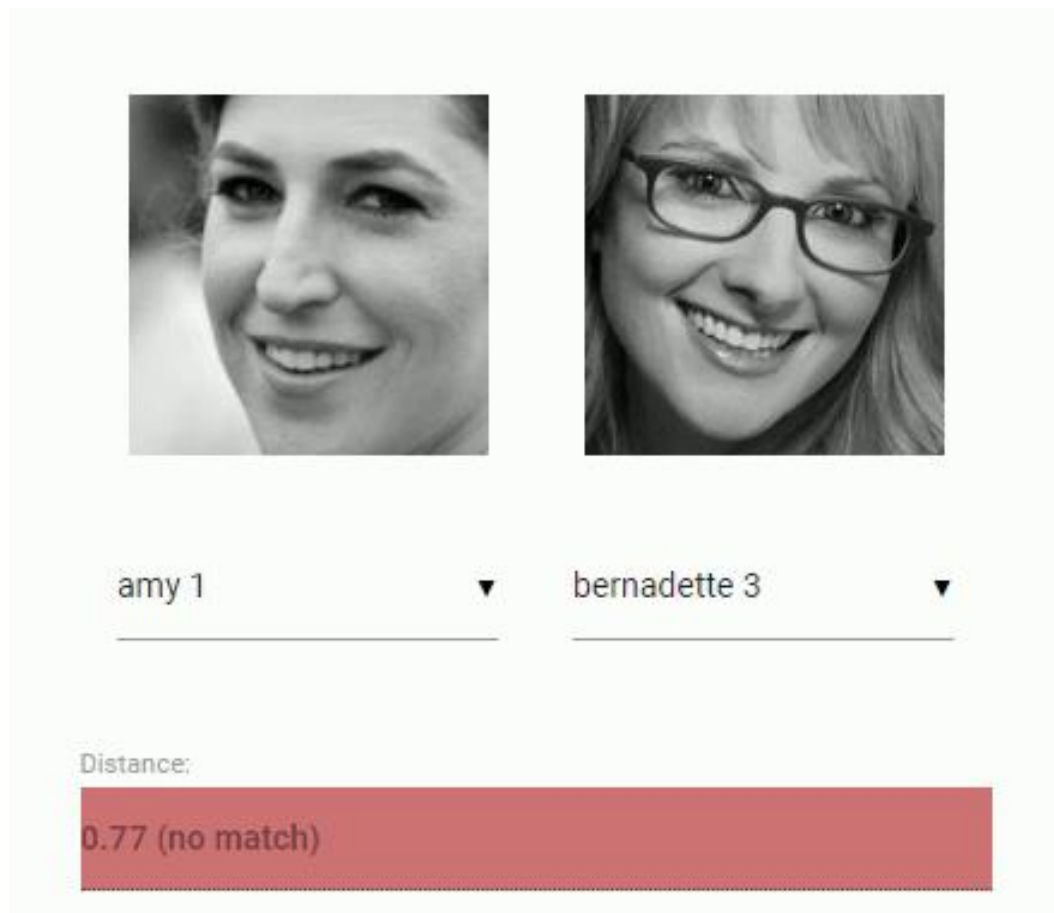


Рисунок 3.10 – Тестування негативного кейсу з жінкою

Далі необхідно провести аналогічні експерименти із зображеннями чоловіків. В першу чергу порівняємо двох різних людей (рис. 3.11), а потім порівняємо обличчя також різних, але схожих чоловіків (рис. 3.12).

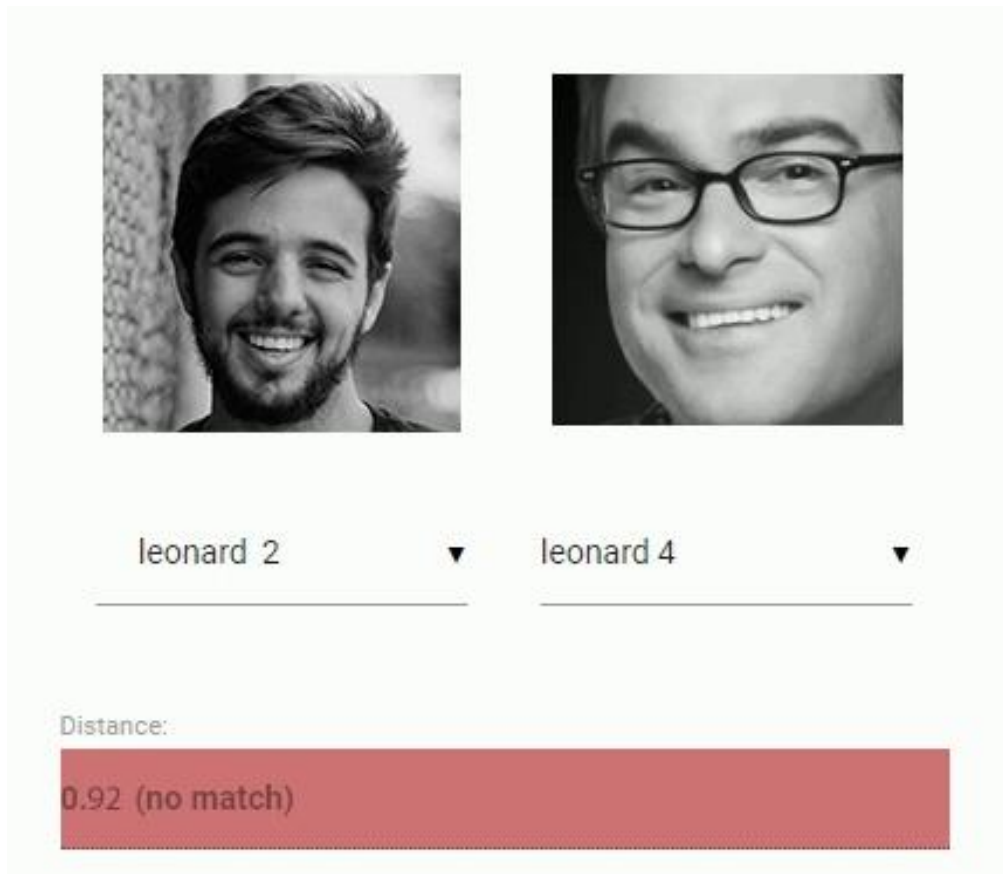


Рисунок 3.11 – Тестування негативного кейсу з чоловіком

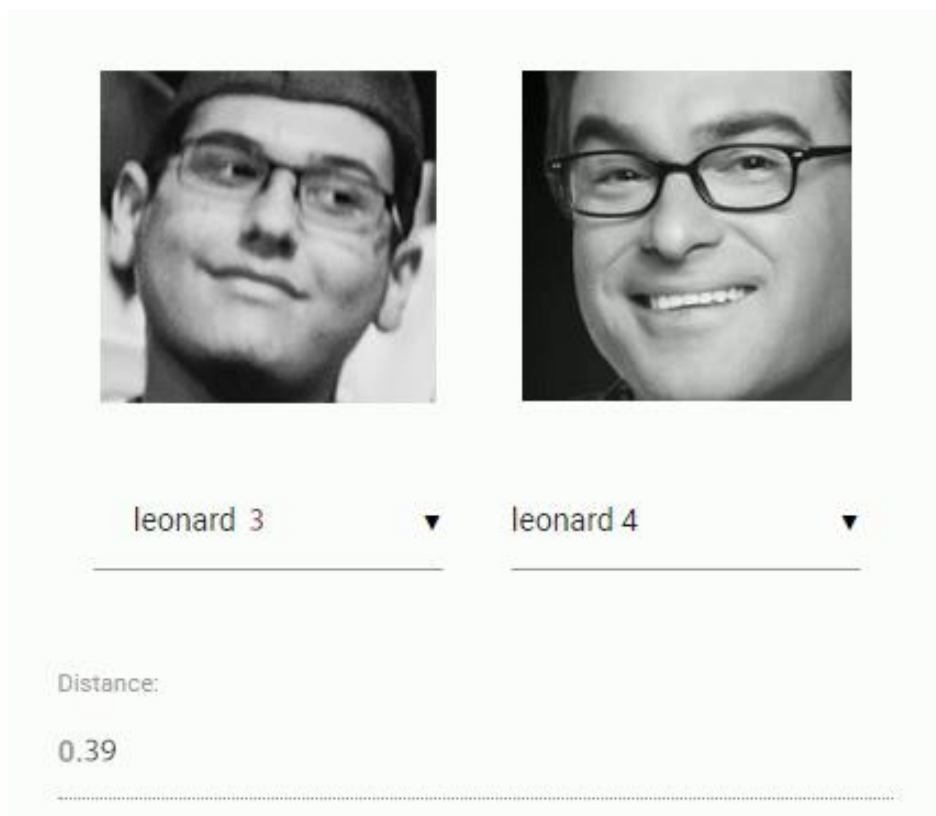


Рисунок 3.12 – Тестування негативного кейсу з чоловіком

У ході даного експерименту виявлено, що різних чоловіків з різними рисами обличчя алгоритм коректно визначає як різних людей зі значенням індексу 0,92. Що стосується зображень схожих чоловіків, то наразі ми отримали помилку. Індекс має значення 0,39, тобто алгоритм виявив двох різних людей як одного чоловіка. Даний результат може бути пов'язаний з тим, що ракурси є зовсім різними, якість зображення не є максимальною, через що базові точки на обличчя встановлено некоректно.

Для запобігання такого вигляду проблем необхідно збільшувати навчальну вибірку з різними ракурсами та розробити метод запобігання виявлення двох схожих людей як одного й того ж. Це необхідно, бо за умов помилки можливий негативний кейс ідентифікації, наприклад, близнюків або двійників.

ВИСНОВКИ

У рамках атестаційної роботи було розглянуто та проаналізовано метод ідентифікації обличчя для проходження ідентифікації у CMS системах фармацевтичних компаній.

В даний час напрямок «розпізнавання образів» набуло величезної популярності серед дослідників. Це пов'язано з тим, що рішення задач в даній області знаходить все більше прикладне застосування в повсякденній діяльності людини. Як приклад можна привести розпізнавання: штрих кодів, автомобільних номерів, осіб, облич, тощо. Особливий інтерес використання методів розпізнавання проявляється в дослідженні соціально-економічних систем. Це обумовлено тим, що завдання, які вирішуються в розглянутих областях, слабо формалізовані і класичні підходи не дозволяють отримати достовірні результати, які мали б практичну користь. У зв'язку з цим зараз спостерігається стійка тенденція до зростання популярності використання способів розпізнавання образів в соціально-економічних системах.

Постійний розвиток теоретичного, математичного апарату також є важливим фактором у популярності комп'ютерного зору та розпізнавання образів. Серед основних підходів для вирішення завдань розпізнавання образів варто виділити: класифікації за допомогою вирішальних функцій і функцією відстані; різні алгоритми кластеризації; машини опорних векторів; нейронні мережі і різні статистичні методи.

Алгоритм, який необхідно розробити повинен бути застосован у системах фармацевтичних компаній з метою ідентифікації людей перед замовленням ліків. Наразі, розвивається ринок онлайн-продажів медикаментів. Багато ліків можливо купити лише за рецептом. Також, в Україні починає розвиватися страхова медицина. До цього, в Україні розвивається використання цифрових документів. Виходячи з усіх цих пунктів обов'язково потрібно ідентифікувати людей перед оформленням замовлення ліків. У системах необхідно використовувати методи порівняння людей, що

замовляти медикаменти з їх фото з офіційних документів, таких як паспорт громадянина країни.

У процесі написання роботи розроблено згорткову нейронну мережу. Для цього було використано мову програмування Java Script, бібліотеки Node JS та Tensor Flow. Нейронна мережа заснована на базі детектора Single Shot MultiBox Detector.

Перед початком роботи було розглянуто актуальність проблеми розпізнавання облич у системах реєстрації. Розглянуто основні задачі розпізнавання образів та методи їх вирішення. Після аналізу існуючих методів виявлено, що для задачі, яка була поставлена на початку роботи кращим методом є створення нейронної мережі.

На наступному етапі було проаналізовано типи нейронних мереж та обрано саме згорткову нейронну мережу. Перед реалізацією було створено тренувальний набір даних, в якому були присутні обличчя людей. На наступному етапі обрано базові контури для задоволення різних умов на зображеннях. Наступним кроком було визначення карти з контрольними точками та контурами об'єктів, необхідних для тренування. Після етапа тренування нейронної мережі було реалізовано інтерфейс програми та проведено тестування.

Для тестування обрано декілька зображень людей з різних ракурсів та з різним освітленням. Результатом роботи програми є відповідь системи чи схожі завантажене зображення людини та його базове фото для порівняння. Таким чином, було виявлено, що система добре розпізнає однакових людей на різних фото.

Показано, що для вирішення задачі розпізнавання образів оптимальні згорткові нейронні мережі. Розроблено топологію згорткової нейронної мережі, що забезпечує інваріантність до змін умов зйомки, поворотам зображень, якістю зйомки.

На наступних етапах роботи метою є саме інтегрування даного методу проходження ідентифікації до різноманітних CMS систем на фармацевтичному ринку.

Результати даної роботи апробовані на 24-му Міжнародному молодіжному форумі «Радіоелектроніка і молодь в XXI столітті» на конференції «Сучасні методи обробки зображень» [41].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
2. Путятін, Є. П., Гороховатський, В. О., & Матат, О. О. (2006). *Методи та алгоритми комп'ютерного зору: навч. посіб.* Харків: ТОВ «Компанія СМІТ».
3. Frago Santos, J., Maksimović, P., Naudžiūnienė, D., Wood, T., & Gardner, P. (2017). JaVerT: JavaScript verification toolchain. *Proceedings of the ACM on Programming Languages*, 2(POPL), 1-33.
4. Дорогой, Я. Ю. (2011). *Архитектура обобщенных сверточных нейронных сетей.*
5. Tilkov, S., & Vinoski, S. (2010). Node.js: Using JavaScript to build high-performance network programs. *IEEE Internet Computing*, 14(6), 80-83.
6. Lytvynenko, V., Babichev, S., Wójcik, W., Vynokurova, O., Vyshemyrskaya, S., & Radetskaya, S. (Eds.). (2019). *Lecture Notes in Computational Intelligence and Decision Making: Proceedings of the XV International Scientific Conference «Intellectual Systems of Decision Making and Problems of Computational Intelligence»(ISDMCI'2019), Ukraine, May 21–25, 2019 (Vol. 1020).* Springer.
7. Cantelon, M., Harter, M., Holowaychuk, T. J., & Rajlich, N. (2014). *Node.js in Action* (pp. 17-20). Greenwich: Manning.
8. Ye. Bodyanskiy A. Shafronenko. Online algorithm for possibilistic fuzzy clustering based on evolutionary cat swarm optimization. *Science and Education a New Dimension. Natural and Technical Sciences*, 2019, 193, P. 86-88.
9. Bodyanskiy Ye. V., Shafronenko A. Yu., Rudenko D. O., Klymova I. M. Online recurrent method of credibilistic fuzzy clustering. *Topical issues of the development of modern science. 5th International scientific and practical conference.* Publishing House «ACCENT». Sofia, Bulgaria. 2020. P. 37-40.

10. Bodyanskiy Ye, Shafronenko A., Mashtalir S., Online robust fuzzy clustering of data with omissions using similarity measure of special type - Lecture Notes in Computational Intelligence and Decision Making-Cham: Springer, 2020-P.637-646

11. Shafronenko, A., Bodyanskiy, Ye., Rudenko, D.: Neuro-fuzzy clustering of Distorted Data Using Cat Swarm Optimization. Saarbrücken, LAP LAMBERT Academic Publishing (2020).

12. Gorokhovatskyi, V.O., Tvoroshenko, I.S., and Peredrii O.O. (2020) Image classification method modification based on model of logic processing of bit description weights vector, Telecommunications and Radio Engineering, 79(1), pp. 59-69. DOI: 10.1615/TelecomRadEng.v79.i1.60. Available online: <http://www.dl.begellhouse.com/ru/journals/0632a9d54950b268,706f31e606bab3dd,2492ae044887304a.html>. CMS collaboration. (2016). The CMS trigger system. arXiv preprint arXiv:1609.02366.

13. Романов, А. А. (2018). Сверточные нейронные сети. In 21-я Международная научно-практич. конф.«Научные исследования: ключевые проблемы III тысячелетия»: сб. тр. конф (р. 5).

14. Kim, Y. (2014). Convolutional neural networks for sentence classification. arXiv preprint arXiv:1408.5882.

15. Гороховатский, В. А. (2014). Структурный анализ и интеллектуальная обработка данных в компьютерном зрении.

16. Kobylin, O., Lyashenko, V., & Shafronenko, A. (2019, August). Wavelet Ideology and Morphological Operations for Analyzing RGB Images of Cytological Preparations. In 2019 7th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW) (pp. 119-122). IEEE.

17. Dayley, B. (2014). Node. js, MongoDB, and AngularJS web development. Addison-Wesley Professional.

18. Тропченко, А. А., & Тропченко, А. Ю. (2012). Нейросетевые методы идентификации человека по изображению лица. Известия высших учебных заведений. Приборостроение, 55(10).

19. Lyashenko, V., Kobylin, O., & Ahmad, M. A. (2014). General methodology for implementation of image normalization procedure using its wavelet transform. *International Journal of Science and Research (IJSR)*, 3(11), 2870-2877.

20. Ding, C., & Tao, D. (2017). Trunk-branch ensemble convolutional neural networks for video-based face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 40(4), 1002-1014.

21. Gorokhovatskyi, V., Gorokhovatskyi, O., Yevgenyi, P., & Olena, P. (2018, August). Quantization of the Space of Structural Image Features as a Way to Increase Recognition Performance. In *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)* (pp. 464-467). IEEE.

22. Гороховатский, В. А., & Путятин, Е. П. (2008). Структурное распознавание изображений на основе моделей голосования признаков характерных точек. Реєстрація, зберігання і обробка даних.

23. Bodyanskiy, Ye. Adaptive clustering of incomplete data using neuro fuzzy Kohonen network / Ye. Bodyanskiy, A. Shafronenko, V. Volkova // *Artificial Intelligence Methods and Techniques for Business and Engineering Applications*. – ITHEA 2012, Rzeszow, Poland; Sofia, Bulgaria. – P. 287-296.

24. Shafronenko, A., Dolotov, A., Bodyanskiy, Y., Setlak, G.: Fuzzy clustering of distorted observations based on optimal expansion using partial distances. In: *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, pp. 327 – 330

25. Бодянский Е.В, Шафроненко А.Ю. Рандомизированная модификация метода оптимизации на основе кошачьих стай. - *Системы обработки інформації*. – 2018. – № 1(152). – С. 142-147.

26. Hu, Zh., Bodyanskiy, Ye, Tyshchenko, O., Shafronenko, A.: Fuzzy clustering of incomplete data by means of similarity measures. In: *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering UKRCON - 2019, Conference Proceedings, July 2-6, 2019, Lviv, Ukraine*, pp.149-152 (2019) doi: 10.1109/UKRCON.2019.8879844.

27. Delcev, S., & Draskovic, D. (2018, May). Modern JavaScript frameworks: A Survey Study. In 2018 Zooming Innovation in Consumer Technologies Conference (ZINC) (pp. 106-109). IEEE.

28. Михалин, Д. А., & Белов, Ю. С. (2018). Сравнение использования многослойных и сверточных нейронных сетей для задачи идентификации по лицу. In НАУЧНЫЕ ИССЛЕДОВАНИЯ В ОБЛАСТИ ТЕХНИЧЕСКИХ И ТЕХНОЛОГИЧЕСКИХ СИСТЕМ (pp. 171-175).

29. Gorokhovatskyi, V.O., Tvoroshenko, I.S., and Peredrii O.O. (2020) Image classification method modification based on model of logic processing of bit description weights vector, Telecommunications and Radio Engineering, 79(1), pp. 59-69. DOI: 10.1615/TelecomRadEng.v79.i1.60. Available online: <http://www.dl.begellhouse.com/ru/journals/0632a9d54950b268,706f31e606bab3dd,2492ae044887304a.html>.

30. Шапиро, Л., Стокман, Д., Богуславский, А. А., & Соколов, С. М. (2013). Компьютерное зрение.

31. Шубин, И., Чернов, В., Гриценко, В., & Кириченко, И. (2012). Модели интеллектуальной адаптивной поддержки навигации в компьютерных обучающих системах. INFORMATION MODELS & ANALYSES, 194.

32. Есиков, А. В. (2013). Сравнительный анализ систем управления контентом (Content management system–CMS). Перспективы развития информационных технологий, (13), 62-65.

33. Setlak, G., Alexandrov, M., & Markov, K. (2012). Artificial Intelligence Methods and Techniques for Business and Engineering Applications. Rzeszow, Poland.

34. Гороховатский, В. А. (2014). Структурный анализ и интеллектуальная обработка данных в компьютерном зрении.

35. Захожай, О. І. (2013). Інформаційна технологія розпізнавання образів в задачах автоматизованої обробки інформації управління складними системами. Проблеми інформаційних технологій, (1), 61-68.

36. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... & Ghemawat, S. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.
37. Алексюк, Ю. А. (2018). Детектування і розпізнавання об'єктів за допомогою комп'ютерного зору.
38. Рябенський, В. М., & Захожай, О. І. (2011). Комбіновані системи розпізнавання образів. Проблеми інформаційних технологій, (1)
39. De G. Matthews, A. G., Van Der Wilk, M., Nickson, T., Fujii, K., Boukouvalas, A., León-Villagrà, P., ... & Hensman, J. (2017). GPflow: A Gaussian process library using TensorFlow. The Journal of Machine Learning Research, 18(1), 1299-1304.
40. Bodyanskiy, Y., Shafronenko, A., Mashtalir, S. Online Robust Fuzzy Clustering of Data with Omissions Using Similarity Measure of Special Type. Advances in Intelligent Systems and Computing. 2019. P.637-646. DOI: https://doi.org/10.1007/978-3-030-26474-1_44
41. Лапченко В. О. (2020). Технології розпізнавання облич.