

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**Факультет інформаційних технологій
Кафедра інтелектуальних технологій**

**КВАЛІФІКАЦІЙНА РОБОТА
на здобуття освітнього ступеня «магістр»**

НА ТЕМУ:

Система оперативного пошуку в неоднорідному
інформаційному просторі в задачах ідентифікації людей

Галузь знань: 12 «Інформаційні технології»

Спеціальність: 122 «Комп'ютерні науки»

Освітньо-наукова програма «Технології штучного інтелекту»

Виконав:

студент 2 курсу магістратури, групи ТШП-21

Бондаренко Богдан Сергійович
(ПІБ)

Науковий керівник:

Самохвалов Юрій Якович
(ПІБ)

Д.Т.Н., професор
(науковий ступінь, вчене звання)

Засвідчую, що в цій кваліфікаційній роботі
немає запозичень з праць інших авторів без
відповідних посилань

Студент

_____ підпис

Кваліфікаційна робота допущена до захисту
рішенням кафедри *інтелектуальних технологій*

Протокол № ____ від « ____ » травня 2021 р.

Зав. кафедри _____ доц. Іларіонов О.Є.

Київ 2021

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	3
ВСТУП.....	4
РОЗДІЛ 1. РОЛЬ ТА МІСЦЕ ПОШУКУ ІНФОРМАЦІЇ В ЗАДАЧАХ ІДЕНТИФІКАЦІЇ ЛЮДЕЙ.....	7
1.1 Аналіз особливостей ідентифікації людей та афілійованих задач	7
1.2 Аналіз процесів пошуку інформації в задачах ідентифікації людей.....	14
1.3 Аналіз моделей, технологій та засобів пошуку інформації.....	19
1.4 Постановка задачі.....	26
1.5 Висновки до розділу 1.....	27
РОЗДІЛ 2. РОЗРОБКА ТЕХНОЛОГІЇ ПОШУКУ В НЕОДНОРІДНОМУ ІНФОРМАЦІЙНОМУ ПРОСТОРІ.....	29
2.1 Опис технології пошуку інформації	29
2.2 Розробка алгоритму ідентифікації особи на основі порівняння зображень.....	31
2.3 Розробка алгоритму ідентифікації особи за текстовим запитом ...	42
2.4 Розробка алгоритму автоматичного розрахунку параметрів зовнішності.....	44
2.5 Висновки до розділу 2.....	50
РОЗДІЛ 3. РОЗРОБКА СИСТЕМИ ПОШУКУ ІНФОРМАЦІЇ....	52
3.1 Розробка архітектури системи пошуку інформації.....	52
3.2 Визначення структури баз даних	58
3.3 Висновки до розділу 3.....	61
РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНА ВЕРИФІКАЦІЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ СИСТЕМИ ПОШУКУ ІНФОРМАЦІЇ.....	62
4.1 Експериментальна верифікація роботи системи на основі контрольного прикладу.....	62
4.2 Аналіз ефективності розробленої системи пошуку інформації.....	71
4.3 Висновки до розділу 4.....	78
ВИСНОВКИ.....	79
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	82

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, СКОРОЧЕНЬ І ТЕРМІНІВ

AIC – автоматизована інформаційна система

ІІМ – інформаційно-пошукова мова

ІІС – інформаційно-пошукова система

ІС – інформаційна система

СУБД – система управління базами даних

CNN – Convolution Neural Network (згортова нейронна мережа)

HOG – Histogram of Oriented Gradients (гістограма напрямлених векторів)

SQL – Structured Query Language (мова структурованих запитів)

SVM – Support Vector Machines (метод опорних векторів)

ВСТУП

Актуальність теми

За останні кілька десятиліть інтерес до розробки алгоритмів ідентифікації людей стрімко зростає. Відеоспостереження, ідентифікація та пошук злочинців або зниклих людей, контроль доступу до об'єктів критичної інфраструктури, а також безпілотні транспортні засоби – лише декілька прикладів конкретних задач, в яких необхідно розпізнавання людей. Завдяки розробці нових інформаційних технологій та алгоритмів безпеки, для ідентифікації людей використовуються біометричні ознаки, такі як райдужна оболонка, відбитки пальців, голос та обличчя для розпізнавання людей. Використання біометричних характеристик набуло популярності через складність їх моделювання та простоту застосування.

Найбільш поширеним в процесі ідентифікації людини є розпізнавання обличчя, через високу точність встановлення особи при їх порівнянні. Проблема формалізації і автоматизації процесу ідентифікації людських облич була сформульована ще на самих ранніх стадіях розвитку систем розпізнавання та залишається актуальною і у теперішній час. Однак протягом останніх років кількість наукових досліджень і публікацій з даної тематики різко збільшилась, що свідчить про зростання актуальності цієї проблеми.

Останнім часом стали важливими питання розробки інформаційно-пошукових систем розпізнавання обличчя людини з метою ідентифікації її особистості. Підсистеми ідентифікації особи з використанням технологій розпізнавання обличчя, можуть бути використані не тільки в системах безпеки, а також при ідентифікації та проведенні пошуку розшукуваних осіб, з подальшою передачею відповідної оперативної інформації спеціальним службам.

Розроблені підходи ідентифікації особи по зображенням обличчя практично устоялися. Проте на даний момент відсутні системи, які могли б порівнювати фотографію людини з параметрами зовнішності, які виражені в словесній формі

або зберігаються у формі числових значень. Зокрема, актуальною є розробка системи, яка могла б обробляти не лише зображення, а й використовувати словесний портрет для встановлення особи людини. При цьому необхідна розробка підходу до ідентифікації особи за рахунок використання ключових ознак зовнішності (наприклад, висота носу, товщина губ, висота лобу та інше), що вилучаються та розраховуються автоматично в процесі обробки фотографії обличчя людини.

Об'єктом дослідження є процеси пошуку інформації в задачах ідентифікації людей.

Предметом дослідження є моделі, методи та технології пошуку в гетерогенних базах даних.

Метою дослідження є підвищення ефективності пошуку інформації при ідентифікації особи за рахунок розробки інформаційно-пошукової системи в неоднорідному інформаційному просторі, яка буде забезпечувати ідентифікацію особи в реальному часі за фотографією її обличчя або словесним портретом.

Завдання дослідження, які було поставлено і вирішено під час дипломного проектування у відповідності до поставленої мети:

- аналіз процесу пошуку інформації та визначення його ролі і місця в задачах ідентифікації людей;
- розробка технології пошуку в неоднорідному інформаційному просторі для вирішення задачі ідентифікації людей;
- створення методу автоматичної оцінки параметрів зовнішності людини;
- розробка архітектури інформаційно-пошукової системи та її імплементація;
- проведення експериментальної верифікації роботи розробленої системи та оцінки аналізу отриманих результатів.

При написанні дипломної магістерської роботи було використано такі **наукові методи**, як метод опорних векторів, використання гістограми напрямлених векторів, метод каскаду дерев регресії, нейронні мережі та ін.

Наукова новизна полягає в створенні технології пошуку в неоднорідному інформаційному просторі для вирішення задачі ідентифікації людей, створенні методу автоматичної оцінки параметрів зовнішності особи, а також розробці інформаційно-пошукової системи, яка здатна виконувати ідентифікацію осіб за зображенням та словесним портретом.

Практична значущість роботи полягає у створенні інформаційно-пошукової системи, яка дозволяє вирішувати задачі ідентифікації осіб і може використовуватися в системах відеоспостереження та контролю доступу, а також проведення розшуку злочинців або зниклих людей.

Апробація результатів дослідження Отримані результати були представлені і одержали схвалення на VII Міжнародній науково-практичній конференції “Інформаційні технології та взаємодії”; м. Київ, 2-4 грудня 2020 року, IV Міжнародній науково-практичній конференції “Проблеми кібербезпеки інформаційно-телекомунікаційних систем” (PCSITS), м. Київ, 15-16 квітня 2021 року та опубліковано у фаховому журналі «Безпека інформаційних систем і технологій», випуск №4, 2020 рік.

Структура та обсяг роботи: магістерська робота складається із вступу, чотирьох розділів, висновків, списку використаних джерел із 43 найменувань. Магістерська робота містить 17 рисунків та 2 таблиці. Загальний обсяг дипломної роботи складає 81 сторінку без списку літератури.

РОЗДІЛ 1

РОЛЬ ТА МІСЦЕ ПОШУКУ ІНФОРМАЦІЇ В ЗАДАЧАХ ІДЕНТИФІКАЦІЇ ЛЮДЕЙ

1.1 Аналіз особливостей ідентифікації людей та афілійованих задач

Ідентифікація – це процес встановлення відповідності невідомого об'єкта відомому на підставі збігу деяких ознак. Ідентифікація особи – процес однозначного визначення особистості людини. Ідентифікація особи може бути необхідною в багатьох задачах, наприклад в процесах її пошуку, прийняття рішення про наявність доступу або відстеження переміщення тощо. Особу можна однозначно визначити за допомогою деякого унікального ідентифікатору, яким може служити фізичний предмет, наприклад паспорт, ідентифікаційна картка, або введення таємного слова – паролю або персонального ідентифікаційного номеру (код, який складається тільки з цифр). У випадку, коли потрібно визначити людину без залучення таких ідентифікаторів, наприклад в процесі встановлення її місцезнаходження за допомогою записів з камер відеоспостереження, або в разі відсутності спеціального ідентифікатору, використовують біометричні дані. До біометричних даних відносяться такі дані, як відбитки пальців, форма та ознаки обличчя, голос, райдужна оболонка ока, малюнок сітківки ока, геометрія руки, почерк, особливості друку на клавіатурі і навіть візерунок вен на зап'ясті. Біометрична ідентифікація – це процес порівняння та визначення подібності між даними людини і її біометричним «шаблоном». Біометрія дозволяє ідентифікувати та проводити верифікацію особи на основі набору унікальних ознак, властивих їй від народження.

Завдання систем ідентифікації на основі біометричних даних, таких як розпізнавання обличчя, останнім часом набуває важливого значення. Крім того, багато вчених та інженерів у всьому світі зосереджуються на створенні все більш надійних і точних алгоритмів та методів для таких типів систем та їх застосуванні в повсякденному житті. Завдяки розробці інформаційних технологій, багато систем безпеки починають використовувати біометричні дані для вирішення завдання розпізнавання [1 – 4]. Біометричні властивості людини дозволяють ідентифікувати особистість за її фізіологічними або поведінковими

характеристиками. Вони мають значні переваги використання, наприклад, для ідентифікації достатньо присутності людини перед камерою або спеціальним датчиком, тому зникає потреба запам'ятовувати паролі та конфіденційні коди. У цьому контексті впродовж останніх років було застосовано багато систем розпізнавання, заснованих на різних біометричних факторах, таких як райдужна оболонка, відбитки пальців [5], голос [6], хода [7], та обличчя. Всі наведені підходи мають свої переваги та недоліки. Наприклад, системи ідентифікації, які використовують розпізнавання по відбиткам пальців або голосу, не можуть бути використані в задачі пошуку зниклої людини. В той же час, як використання особливостей ходи або зовнішності малоефективні при низькій роздільній здатності відеозапису.

Системи, що ідентифікують людей на основі їх біологічних особливостей, дуже перспективні, оскільки ними легко користуватися. З цієї причини за останні роки вони стали однією з найбільш широко використовуваних біометричних систем автентифікації, враховуючи її потенціал у різних додатках та галузях (відеоспостереження, безпека дому, прикордонний контроль тощо) [8 – 10]. Системи розпізнавання обличчя, як посвідчення особи вже пропонуються в місцях реєстрації в аеропорту, на спортивних стадіонах та концертах. Крім того, дані системи не вимагають втручання людей для роботи, що дозволяє ідентифікувати людей лише за зображеннями, отриманими з камери. Багато біометричних систем ідентифікації, розроблені з використанням різних типів пошуку, забезпечують хорошу точність визначення особи.

Існує багато методів комп'ютерного зору, які призначені для вирішення завдань виявлення або розпізнавання обличчя з високою інваріантністю при спотворенні даних та точністю, таких як локальний, поділ на регіони та гібридний підходи [11 – 14]. Однак існує низка проблем при ідентифікації особи за зображенням, які все ще потрібно вирішити, такі як орієнтація голови, умови освітлення та вираз обличчя. Багато методів розроблено, щоб протистояти цим проблемам, а отже, створювати надійні системи розпізнавання облич. Тим не

менше, вони вимагають значного часу обробки, великого споживання пам'яті та відносно складні.

Визначення людини на основі біометричних ознак є останніми доповненнями до існуючих механізмів, що використовуються для захисту інформаційних систем та контролю доступу. Серед багатьох інших, розпізнавання облич є найефективнішою біометричною технікою для ідентифікації та перевірки особистості людини. Розпізнавання облич із відеозапису привернуло увагу завдяки своїй популярності та простоті використання в системах безпеки на базі систем спостереження. Автоматизована система розпізнавання облич на основі відео пов'язана з величезним спектром завдань, оскільки необхідно проводити перевірку обличчя за різних обставин. Для багатьох систем безпеки, як визначено в роботах [15 – 18], характеристики, які роблять систему розпізнавання обличчя ефективною, є: здатність працювати як з відео, так і з зображеннями, виконувати ідентифікацію в режимі реального часу, мати стійкість до різних умов освітлення та вміти розпізнавати обличчя людей, які знаходяться під різними кутами повороту до камери. Розпізнавання обличчя у відео продовжує привертати велику увагу дослідників у всьому світі, тому в цій галузі зафіксовано значний прогрес.

У традиційних сценаріях спостереження, користувачі повинні переглядати відеозаписи, що відповідають тривалим періодам часу, щоб знайти потрібну подію. Однак цей процес вимагає ресурсів і потребує затрат на наймання персоналу охорони. Область інтелектуального візуального спостереження прагне вирішити ці проблеми, застосовуючи методи комп'ютерного зору для автоматичного виявлення конкретних подій у відеопотоках. Знайдені події можуть бути представлені користувачеві або проіндексовані в базі даних, щоб дозволити проводити пошук за деяким запитом, наприклад "знайти людей, які заїхали на стоянку з 19:00 до 21:00 у понеділок" або "показати обличчя всіх осіб, які відвідували дане місце минулого тижня".

Сучасні дослідження теми ідентифікації особи зосереджені на підходах, заснованих на розпізнаванні обличчя, де метою є встановлення особистості

людини, за зображенням її зовнішності. Однак розпізнавання обличчя все ще є дуже складною проблемою, особливо на знімках з низькою роздільною здатністю, з різним кутом повороту до камери та освітленням, що часто трапляється у даних з камер спостереження. Сучасні системи розпізнавання обличчя вимагають достатньої роздільної здатності для отримання надійних результатів.

Системи розпізнавання осіб не завжди здатні точно зіставити образи осіб з образами, збереженими в базі даних. Як правило, помилки виникають в результаті поганої якості зображень або недостатньої кількості інформації в базі даних. Погане освітлення або низька якість зображення можуть ускладнити виконання точного аналізу вузлових точок людини. Наприклад, на дані може вплинути затемнення деяких областей обличчя. Це створює при розрахунку ознак особи, в результаті чого буде неможливо зіставити його з правильними даними в базі даних.

В інших випадках база даних може бути недостатньо велика, щоб мати точне зіставлення для кожної особи. Наприклад, система розпізнавання, яка використовується ФБР (FACE Services) не змогла розпізнати 153 636 фотографій за період з серпня 2011 до квітня 2019 року.

Процесом розпізнавання осіб прийнято називати набір різних завдань, що служать для ідентифікації людини по цифровому зображенню або фрагменту відеозапису. У загальному вигляді цей процес виглядає наступним чином: після того, як система отримала зображення з камери, за допомогою алгоритмів визначаються межі регіону розташування особи на ньому (етап виявлення). Далі йде етап розпізнавання, на якому фотографія трансформується (регулюється яскравість, нормалізується, вирівнюється обличчя людини, масштабується, і т.д.) і приводиться до деякого заданого стандарту, з яким працює система. Після чого, відбувається обчислення ознак та безпосереднє порівняння їх із збереженими в базі даних еталонами. Цей заключний етап порівняння називається ідентифікація або верифікація, в залежності від мети застосування системи.

Хоча технології ідентифікації людей можуть здатися чимось футуристичним, тим не менш, вони вже активно використовуються в самих різноманітних сферах нашого життя. До задач, в яких застосовується ідентифікація людини на фото можна віднести:

- **Безпека доступу до даних електронного пристрою.** Деякі додатки використовують розпізнавання осіб для захисту даних, що зберігаються. Навіть безпечний пароль не може захистити акаунти та інформацію від досвідчених зловмисників, тому розробники вирішили застосувати технологію розпізнавання особи по фотографії. Ці додатки вимагають користувача показати обличчя, щоб розблокувати смартфон або авторизуватися в системі та отримати доступ до особистих даних.
- **Виявлення медичних порушень.** Існують спеціальні медичні програми, такі як Face2Gene і DeepGestalt, які використовують розпізнавання осіб для виявлення генетичних порушень. Вони аналізують обличчя та порівнюють його з базою даних обличчя тих осіб, у яких є різні порушення.
- **Забезпечення безпеки.** Деякі магазини оснащені системами розпізнавання людей, які визначають потенційно небезпечних людей, якщо вони щось крали в магазинах. Така система може ідентифікувати магазинного злодія та повідомити власника магазину про його минулі злочини, навіть якщо даний зловмисник ніколи не бував в даному магазині раніше. Хоча така система може надавати значні вигоди для власників магазинів, але часто ефективність таких систем ставиться під сумнів. Якщо невинна людина буде визначена як зловмисник, то це може вплинути на її життя.
- **Контроль віку.** Деякі продуктові магазини і бари в Великобританії використовують системи розпізнавання людей, щоб визначити, чи достатньо років покупцеві, щоб мати право покупки алкоголю. Продуктові магазини дозволяють покупцям використовувати систему самоперевірки, без необхідності в додатковій перевірці співробітників, перевіряючим паспорт. Якщо система вважатиме, що клієнту менше 25 років, то він повинен буде пред'явити паспорт для перевірки.

- **Безпека в школах.** Розпізнавання облич починають впроваджувати в школах. Одна школа в Швеції використовує технологію розпізнавання облич для перевірки відвідуваності на уроках. Школи в США, особливо в Нью-Йорку, починають тестувати використання технологій розпізнавання осіб в якості «системи раннього сповіщення» проти загроз з боку відомих злочинців. Технологія також може розпізнавати 10 видів зброї для запобігання актів насильства в школах.

- **Встановлення права на отримання послуг.** Такі авіакомпанії як Delta і JetBlue використовують розпізнавання осіб для ідентифікації пасажирів. Біометричне сканування особи є необов'язковим, але дозволяє пасажирам використовувати свої обличчя в якості квитка, економлячи час на перевірку квитків.

Ідентифікація людей може бути необхідною і у випадку, коли потрібно знайти особу, що зникла. Наприклад в правоохоронних органах, в разі звернення з заявою про зникнення людини, може з'явитися потреба пошуку зниклої особи на даних з камер відеоспостереження. Також ідентифікація особистості необхідна в системах нагляду на підприємствах, коли потрібно проводити моніторинг та контроль роботи співробітників. В таких випадках, в базі даних зображень мають міститись фотографії необхідної людини, або опис її зовнішності. Оскільки, зображення особистості не завжди є в наявності, необхідно мати систему, яка може виконувати пошук людини по опису її зовнішності, вираженому в словесній формі. Для цього необхідно мати в базі даних лише опис деяких параметрів зовнішності людини, наприклад пропорції обличчя (висота лобу, розмір носу, ширина губ і т.д.) або фізичних параметрів (вага, зріст). В такому випадку, з'явиться можливість, виконавши аналіз зображення або відео, на якому присутня необхідна особа, провести її пошук та ідентифікацію.

Коли трапляється злочин або зникнення людини, головною підказкою, яка веде до знаходження потрібних осіб, є словесний опис свідками. Відділи поліції створюють ескіз обличчя з опису для того, щоб встановити особу злочинців або опитати інших людей для встановлення місцезнаходження особи. Процес ідентифікації в такому випадку передбачає перегляд бази фотографій свідками. Деякі відділи поліції використовують обчислювальну техніку для допомоги у пошуку та ідентифікації особи. Співробітник служби ідентифікації вводить зріст, вагу та стать підозрюваного в програму, яка здійснює пошук у базі даних та видає список кандидатів. Потім офіцер переглядає зображення кожної людини з цього списку і порівнює його з ескізом, отриманим з описів свідків. Зображення отримані з баз даних реальних фотографій, які формує поліція та інші державні установи. Проблема, яку ми розглядаємо, є однією з найскладніших випадків ідентифікації за обличчям. Потрібно, отримуючи в якості вхідних даних зображення особи з камер відеоспостереження або словесного портрету, виконувати пошук по базі даних фотографій людей та базі параметрів зовнішності особи .

Протягом останніх десятиліть існує інтерес до проблеми автоматичного порівняння словесного портрету, складеного за показаннями свідків та оригінальними фотографіями осіб [19]. Розроблені системи не можуть підтримувати пошук та ідентифікацію особи не тільки за зображенням, але і у випадку наявності лише бази даних опису людей. Вхідною при цьому є словесний портрет людини або фотографія потрібної особи. Вирішення даної проблеми зможе допомогти у випадку оперативного пошуку зниклої людини. Отриманий словесний портрет або зображення-приклад можуть бути використанні при обробці відео з камер відеоспостереження та автоматичному пошуку особи в деякому районі. Для цього, система ідентифікації особи має вміти аналізувати зображення людей та встановлювати їх відповідність не тільки за зображенням, але й за числовими або словесними параметрами оцінки зовнішності людини, що зберігаються в базі даних.

1.2 Аналіз процесів пошуку інформації в задачах ідентифікації людей

Задача ідентифікації людей має різні підходи до вирішення. В залежності від методу ідентифікації, можуть використовуватись різні типи інформації та файлів, які необхідні в ході проведення процесу визначення особи. Такими даними можуть бути мультимедійні файли (зображення, відео, аудіо), структуровані записи, що зберігають дані про особу та її ознаки, або вихідні файли інформації, необхідні для роботи алгоритму. Наприклад, для ідентифікації особи за обличчям або ходою, потрібно мати наявності інші зображення людей, що будуть аналізуватися для тренування методу розпізнавання. Питання пошуку інформації та проведення ідентифікації людей в системах відеоспостереження розглядались в [20 – 24].

Під час ідентифікації особи, в будь-якому разі використовуються деякі файли або інформація, що внесена в базу даних. Це вимагає проведення пошуку відповідних файлів в базі даних.

Інформаційний пошук – процес знаходження, відбору і видачі певної інформації (документів, файлів, записів) за заздалегідь заданими ознаками з масивів даних будь-якого виду та на будь-яких носіях [25]. Причиною здійснення інформаційного пошуку є інформаційна потреба, виражена у формі інформаційного запиту.

Автоматизовану інформаційну систему (АІС), або інформаційну систему (ІС), можна визначити як сукупність інформаційних ресурсів, процесів і технологій, які збирають, перетворюють і поширюють інформацію. Узагальненою метою ІС є накопичення, зберігання, перетворення та пошук інформації для використання її в різних системах ідентифікації людей або процесі прийняття рішень. Сучасні ІС володіють складною, часто гетерогенною структурою і призначені для вирішення великої кількості різнотипних завдань автоматизації, що виникають.

В автоматизованих інформаційних системах, пошук здійснюється із залученням лінгвістичних, інформаційних, програмно-технічних,

технологічних, організаційних засобів і складених з них комплексів. Безпосередньо інформаційний пошук проводиться засобами інформаційно-пошукової системи, яка є підсистемою АІС.

Інформаційні системи можуть відрізнятися алгоритмами обробки інформації, принципами організації зберігання даних і т.д. Також необхідно відзначити, що завдання даних систем відрізняються [26].

Структура і функціональність ІС, як правило, визначається характером організації джерел і споживачів інформації, серед яких можуть бути як користувачі, так і інші ІС. Система пошуку інформації може бути частиною системи ідентифікації, яка використовує її в разі необхідності пошуку особи в різних базах даних.

Інформаційно-пошукова система (ІПС) це підсистема автоматизованої системи, яка реалізована на засобах електронно-обчислювальної техніки і призначена для знаходження та видачі необхідної інформації по заданим критеріям [25]. ІПС являє собою сукупність інформаційно-пошукової мови (ІПМ), (програмних) засобів та правил перекладу інформації на цю мову (для індексування), забезпечення її пошуку і критеріїв відповідності.

Системи пошуку інформації спочатку обробляли та виконували пошук лише тексту, оскільки вони працювали лише з текстовими файли. Проте сучасні ІС, особливо у випадку роботи в складі систем ідентифікації особи, мають справу з мультимедійною інформацією, що включає аудіо, зображення та відео [27]. Хоча деякі зі звичайних методів пошуку тексту можна застосувати до пошуку мультимедійної інформації, специфічний характер аудіо, зображень та відео інформації вимагає розробки нових інструментів та методів пошуку.

На відміну від звичайної системи управління базами даних (СУБД), система пошуку інформації повинна працювати з неоднорідними даними. Основною метою системи пошуку інформації є отримання інформації – фактичної інформації, або файлів, що містять інформацію, яка повністю або частково відповідає запиту. База даних може містити тези або повні тексти

документів, а також зображення, аудіо та відеоінформацію, а отже система пошуку повинна вміти виконувати пошук по всіх цих типах даних.

Система пошуку інформації, з якою працює система ідентифікації людей, може зберігати зображення, відео та аудіозаписи, необхідні в процесі визначення особисті, а також містити записи в структурованому вигляді та організованому за деякою моделлю даних.

Найчастіше записи, як використовуються при ідентифікації особи, зберігаються в реляційній базі даних. Реляційна база даних це сукупність даних, що організовані у вигляді набору формально описаних таблиць. Таблиці складаються із стовпців і рядків. У таблицях зберігається інформація про об'єкти, представлених в базі даних. У кожному стовпчику таблиці зберігається певний тип даних, в кожній комірці – значення атрибуту. Кожний рядок таблиці являє собою набір пов'язаних значень, що відносяться до одного об'єкту або сутності. Стовпець може бути позначеним унікальним ідентифікатором, який називається первинним ключем. Рядки з декількох таблиць можуть бути пов'язані з допомогою зовнішніх ключів. Зберіганням та пошуком потрібних записів в базі, в цьому випадку, займаються системи управління базами даних (СУБД), які за необхідності повертають необхідні записи з таблиці, у відповідь на запит пошукової системи. Запити до СУБД формуються на деякій структурованій мові. Найбільш поширеною мовою запитів до системи управління базами даних є SQL (Structured query language, мова структурованих запитів).

Інформаційно-пошукова система складається з шести основних підсистем[25]:

- підсистема документів;
- підсистема індексації;
- підсистема словникового запасу;
- пошукова підсистема;
- інтерфейс користувачької системи;

- підсистема визначення відповідності.

Коли інформаційно-пошукова система отримує запит, сформульований на природній мові, вона перетворює його в вигляд інформаційно-пошукової мови (ІПМ) – формальна мова, призначена для опису змісту файлів, що зберігаються в ІС. Інформаційно-пошуковій мові це знакові системи зі своїм алфавітом, лексикою, граматиною і правилами користування [26]. Процедура опису файлу та записів інформації на ІПМ називається індексуванням. В результаті індексування кожного файлу, йому приписується формальний опис – пошуковий образ. Аналогічним чином індексується і пошуковий запит, якому ставиться у відповідність його пошуковий образ на ІПМ. Алгоритми інформаційного пошуку засновані на порівнянні пошукового образу документу з пошуковим образом запиту. Загальну схему функціонування ІПС наведено на рис. 1.1.

В процесі індексування, для кожного файлу та запису, що зберігається в системі, будується пошуковий образ. Для кожного типу файлів потрібно застосовувати свій метод індексації. Наприклад, в системах ідентифікації людей за зображенням, в базі даних можуть зберігатися інші фотографії даної особи, що потребує свого підходу в процесі індексації.

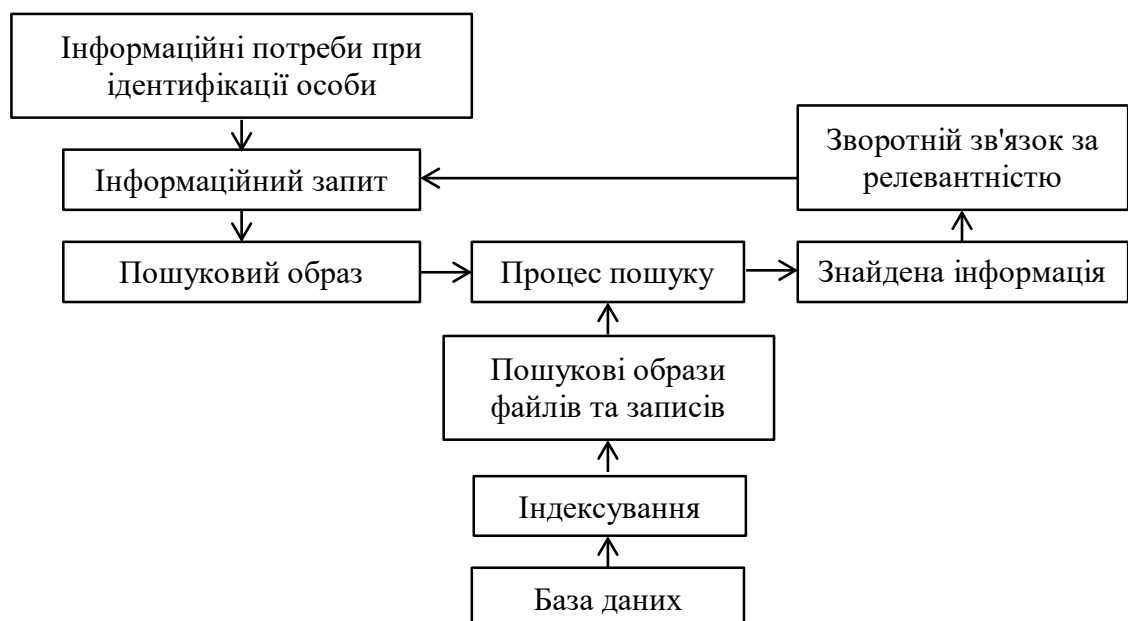


Рис. 1.1. Процес функціонування ІПС

Критерій відповідності знайдених файлів та записів до запиту називається релевантністю. Знайдені результати можна сортувати за даним критерієм, при цьому в процесі інформаційного пошуку можна відфільтрувати у видачі значний інформаційний шум – безліч файлів та записів, формально відповідних пошуковому запиту, але не відповідних за змістом.

Після отримання та перекладу запиту на ІПМ, система виконує його порівняння з образами файлів в базі даних, враховуючи критерії пошуку. Знайденні документи сортуються за релевантністю та відображаються користувачеві.

З точки зору функціональності в складі системи пошуку можна виділити такі блоки [25]:

- блок попередньої обробки – проводить індексацію файлів та запитів;
- блок пошуку – підбір файлу по пошуковому образу запиту з файлів, які задовольняють критеріям пошуку;
- блок пост-обробки – впорядкування знайдених файлів та записів за релевантністю.

Результати пошуку характеризують двома параметрами: повнота та точність. Повнотою пошуку (recall) називається міра, яка обчислюється як відношення кількості виданих релевантних файлів до загальної кількості файлів даного типу, що містяться в інформаційному просторі. Точність пошуку (precision) – це відношення кількості виданих релевантних файлів до загальної кількості файлів у виданому пошуковому результаті.

1.3 Аналіз моделей, технологій та засобів пошуку інформації

Інформація в базі даних, з якою працює інформаційно-пошукова система, може бути структурованою або неструктурованою. Саме структура даних

вимагає унікального підходу до обробки для отримання корисних знань з величезного потоку даних.

Структуровані дані – це інформація, упорядкована певним чином і організована з метою її подальшого аналізу. Така інформація не вимагає обробки для отримання з неї будь-яких необхідних даних. До такої інформації в задачах ідентифікації людей можна віднести таблиці реляційних баз даних, які зберігають записи про зовнішність людини, її фізичні характеристики, або вилучених з інших файлів (зображення, аудіо), за допомогою деякого методу даних.

Іншим типом даних є неструктуровані дані – це інформація, яка або не має наперед визначеної структури даних, або не організована в установленому порядку, зрозумілому для машини. Неструктуровані дані непридатні для обробки безпосередньо методами аналізу даних, тому такі дані обробляються методами структуризації, причому сам характер даних в процесі структуризації може істотно змінитися. До неструктурованих даних відносять дані на природній мові, машинні, графіки, потокові, аудіо, відео та графічні дані [27].

При індексуванні тексту, елементи що використовуються як основа індексації, включають символи, слова та фрази. Однак, маючи справу з неструктурованими даними, а саме мультимедіа файли – зображеннями, аудіо чи відео, повинні використовуватись інші методи обробки даних, такі як анотування та класифікація [28].

Система пошуку, що працює з неструктурованою інформацією, що необхідна в процесі ідентифікації особистості людини, повинна автоматично визначати значення вмісту файлу інформації. Наприклад, якщо в база даних зберігає файли зображень, то система повинна вміти визначати особу, що зображена та підтримувати доступ до колекцій зображень за запитом, який описує, хто саме зображений на фото або його основні візуальні параметри. В разі ідентифікації за записом голосу, може бути необхідним класифікація глибини голосу мовця. Кожен тип файлу даних вимагає свого підходу до зберігання та

подальшого пошуку. Проблеми збереження в базі даних та виконання подальшого пошуку неструктурованої інформації розглядалися в [28 – 32].

В дослідженнях по обробці зображень розглядали конкретні типи проблем, пов'язаних із пошуком файлів з врахуванням їх вмісту. В тому числі обробку обличчя, де розрізняють: розпізнавання обличчя, як ідентифікацію обличчя чи облич на сцені; розпізнавання обличчя, як підтвердження того, що дане обличчя належить певній особі; пошук людей зі схожим обличчям (знайти найближче відповідне обличчя у сховищі, що відповідно до прикладу або відповідає деяким критеріям пошуку). Також розроблено системи для відстеження людських рухів (наприклад, голови, рук, ніг) та ідентифікування людських виразів облич, таких як посмішка, здивування або гнів [33]. Для кожної конкретної задачі, в якій потрібно проводити процес пошуку інформації, необхідно використання окремих методів аналізу та пошуку файлів.

Залежно від форми запитів, алгоритми пошуку зображень поділяють на дві категорії: текстові підходи та методи, які обробляють вміст файлу для його подальшого пошуку. Текстові підходи засновані на зберіганні ключового слова, набору ключових слів або текстового опису вмісту зображення. Процес створення текстового опису для зображення називається анотуванням, а побудований опис – анотацією. Потім пошук зображень переходить до стандартних методів пошуку інформації в тексті. Легше реалізувати систему пошуку зображень на основі ключових слів або повнотекстових описів, ніж за змістом зображення, за умови, що можна отримати анотації зображень [27]. Обробка запитів у таких пошукових системах зазвичай дуже швидка завдяки існуючим ефективним технологіям управління базами даних.

Анотування великих колекцій зображень може бути довгим та ресурсозатратним процесом. Значення зображення може бути не зрозумілим, що ускладнює додавання анотацій до зображення [31]. Пошук зображень на основі вмісту націлений на ефективне отримання відповідних зображень із баз даних зображень на основі автоматично отриманих особливостей зображень. Ці особливості виділяються із властивостей форми, текстури, інтенсивності

освітлення або кольору зображення-запиту та зображень у базі даних. Система пошуку зображень розглядає зображення запиту та зображення в базі даних як сукупність ознак. Вона класифікує відповідність між запитом та цільовим зображенням пропорційно мірі подібності, розрахованої на основі ознак. Міра подібності кількісно визначає схожість за змістом ознак між парою зображень.

Вилучення ознак є фундаментальним компонентом будь-якої системи пошуку зображень на основі вмісту. Основним недоліком такого підходу є чутливість до змін інтенсивності, спотворень кольору та зсуву об'єкту, який потрібно визначити. Для усунення цього недоліку було запропоновано багато підходів до вилучення особливостей. Підхід вилучення ознак на основі поділу зображення на регіони застосовує сегментацію зображення для розділення зображення на колекцію регіонів. Ознаки визначаються для кожного регіону, в результаті чого, зображення представляється як сукупність векторів, кожен з яких відповідає регіону.

Низькорівневі мультимедійні потоки бітів не підходять для безпосереднього пошуку, і розпізнавання шаблонів потрібне для отримання більш значущих та корисних описів даних. Традиційно мультимедійні пошукові системи базуються на ручних анотаціях вмісту. Враховуючи типові кількості виробленого матеріалу, така анотація, як правило, є трудомістким та дорогим завданням. Для багатьох систем ручна анотація не дуже практична, а отже існує потреба в автоматизації цього процесу. І хоча в даний час поки що неможливо розробити системи, які анотують з таким самим рівнем деталізації, як людина. Тим не менше, існує велика кількість застосувань для систем пошуку, які автоматично надають деякі основні анотації.

З огляду на такий чітко визначений набір тематики файлів, які зберігаються в базах даних, підхід до класифікації файлів для автоматизованої анотації полягає у розробці класифікаторів для кожного з понять у наборі тем. На практиці це означає, що невелика частина мультимедійної колекції все ще анотована вручну, щоб отримати набір анотованих прикладів для навчання

класифікаторів. Після навчання класифікаторів, основна частина колекції анотується автоматично.

Одним з підходів до класифікації файлів є розпізнавання шаблонів, що має на меті класифікувати дані на основі апріорних знань, або статистичної інформації, вилученої з них. Система розпізнавання шаблонів складається з екстрактору ознак, який обчислює числові або символічні ознаки на основі даних і класифікатору, який використовує вилучені ознаки для відношення їх до відповідних категорій. В якості екстракторів ознак можуть виступати статистичні методи, які, наприклад, оцінюють розподіл кольорів та яскравість зображення, або частоту звуку аудіофайлів.

Ознаками, що аналізуються при обробці файлів зображення, можуть виступати інтенсивність або значення градієнтів пікселів. При врахуванні інтенсивності зображення, використовують гістограму інтенсивності, яка підраховує кількість пікселів для кожного різного значення інтенсивності, знайденому на зображенні. Використовуючи значення градієнтів як ознаки зображення, будують гістограму орієнтованих градієнтів (HOG, Histogram of Oriented Gradients), яка може бути використана як дескриптор об'єктів при його пошуку, де випадки орієнтації градієнта в локалізованих частинах зображення відіграють важливу роль. Гістограма орієнтованих градієнтів оперує на сітці пікселів та використовує локальну нормалізацію контрасту для підвищення точності. Дескриптори HOG вперше описані в [34].

Після вилучення ознак деяким методом, до отриманих значень застосовують деякий метод класифікації. До таких належать методи:

- k-найближчих сусідів,
- метод опорних векторів (SVM, Support vector machines)
- наївний баєсів класифікатор.

Отримані класифіковані файли даних поміщуються в базу даних для їх подальшого пошуку.

Методи, що використовують статистичні характеристики файлів для їх аналізу та пошуку мають свої переваги та недоліки. При цьому можна виділити загальні особливості їх застосування. Перевагами застосування таких методів є:

- висока швидкість обробки кожного файлу;
- низькі затрати електроенергії на обчислення та обчислювальна складність для вилучення ознак файлу;
- легко імплементувати.

До основних недоліків використання статистичних методів при обробці файлів відносять:

- відсутність або низька стійкість до класифікації при спотворенні даних, наприклад повороті або масштабуванні об'єкта на зображенні;
- відносно низькі показники точності класифікації – в середньому 60%.

Для пошуку мультимедіа файлів в неоднорідному інформаційному можуть застосовуватися нейронні мережі. Вони мають вищу точність пошуку, в порівнянні з традиційними методами, адже можуть виконувати пошук, враховуючи вміст файлу, при цьому володіють інваріантністю при спотворенні даних. Розроблені прості методи, такі як аналіз гістограми кольорів зображення або частоти звуку, мають високу швидкість обробки, але не здатні виконувати пошук при незначних перетвореннях даних.

Основною причиною використання нейронних мереж для пошуку зображень є їх адаптивна здатність до навчання [32]. Нейронні мережі пропонують некеровану кластеризацію, а також контрольовані механізми навчання для пошуку зображень. При створенні анотації зображень за допомогою нейромереж, класифікують об'єкти та явища, що зображені на них та визначають пов'язані з ними ключові слова.

При обробці зображень нейронними мережами, створюють їх анотацію, для цього виконується класифікація. Існують різні архітектури нейронних мереж для класифікації. Майже всі вони використовують як основу згорткові шари та шари агрегування. За допомогою спеціально натренованих згорткових

нейромереж можна виконувати класифікацію та опис зображень. Після класифікації зображення, створюється його текстовий опис. В подальшому, виконується семантичний пошук по описам зображень та знаходяться не лише конкретні об'єкти з зображень, а близькі до них за значенням. Прикладом мережі класифікації зображень є Inception-ResNet-v2 – це згорткова нейронна мережа, яка натренована на більш ніж мільйоні зображень з бази даних ImageNet. Її архітектуру наведено на рис. 1.2. Мережа складається з 164 шарів і може класифікувати зображення за 1000 категоріями різних предметів та тварин.

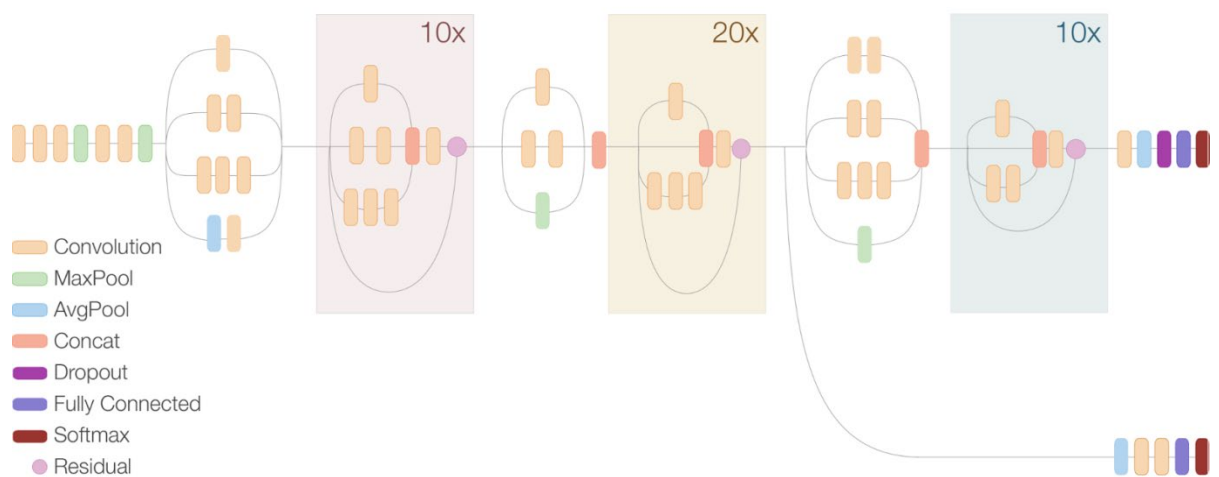


Рис. 1.2. Архітектура мережі для класифікації зображень

Мережа Inception-ResNet-v2 приймає на вхід зображення розміром 299 на 299 пікселів та визначає що зображено на ньому. Її глибока архітектура складається з більш ніж 30 згорткових шарів, що дозволяє показувати високу точність класифікації, але потребує значних затрат ресурсів на обробку одного зображення.

Окрім класифікації, нейронні мережі можуть виконувати сегментацію, тобто визначати розмір та положення об'єкта на зображенні. Результат сегментації також може бути корисним при побудові анотації. Створення нейронної мережі, яка класифікує абсолютно всі об'єкти які можуть бути зображені зараз неможливо. Тому можна використовувати декілька нейромереж, натренованих для класифікації різних типів об'єктів, а також експерт може

допомагати здійснювати класифікацію. Створення універсальної системи класифікації – це ще один напрям для розвитку.

Для пошуку мультимедійної інформації, за допомогою нейронних мереж, будують вектор ознак медіа файлу. Використовуючи мережу зі спеціально побудованою архітектурою, мультимедіа файл перетворюють в n -вимірний вектор, створюючи його пошуковий образ [33]. На рис. 1.3 наведено загальну архітектуру нейронної мережі перетворення зображення в вектор ознак. Перші шари таких мереж зазвичай є шарами згортки та агрегування, які потрібні для зменшення розмірності вхідного файлу. Після цих шарів розташовують повнозв'язний шар, вихідне значення елементів якого стає вектором ознак. Для покращення точності пошуку, між шарами згортки та повнозв'язним шаром доцільно розташовувати декілька прихованих шарів.

Побудований вектор ознак для мультимедіа файлу зберігає в собі особливості даних, наприклад, для зображення це різноманіття кольорів, пропорції фігур, текстура і т.д.

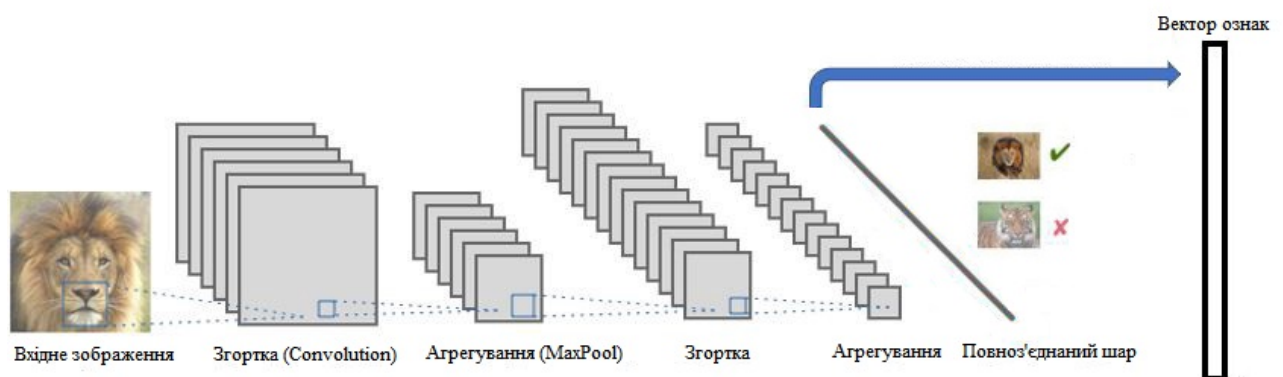


Рис. 1.3. Перетворення зображення в вектор ознак

Отриманий від користувача пошуковий запит також обробляється нейронною мережею та будується вектор його ознак. Пошуковий запит може бути як в текстовій формі, так і у вигляді відповідному формату шуканому медіа файлу – при пошуку зображення, порівнювати зі схожим зображенням; шукаючи аудіо файл, виконувати пошук за схожим відрізком звуку або схожою ритмічністю тощо.

Даний підхід дозволяє зменшити витрати часу на аналіз файлу, але він потребує налаштування архітектури мережі для досягнення найбільшої точності

пошуку. Обробка великих векторів нейронною мережею потребує значних затрат ресурсів, тому зменшення розмірів векторів має значення для підвищення ефективності роботи системи при використанні даного методу.

1.4 Постановка задачі

Проведений аналіз сучасних систем ідентифікації особи на зображенні показує, що наявні системи не здатні визначати особу на фотографії порівнюючи її з параметрами зовнішності, які зберігаються в базі даних. Також сучасні системи не здатні автоматично встановлювати можливу особу людини за словесним портретом.

Виходячи з цього, актуальною задачею є створення пошукової системи, яка зможе працювати з різнорідними базами даних, та виконувати ідентифікацію особи за фотографією обличчя або словесним портретом. Така система має приймати та обробляти в якості пошукових запитів зображення або текстовий опис зовнішності людини. Необхідно, щоб система мала змогу виконувати ідентифікацію людини в реальному часі на основі використання нейронних мереж.

Така система повинна вирішувати наступні задачі:

- ідентифікація особи на фотографії, порівнюючи його з наявними зображеннями в базі даних;
- ідентифікація особи на фотографії, у випадку наявності лише бази даних параметрів зовнішності людей;
- пошук людини за словесним портретом, якщо є база даних фотографій осіб;
- пошук людини за словесним портретом, якщо в наявності є лише база даних параметрів зовнішності людей.

Вирішення даних задач є необхідним в процесах ідентифікації особи, що робить розробку даної системи актуальною.

Для вирішення поставлених система повинна:

- виконувати порівняння людини на зображенні з іншими фотографіями людей, присутніми в базі, та визначати їх відповідність;
- виконувати пошук за словесним портретом особи, вираженим у текстовій формі, в базі даних параметрів зовнішності осіб;
- автоматично визначати параметри зовнішності людини на фотографії та порівнювати їх з наявними даними, в базі параметрів зовнішності для встановлення їх відповідності;
- автоматично визначати параметри зовнішності людини на фотографії та заносити їх до бази даних опису параметрів зовнішності, для подальшого можливого пошуку.

При цьому система пошуку повинна мати високі показники точності ідентифікації особи та швидкості обробки вхідних пошукових запитів для забезпечення можливості проведення оперативного визначення особи людини.

1.5 Висновки до розділу 1

В розділі було розглянуто задачі, в яких необхідно виконувати ідентифікацію особи та визначено основні з них. Основними сферами, в яких застосовують ідентифікацію людей, є: в системах відеоспостереження та моніторингу, в автоматизованих системах безпеки і контролю доступу, в медичній сфері.

Встановлено роль та місце процесу пошуку інформації в задачах ідентифікації людей. Пошук даних при визначенні особистості людини необхідний для швидкого доступу та отримання даних, які підходять за деякими критеріями.

Розглянуто процес пошуку інформації в задачах ідентифікації осіб та наведено основні етапи роботи системи пошуку. До них відносяться: етап індексації інформації та поміщення її в базу даних; етап пошуку та підбору даних за запитом і критеріями пошуку; етап сортування знайденої інформації та її повернення.

Визначено, що основними даними, якими оперують в процесі ідентифікації є мультимедійні файли та реляційні бази даних, тому система пошуку інформації повинна вміти аналізувати та обробляти різні типи файлів.

Виділено основні методи обробки та проведення пошуку мультимедійних файлів. При пошуку мультимедійної інформації в базі даних, можливо: створювати текстову анотацію файлів, використовувати класифікацію даних, для визначення їх можливого значення, та застосовувати нейронні мережі для побудови векторів ознак з метою їх подальшого порівняння.

Виявлено недолік сучасних систем ідентифікації людей на зображенні. Він полягає у відсутності можливості пошуку особи за фотографією або словесним портретом, порівнюючи їх зі збереженими значеннями параметрів зовнішності. Було зроблено висновок про необхідність розробки системи пошуку осіб, яка зможе виконувати ідентифікацію людей за фотографією обличчя або словесним портретом в реальному часі.

РОЗДІЛ 2

РОЗРОБКА ТЕХНОЛОГІЇ ПОШУКУ В НЕОДНОРІДНОМУ ІНФОРМАЦІЙНОМУ ПРОСТОРИ

2.1 Опис технології пошуку інформації

Система пошуку інформації, що розробляється в даній роботі, повинна вміти обробляти пошукові запити які можуть мати одну з двох форм: запит у формі зображення та текстовий запит на природній мові. Кожна форма запиту обробляється окремо. Схему обробки обох варіантів запиту системою наведено на рис. 2.1.

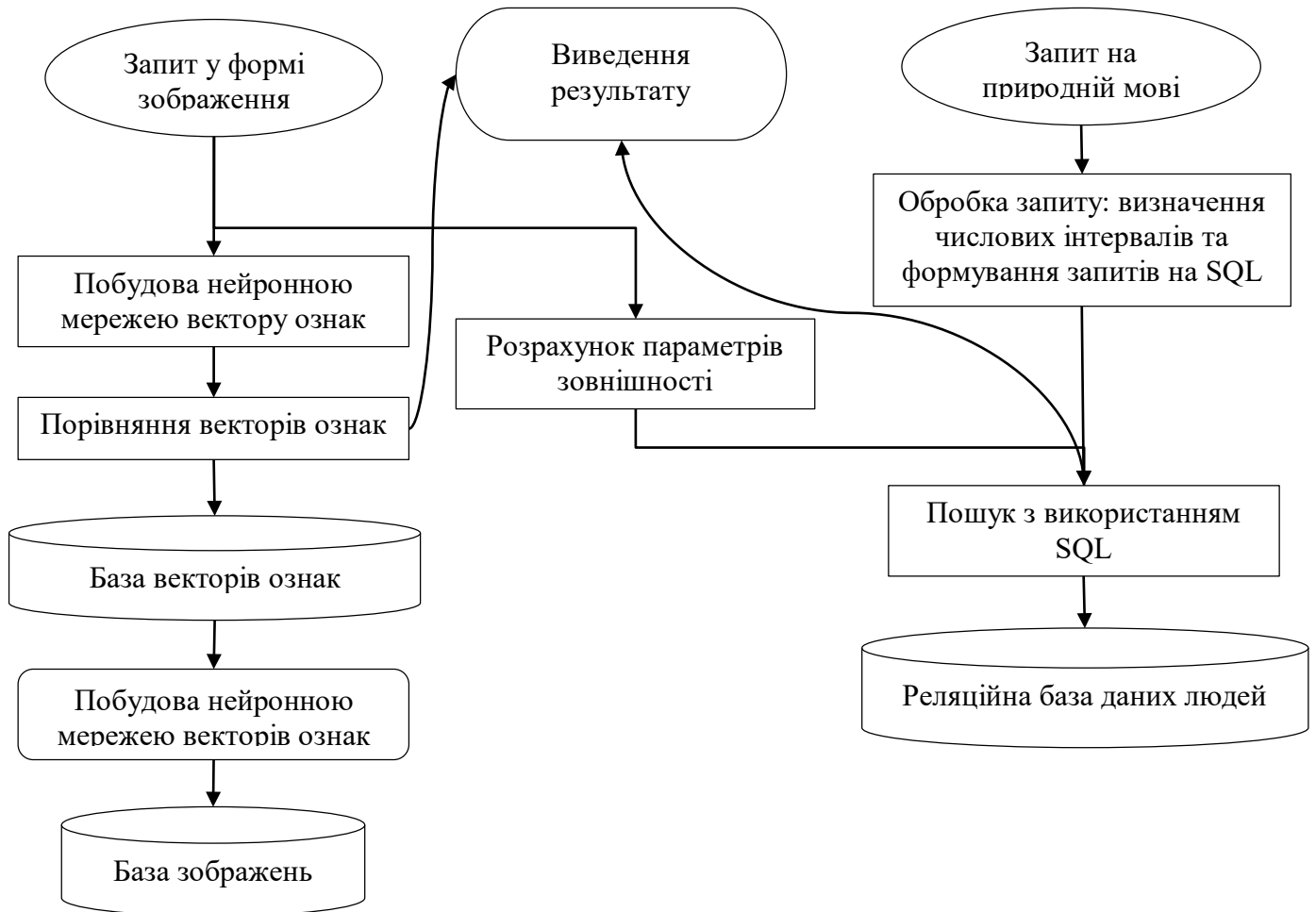


Рис 2.1. Схема обробки запитів системою

При отриманні пошукового запиту у форматі зображення, воно обробляється та виконується побудова його вектору ознак. Побудований вектор система по чергово порівнює зі всіма векторами з бази векторів ознак, використовуючи функцію подібності. Якщо в базі знайдено деяку кількість векторів, які знаходяться достатньо близько до пошукового на задане значення, система вважатиме людину на відповідних зображеннях ідентифікованою. У разі, якщо в базі зображень існують фотографії схожих за зовнішністю людей, то створені вектори ознак можуть одночасно бути достатньо близькими до цільового вектору із запиту. Тобто, на одне запит-зображення може бути знайдено декілька відповідних векторів з бази. В такому випадку, система порахує кількість схожих зображень кожної людини. Особа, більше векторів

зображень якої, відповідає вектору запиту, в результаті вважається зображеною на вхідній фотографії. Саме тому, необхідно додавати до бази зображень декілька екземплярів фото однієї людини, а загальна кількість внесених фотографій кожної особи має бути приблизно однаковою.

В результаті пошуку за запитом у формі зображення, система повертає ім'я особи, що присутня на ньому та саме зображення з позначеним розташуванням її обличчя і підписом її імені. У випадку, якщо на вхідному фото присутні декілька людей одразу, та система змогла ідентифікувати їх, буде повернуто список їх імен і, відповідно, зображення з відміченими положеннями їх розташування.

В разі необхідності пошуку за запитом сформованим на природній мові, система буде працювати зі реляційною базою даних, в якій зберігаються значення параметрів зовнішності людей. Отриманий пошуковий запит обробляється та визначається його зміст, а саме вилучаються параметри за якими відбувається пошук та їх значення. Вигляд пошукового запиту має визначену структуру. В разі введення запиту, сформованого в не відповідному вигляді, оброблятися він не буде та система повідомить про це користувача. Якщо в запиті будуть присутні слова, які не занесені до системи, то їх буде проігноровано.

Пошукові запити до реляційної бази даних будуть формуватись з використанням мови запитів SQL. Для отриманого запиту на природній мові, після аналізу системою, буде сформовано запит до бази даних на SQL. З отриманої відповіді, буде вилучатись ім'я людини. Таким чином, результатом пошуку за запитом на природній мові, будуть імена осіб, які підходять під задані параметри.

Третім можливим варіантом пошуку, є визначення записів всіх наявних людей в базі описів, які відповідають параметрам особи, що зображена на фотографії-запиті. Завдяки можливості автоматичного оцінювання пропорцій обличчя людини, система зможе за бажанням користувача аналізувати зображення-запит та виконувати пошук в записів в базі параметрів зовнішності.

В результаті такого пошуку, може бути виведено не лише ім'я конкретної особи, що зображена на фотографії, а всі люди, що мають схожі пропорції обличчя. Такий варіант пошуку не гарантує точне визначення особи на фото, але потрібен, якщо в базі зображень немає фотографій потрібної людини.

2.2 Розробка алгоритму ідентифікації особи на основі порівняння зображень

Однією з поставлених задач даної роботи є ідентифікація особи на зображенні, шляхом порівняння його з наявними фотографіями в базі даних. Для цього необхідно визначити алгоритм порівняння вхідного зображення з іншими зображеннями в базі.

При пошуку зображень в базі даних, як було визначено в першому розділі, розрізняють три основні підходи – побудова вектору ознак зображення, його класифікація або створення анотації. Оскільки метою даного дослідження є розробка системи пошуку, яка працює із зображеннями облич людей, тому буде застосовано підхід побудови вектору ознак за допомогою нейронних мереж. Ознаками в даному випадку будуть виступати параметри обличчя, такі як розміщення очей, носу, роту, щік та їх пропорції. Це дозволить приймати в якості пошукового запиту файл зображення, на якому присутнє обличчя потрібної людини та виконувати пошук схожих фотографій з подальшою ідентифікацією особи на фотографії.

Задачу ідентифікації особи на фотографії можна розбити на дві підзадачі – пошук обличчя людини на зображенні та побудова вектору ознак для знайденої області обличчя. Побудовані вектори різних облич порівнюються для визначення особистості зображеної людини.

Для пошуку обличчя на зображенні можна використовувати різні методи комп'ютерного зору, такі як використання гістограми напрямлених векторів (HOG, Histogram of Oriented Gradients), метод Віюли-Джонса [35] або нейронні мережі. Перші два методи мають високу швидкість обробки зображення, в

порівнянні з використанням нейронних мереж, але при цьому значно меншу точність пошуку необхідних об'єктів, через їх нестійкість до спотворень вхідних даних. Оскільки передбачається, що система переважно буде опрацьовувати зображення з камер відео нагляду, найбільш доцільними є прості методи пошуку об'єктів. Обличчя людей на зображеннях з таких камер зазвичай мають приблизно одне розташування та напрям погляду. В даній роботі буде використано гістограму напрямлених векторів для знаходження місця розташування обличчя на фотографії.

Краї та ребра об'єктів на зображеннях є найбільш часто використовуваними ознаками при пошуку та класифікації об'єктів. Основна ідея застосування гістограми напрямлених векторів полягає в тому, що краї мають найбільші значення градієнту на всьому зображенні, тому ознаки ребер та країв можуть бути представлені різними способами за допомогою градієнтів [34]. Розраховані градієнти поєднуються для побудови дескриптору гістограм напрямлених векторів (дескриптору HOG).

Використання гістограми напрямлених векторів, для пошуку обличчя на фотографії, має декілька ключових переваг над іншими методами. Її застосування призводить до інваріантності знаходження об'єкту при геометричних та фотометричних перетвореннях даних, оскільки вона обраховується на локальних значення комірок даних, так що ці спотворення будуть впливати на результат лише у більших областях простору. Крім того, глобальне розбиття простору та локальна фотометрична нормалізація дозволяє ігнорувати зміну положення людей на зображеннях, за умови, що вони мають вертикальне положення тіла. Таким чином використання HOG підходить для виявлення людей та обличч на зображеннях. В основі цього методу лежить припущення, що вигляд розподілу градієнтів інтенсивності зображення дозволяє досить точно визначити наявність і форму присутніх об'єктів.

Першим кроком при обрахунку дескрипторів HOG є обчислення градієнтів зображення. Спочатку зображення обробляється за допомогою одновимірних масок як у горизонтальному, так і у вертикальному напрямках, які називаються

ядрами (фільтрами) зі значеннями $[-1, 0, 1]$ і $[-1, 0, 1]^T$. Також є інші маски, такі як маски Собеля, Превітта або Канні, але зазвичай вони показують гірші результати при використанні. За результатом проходження ядер, будуються дві матриці D_x та D_y похідних уздовж осей x та y відповідно. Дані матриці використовуються для обчислення кутів і величин (модулів) градієнтів в кожній точці зображення.

Після розрахунку градієнтів, виконується побудова гістограми. На цьому етапі створюються гістограми для кожної комірки. Комірki це області пікселів, які мають прямокутну або радіальну форму. Крок гістограми рівномірно розширюється від 0° до 180° (або від 0° до 360°). Кожен піксель у комірці зважується та включається в один з 9 стовпців гістограми, до якого належить його орієнтація (похідна по y). Вагою пікселю може бути як сама величина його градієнту, так і деяка функція, наприклад, квадратний корінь.

Щоб отримати стійкість до різного освітлення та контрасту, значення градієнтів повинні бути нормалізовані локально. Для цього виконується групування комірок у більші піксельні області, які називаються блоками. Блоки накладаються на сусідні блоки, так що кожна комірка пікселів може бути врахована більше одного разу. Існує 4 вигляди геометрії блоків, причому найбільш часто використовується – прямокутна. Прямокутні блоки можуть мати розмір 2 на 2 комірки з 8 на 8 пікселів або 3 на 3 комірки з 6 на 6 пікселів.

Отримані блоки піддаються нормалізації. Існує три основні способи нормалізації блоків. Нехай h позначає ненормований блок, який складається зі всіх комірок гістограм з даного блоку. Тоді схеми нормалізації мають такі форми (L_2 або L_1 норма):

$$h_{L_2} = \frac{h}{\sqrt{|h|_2^2 + \varepsilon}} \quad (2.1)$$

$$h_{L_1} = \frac{h}{\sqrt{|h|_1 + \varepsilon}} \quad (2.2)$$

$$h_{\sqrt{L_1}} = \sqrt{h_{L_1}} \quad (2.3)$$

де $|h|$ – норма вектору,
 ε – деяка константа.

Кінцевим дескриптором ознак HOG є вектор, що містить елементи нормалізованих гістограм комірок з усіх областей блоку.

Для тренування, тобто побудови, HOG дескриптору, потрібно мати велику вибірку даних. Вибірка даних для навчання має складатись з зображень, на яких присутній об'єкт, який потрібно знаходити, і зображень, на яких його немає. Всі вхідні дані потрібно проаналізувати і побудувати гістограму значень. Результируюча гістограма складається зі значної кількості векторів – напрямлень градієнтів в кожній комірці зображення. Після чого, дані вектори потрібно поділити на два кластери. Найчастіше для поділу на кластери використовують метод опорних векторів (SVM, Support Vector Machines), так як він гарантує оптимальний розподіл векторів на два кластери. На виході SVM, між дома створеними кластерами утворюється границя, найчастіше лінійна. Два отримані кластери потім використовуються для прийняття рішення про наявність шуканого об'єкту на зображенні.

HOG дескриптори використовуються в поєднанні з різними алгоритмами машинного навчання для розпізнавання об'єктів, найчастіше у методі опорних векторів. Вони подаються на вхід алгоритмів як ознаки шуканого об'єкту. Однак дескриптори HOG не прив'язані до певного алгоритму машинного навчання.

Для знаходження обличч людей на зображенні методом HOG, потрібно побудувати дескриптори на вибірці зображень на яких є обличчя людей та на яких їх немає. Оскільки, під час проходження біля камери відео наглядю, людина не завжди направляє свій погляд в її бік, тому обличчя на зображенні може бути повернуте на деякий кут. Для подолання даної проблеми і підвищення надійності знаходження обличчя при такому спотворенні, під час використання даного методу пошуку, потрібно будувати декілька дескрипторів HOG.

Окрім пошуку обличчя людини, яка дивиться в напрямку камери, потрібно виконувати пошук обличчя, яке повернуте повністю вліво, повністю право,

трохи вліво та трохи вправо. Отже, потрібно мати п'ять вибірок даних для кожного з цих випадків та будувати п'ять НОГ дескрипторів. При обробці вхідного зображення, його потрібно аналізувати всіма створеними дескрипторами.

Аналіз зображення дескриптором НОГ дозволяє лише визначити чи присутній шуканий об'єкт на зображенні. Потрібно визначити область, яка потенційно містить цільовий об'єкт, щоб ідентифікація особи могла бути виконана більш ефективно, тому необхідно також знаходити положення об'єкта та розрахувати координати обмежуючого ящика – координати кутів квадрату в якому присутній шуканий об'єкт.

Визначення обмежуючого ящика шуканого об'єкту на зображенні виконується використовуючи різні методи. Такими методами є метод ковзного вікна або метод вибіркового пошуку (Selective Search) [36]. Ідея методу ковзного вікна полягає в тому, що вхідне зображення потрібно аналізувати “вікном” деякого початкового розміру та визначити присутність об'єкту в даному вікні. Після обробки даної області зображення, вікно зсувається на деяку фіксовану кількість пікселів, та аналізується детектором знову. При знаходженні об'єкту в деякому вікні, розмір вікна інкрементно зменшують для встановлення найбільш точного розташування об'єкту.

Більш ефективним для виявлення місця розташування об'єкту на зображенні вважається метод вибіркового пошуку (Selective Search) – алгоритм генерації запропонованих регіонів, які потенційно містять потрібні об'єкти. Він виконує аналіз вихідних даних сегментації зображення та використовує ознаки регіонів для ієрархічного групування знизу вгору. Попередню сегментацію зображення, для даного методу, зазвичай проводять застосовуючи метод Фельзеншвальба [37].

Використовуючи алгоритм вибіркового пошуку, спочатку, за допомогою методу Фельзеншвальба, зображення сегментується для створення початкових регіонів. Потім, використовуючи жадібний алгоритм, регіони ітеративно групуються разом: розраховується схожість між усіма сусідніми регіонами; два

найбільш подібні регіони об'єднуються, і обчислюються нові значення подібності між отриманим регіоном та його сусідами. Процес групування найбільш подібних регіонів повторюється, поки все зображення не стане єдиною областю.

Після генерації регіонів, дескриптори HOG можуть бути використані для виявлення обличчя людини на фото.

Побудова вектору ознак.

Після того як на зображенні знайдено обличчя людини, система повинна ідентифікувати особу. Задачу визначення людини за фотографією її обличчя, можна вирішити порівнявши вхідне зображення з фотографіями обличч інших людей в базі даних. Для порівняння зображень використовують багато алгоритмів, такі як попиксельне порівняння, порівняння ключових точок або побудова вектору ознак за допомогою нейронних мереж.

В порівнянні з іншими алгоритмами, найбільшою перевагою нейронних мереж в процесі пошуку зображень є їх інваріантність до спотворень вхідних даних, наприклад повороту зображення, різне освітлення або, в нашому випадку, незначний нахил або поворот обличчя людини на зображенні.

Для підвищення точності порівняння обличч на фото, потрібно обробляти не все зображення, а лише знайдену область, на якій присутнє обличчя. Знайдений регіон обличчя вилучається з загального фото та обробляється як окреме зображення, саме для нього будується вектор ознак.

Побудова вектору ознак обличчя, буде виконуватись з використанням згорткової нейронної мережі. Глибокі нейронні мережі дозволяють визначати низько-, середньо- та високорівневі ознаки використовуючи багат шарову архітектуру, поєднуючи шари один за одним, а збільшення кількості прихованих шарів може збагатити «рівні» ознак, тому кількість прихованих шарів має вирішальне значення. Але при збільшенні глибини мережі, точність її передбачення спочатку зростає, а потім швидко зменшується. Щоб вирішити цю проблему, в даній роботі буде застосовано архітектуру залишкових нейронних мереж (Residual Network).

Для побудови векторів ознак зображення буде використано залишкову нейронну мережу ResNet-34 [38], архітектуру якої наведено на рис. 2.2. Вона складається з 34 шарів, майже всі з них – згорткові (conv). На даному рисунку шари агрегації проігноровано, але вони присутні після кожного шару згортки. Кожен шар згортки має розмір фільтру, 3 x 3 або 7 x 7, тобто зображення обробляється квадратним фільтром розміром 3 на 3 та 7 на 7 пікселів відповідно. Ключовою особливістю залишкових нейронних мереж є зв'язки швидкого доступу, які пропускають декілька шарів та додають виходи пропущених шарів до шарів перед ними за формулою:

$$F(x) = f(x) + x \quad (2.4)$$

де x – вхідне значення, $f(x)$ – результат обчислення пропущених шарів, $F(x)$ – результат блоку.

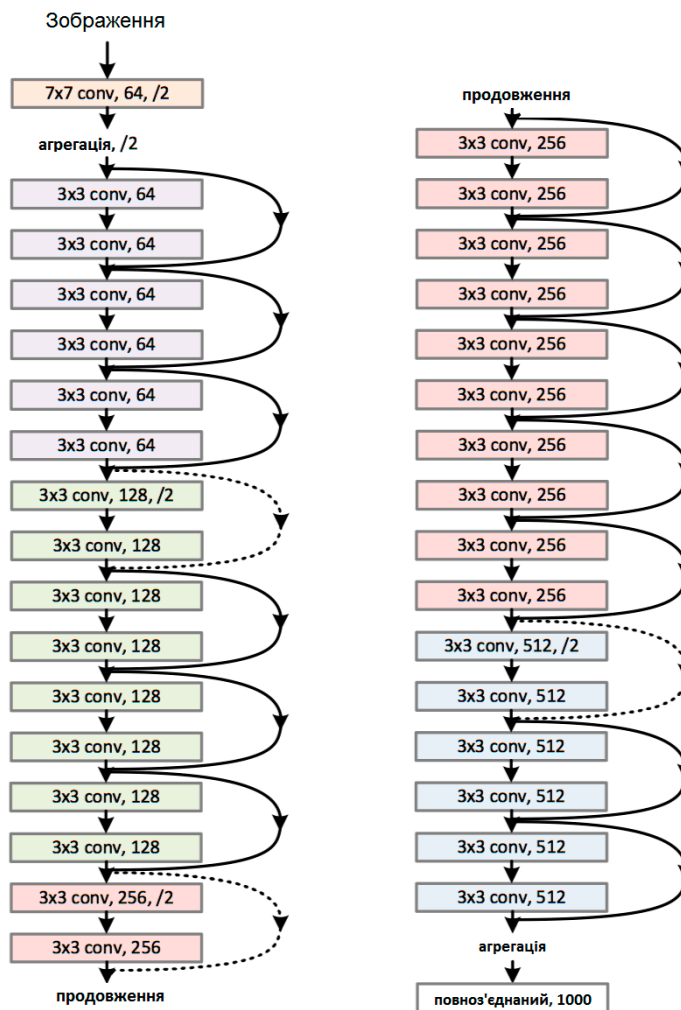


Рис. 2.2. Архітектура мережі ResNet-34

Існують інші варіації архітектури мережі ResNet-34, а саме ResNet-50 та ResNet-101. Вони мають більшу кількість шарів згортки (50 та 101 відповідно), що дозволяє даним архітектурам показувати ще більшу точність при побудові та порівнянні векторів ознак зображення. Але при цьому вони виконують значно більше операцій обчислення, що потребує часу та витрачає більше електроенергії під час роботи. Тому для нашої задачі найбільш оптимальною є саме обрана архітектура з 34-ма шарами.

На вхід даної мережі буде надходити лише вилучена з початкового зображення, область, яка містить обличчя людини, а результатом обчислення мережі буде вектор ознак.

Для оптимізації роботи системи, вимірність кожного шару в мережі буде зменшено на половину – 512 зменшується до розміру 256 елементів, 256 до 128 і т.д. Останній – повноз’єднаний шар – зменшено до розміру 128 елементів.

Останній шар мережі з’єднано з попереднім за схемою всі-до-всіх. Результиуючі значення всіх його елементів є побудованим вектором ознак. Чим менше розмірність даного шару, тим менше значень налічує вектор ознак, що зменшує час, затрачений на їх порівняння. При цьому зменшення розмірності вектору ознак негативно відображається на точності порівняння зображень, адже зменшується кількість описуваних ознак. Однак, оскільки людське обличчя має невелику кількість деталей, то зменшення розмірності елементів останнього шару до 128 дозволяє зберегти достатню кількість ознак зображення для досить точного порівняння обличчя людини. При великій кількості елементів вектору, яке описує обличчя, можливе зменшення точності порівняння через врахування вектором зайвих ознак зображення, такі як деталі освітлення, точне положення і форма деяких об’єктів, та інше.

Тренування нейронної мережі для побудови векторів ознак слід проводити таким чином, щоб для зображень з обличчям однієї людини побудовані вектори ознак були максимально близько, а для зображень іншої – якнайдалі. Тобто повинно виконуватись правило “максимізація різниці, мінімізація збіжностей”. Для досягнення такого результату тренування можна скористатись функцією втрати триплетів (Triplet loss).

Функція triplet loss задається на так званих триплетах – трійках виду (x_i^a, x_i^p, x_i^n) , де x_i^a - зображення конкретного класу (якір), x_i^p - зображення того ж класу (позитивний приклад), x_i^n - зображення класу, відмінного від класу зображень x_i^a та x_i^p (негативний приклад) [39, 40]. Елементи триплета подаються на вхід нейронної мережі $f(x)$, і від виходів обчислюється функція втрат:

$$L = \text{Softplus}(\|f(x_i^a) - f(x_i^p)\|_2 - \|f(x_i^a) - f(x_i^n)\|_2), \quad (2.5)$$

де $\text{Softplus}(x) = \ln(1 + e^x)$.

В результаті мінімізації такої функції втрат вихід нейронної мережі $f(x)$ набуває необхідні властивості: кожен якір $f(x_i^a)$ стає ближче до всіх позитивних прикладів $f(x_i^p)$, ніж до будь-якого з негативних $f(x_i^p)$.

Для визначення особи, яка зображена на фотографії, потрібно попередньо занести до бази зображень декілька різних фотографій з даною людиною, на яких вона буде розташована в різних умовах. Кожне зображення з бази необхідно проаналізувати та побудувати вектор ознак для нього. Чим більше різних зображень однієї людини буде додано до бази та проаналізовано, тим вище ймовірність правильної ідентифікації даної особи на інших зображеннях. При отриманні зображення, на якому потрібно ідентифікувати особу, для нього будується вектор ознак, який в подальшому по чергово порівнюється з іншими векторами ознак в базі. Порівняння векторів полягає в тому, щоб розрахувати кут між ними. Розрахунок кута між векторами A та B проводиться з використанням деякої функції подібності, в даному випадку буде застосовано косинус подібності, який розраховується за формулою:

$$\cos(\theta) = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}, \quad (2.6)$$

де A_i та B_i – координати векторів A та B відповідно.

Після розрахунку значення кута між векторами, воно порівнюється з деякою величиною – пороговим значенням. У разі, якщо кут між векторами менше порогового значення, вважається що відповідні зображення мають однаковий зміст, в нашому випадку є фотографіями однієї людини.

Таким чином, визначено алгоритм ідентифікації людини на зображенні з використанням нейронної мережі та виконанням побудови векторів ознак, його кроки проілюстровано на рис. 2.3.

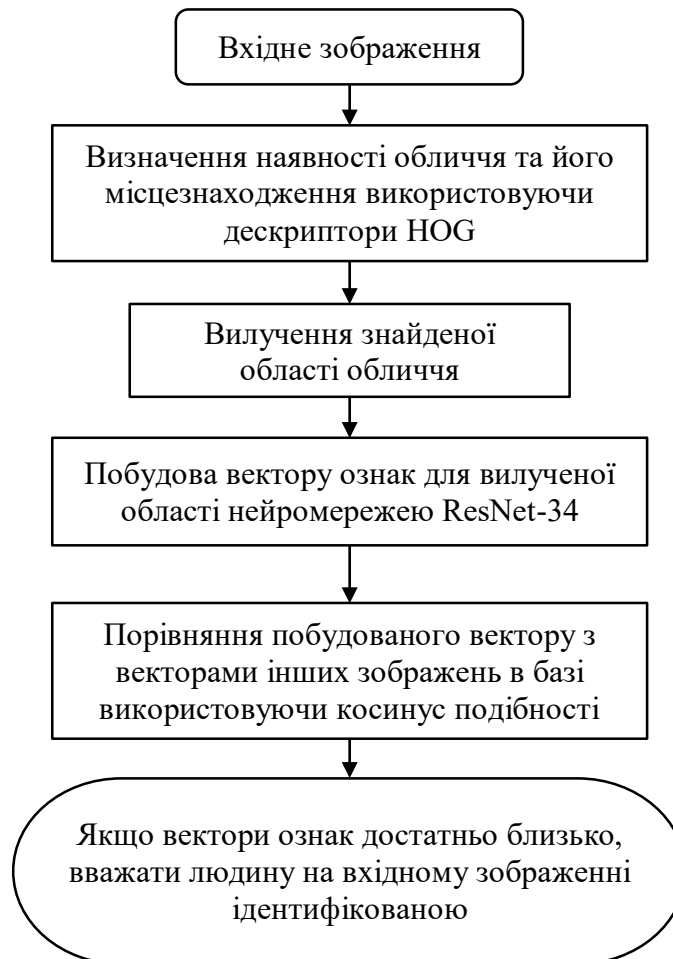


Рис. 2.3. Алгоритм ідентифікації людини на фото

До початку роботи системи, в базу даних будуть завантажені фотографії людей, яких необхідно ідентифікувати. Система проаналізує кожне зображення з набору та побудує для них вектори ознак. Спочатку зображення обробляється (з використанням дескрипторів HOG та методу ковзного вікна) для визначення місцезнаходження обличчя на фото. Після визначення регіону з людським обличчям, він вилучається та далі аналізується окремо. До вилученого регіону застосовується нейронна мережа ResNet-34, яка будує вектор його ознак. При отриманні деякого зображення в якості пошукового запиту, для нього будується

вектор ознак, який потім порівнюється з кожним вектором в базі даних та визначається їх просторова близькість. Порівняння векторів проводиться за допомогою функції косинусу подібності. Якщо вектори ознак розташовані один до одного на деяку порогову відстань, людина, зображена на фотографії якій відповідає вектор, вважається тією ж, що зображена на фотографії-запиті. У випадку, коли жоден з векторів в базі не є достатньо близьким до вектору вхідного зображення, людина на фотографії вважається не ідентифікованою, тобто зображень з нею немає в базі.

2.3 Розробка алгоритму ідентифікації особи за текстовим запитом

Під час роботи з системою пошуку інформації, користувачам потрібно мати можливість вводу пошукового запиту у форматі тексту. Такий пошук здійснюється задаючи одне або декілька ключових слів, які вводяться в поле пошуку. В даному випадку, такі ключові слова будуть характеризувати параметри, які описують людину: її стать, зовнішність – пропорції носу, губ, товщину брів та інше. Система, яка розробляється в даному дослідженні, має вміти аналізувати текстовий запит та виконувати пошук в базі даних, яка зберігає параметри зовнішності внесених людей.

Опис особи, в системах ідентифікації людей, організовується в структурованому вигляді та зберігається за допомогою реляційної моделі даних. База даних людей зазвичай налічує декілька таблиць, кожна з яких зберігає опис відповідної інформації про особу за деякою тематикою, наприклад опис зовнішніх параметрів, опис обличчя, контактна інформація та інше. Ці дані використовуються системою для під час обробки та встановлення особи. Підтримкою цілісності та пошуком записів в таких базах займається деяка система управління базою даних, яка, в разі необхідності, повертає знайдені записи у відповідь на вхідний запит, який надходить від системи ідентифікації.

При цьому інформація в реляційній базі даних може містити числа, тобто точні значення. Користувачу не зручно, а інколи складно, конкретно вказувати

точні дані. Натомість, людина здатна безпосередньо оцінювати параметри, оперуючи словами, наприклад такими термінами як “низька”, “середня”, “висока”, якщо мова йде про швидкість, або “мало”, “багато” – якщо оцінується об’єм. Щоб дати змогу користувачу виражати запит до числових даних у формі слів-оцінок, система має розуміти значення таких слів та бути здатною переводити їх в числові значення, які і будуть відшуковуватись в базі даних.

В даній роботі пропонується введення запитів на природній мові та їх переклад в числові інтервали на основі використання таблиці співвідношень.

Першим кроком обробки текстового запиту є визначення синтаксису введеного тексту. Користувач може ввести запит на пошук людини описуючи параметри її зовнішності, наприклад "високий ніс, широкі губи, чоловік". У користувача має бути змога задати складений запит, який описує одразу декілька параметрів. Загальний синтаксис запитів, який буде підтримувати дана система пошуку, має такий вигляд:

$$\text{ЗАПИТ} = \{ \{ \langle \text{словесна оцінка} \rangle \langle \text{параметр} \rangle \} \}$$

де

$$\text{словесна оцінка} = \{ \langle \text{високий} \mid \text{звичайний} \mid \text{низький} \rangle \}$$

$$\text{параметр} = \{ \langle \text{ніс} \mid \text{рот} \mid \text{лоб} \mid \text{брови} \mid \text{очі} \rangle \}$$

Складений запис може бути вираженням з деяким розділовим знаком між кожним описуваним параметром. В якості розділового знаку найзручніше використовувати символ коми “,”. Наприклад, запит для пошуку та визначення особи може бути вираженням наступним чином:

$$\text{ЗАПИТ} = \{ \text{високий лоб, звичайний ніс} \}$$

Після визначення параметру та його слова-оцінки, потрібно вирахувати числовий інтервал для нього. Для цього буде використана таблиця співставлень. Дана таблиця буде зберігати словесну змінну – параметр зовнішності, який

описує слово, її можливі словесні значення – значення словесної змінної в формі слова, множину інтервалів словесної змінної – значення всіх можливих словесних значень змінної та їх числові інтервали. Використовуючи таблицю співставлення, система буде визначати числові інтервали для кожного, веденого в запиті, параметру та формувати запит до бази даних описів на мові SQL. З отриманої від бази даних відповіді система буде вилучати ім'я людини та повертати його користувачу. За необхідності, можливо виводити всі збережені параметри зовнішності людини.

2.4 Розробка алгоритму автоматичного розрахунку параметрів зовнішності

Для забезпечення можливості пошуку в структурованій базі параметрів зовнішності за пошуковим запитом у вигляді фотографії, система має вміти самостійно розраховувати необхідні пропорції обличчя людини з вхідного зображення. Автоматичний розрахунок пропорцій обличчя на фотографії буде виконуватись з використанням нейронної мережі ViSeNet [41] та методу ансамблю дерев регресії [42].

В даному випадку, база описів зовнішності людей буде зберігати опис пропорцій обличчя кожної особи. Чим більше таких пропорцій розраховується та зберігається, тим більша варіативність параметрів запиту при пошуку. З'являється можливість сортувати результати пошуку по більшій кількості критеріїв. Також від кількості пропорцій, що враховуються при формуванні бази даних, залежить точність пошуку та визначення необхідної особи.

Виконання автоматичного розрахунку пропорцій знайденого обличчя на фотографії можливо лише за умови, що зображене обличчя знаходиться під достатньо прямим кутом до камери. Це важливо через можливу неадекватність результатів обрахунку за рахунок повороту обличчя. Тобто, якщо голова зображеної особи направлена вбік, то її обличчя буде знаходитись під

неправильною перспективою, а отже правильний прорахунок значень пропорцій зовнішності неможливий.

Пропонується алгоритм автоматичного розрахунку параметрів зовнішності людини, який складається з трьох етапів: визначення ключових точок обличчя та перевірка напрямку погляду людини; сегментація обличчя на зони (ніс, рот, очі, брови і т.д.); обрахунок параметрів зовнішності, використовуючи отриманий результат сегментації зображення обличчя.

На першому етапі, для перевірки напрямку погляду людини, необхідно визначити лише ключові точки на зображенні, а саме розташування очей, носу та країв обличчя. Для пошуку ключових точок обличчя буде використано метод ансамблю дерев регресії, який описано в [42].

Метод ансамблю дерев регресії дозволяє визначити ключові точки на зображенні обличчя людини. Він застосовує каскадні регресори [43] для визначення кожної необхідної точки обличчя. З його використанням буде визначено позиції 68 ключових точок обличчя людини, які наведено на рис. 2.4.

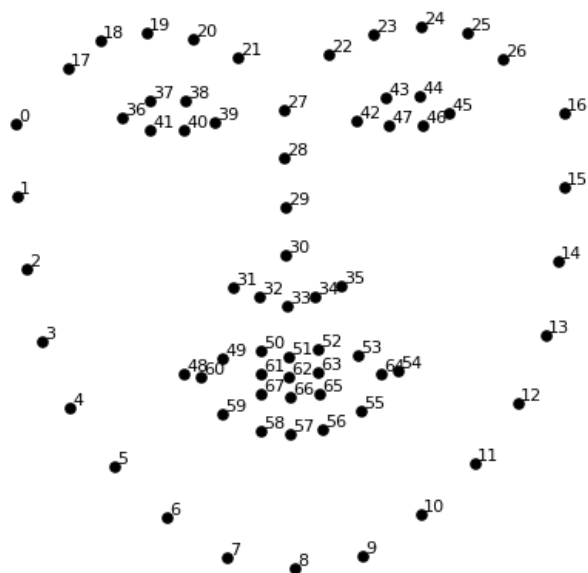


Рис. 2.4. Ключові точки обличчя людини

Використовуючи розраховані координати необхідних точок, можливо встановити напрямок погляду людини. Для цього необхідно використати

координати очей (точки 37 та 43), носу (точки 27 та 33) та країв обличчя (точки 0 та 16). Потрібно порівняти відстані між носом та краями обличчя. Якщо ліва або права частина обличчя менше іншої, на деяке порогове значення, а також верхня точка носу та нижня не знаходяться на одній прямій, то можна сказати, що зображене обличчя не направлене в бік камери.

Якщо визначено, що обличчя на зображенні повернуто до камери під достатнім кутом, треба виконати перевірку та поворот зображення, для того, щоб очі зображеної людини були розташовані на одному рівні. Це необхідно в подальшому, для зручного підрахунку пропорцій обличчя. Отримані координати розташування очей, можна використати для визначення чи знаходяться вони на одному рівні, порівнявши їх значення по вертикальній осі. У випадку, коли очі не розташовані на одній прямій, потрібно повернути все зображення. Напрямок та кут, на який потрібно повернути фотографію, розраховується використовуючи евклідову відстань. Після перевірки напрямку погляду людини на фото та вирівнювання обличчя, можна переходити на етап сегментації.

Другий етап визначення пропорцій обличчя – сегментація обличчя на зони. Сегментація обличчя на зображенні – це процес поділу фотографії людини на декілька сегментів (множину пікселів). Мета даної сегментації полягає у визначенні областей розташування частин обличчя. Точніше, сегментація обличчя – це процес присвоєння кожному пікселю зображення мітки назви частини обличчя. В даному випадку, цей етап використовує перевірене та вирівняне зображення обличчя людини, яке було отримано в результаті роботи попереднього етапу. Поділ зображення обличчя на зони необхідний для розрахунку пропорцій зовнішності людини, яке буде виконано на наступному кроці.

Сегментація обличчя на другому етапі розрахунку пропорцій зовнішності буде виконуватися з застосуванням нейронної мережі ViSeNet, яка дозволяє

швидко виконувати семантичну сегментацію в реальному часі. Семантична сегментація вимагає аналізу кожного пікселю зображення для визначення всієї просторової інформації. Однак сучасні підходи зазвичай ігнорують роздільну здатність зображення для досягнення швидкості виведення, що призводить до низької якості результату сегментації. Нейронна мережа BiSeNet створена для подолання цієї проблеми. Вона має двосторонню архітектуру (рис. 2.5), тобто складається з двох паралельних гілок, які одночасно обробляють зображення [41].

Під час обробки зображення нейромережею BiSeNet, воно проходить два шляхи. Спочатку, на “просторовому шляху”, фотографія обробляється загортковими шарами, щоб зберегти просторову інформацію та вилучити високорівневі ознаки з загальної області. В цей же час, зображення проходить “контекстний шлях”, який слугує для швидкого зменшення розмірності фотографії та поєднання ознак. Крім двох гілок, мережа використовує модуль поєднання ознак та модуль уточнення уваги. Модуль поєднання ознак ефективно виконує групування вилучених особливостей. Модуль уточнення уваги деталізує знайдені на кожному кроці ознаки. Особливістю модулю уточнення уваги є використання шару агрегації, який вилучає особливості зі всього зображення глобально, що дозволяє уникнути подальших операцій збільшення розмірності.

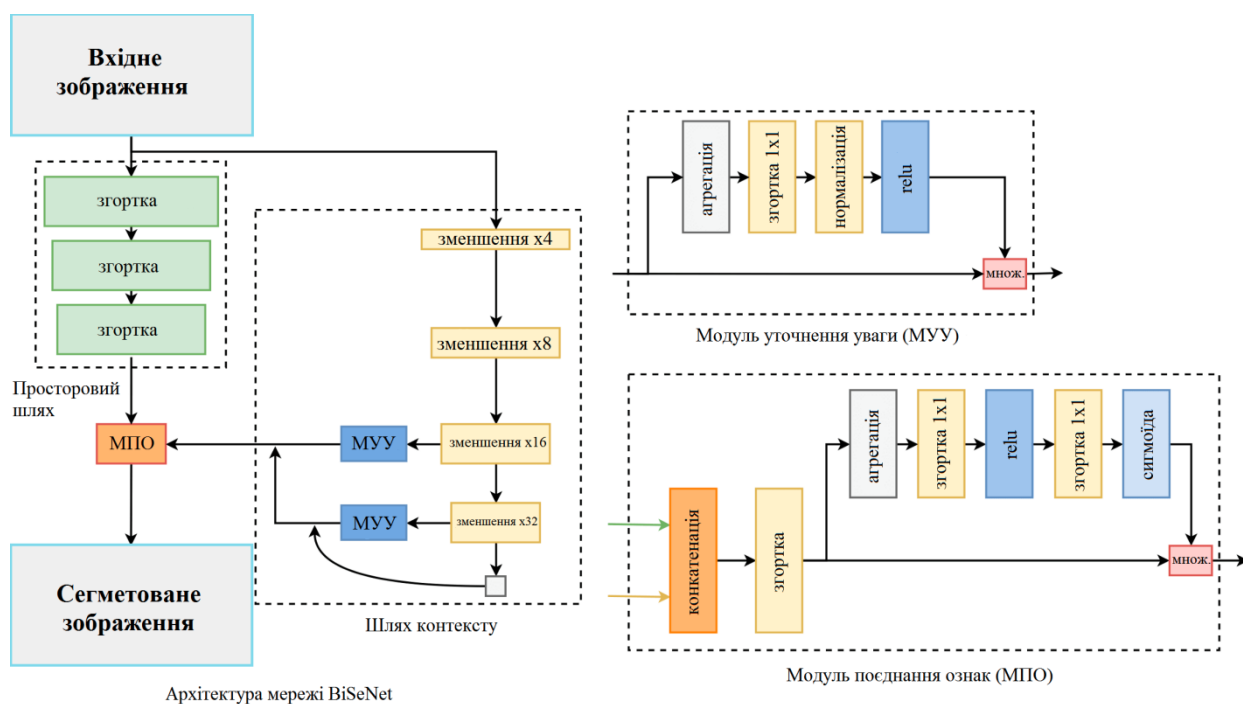


Рис. 2.5. Архітектура мережі BiSeNet

На вхід нейронної мережі BiSeNet подається фотографія, на якій присутнє обличчя людини. Результатом роботи мережі є сегментоване зображення, на якому визначено 16 зон, серед яких задній фон, шкіра, ліва та права брова, ліве та праве око, окуляри, ліве та праве вухо, ніс, верхня губа, нижня губа, шия, волосся, одяг. Приклад процесу пошуку обличчя, вирівнювання та сегментації зображення людини наведено на рис. 2.6.

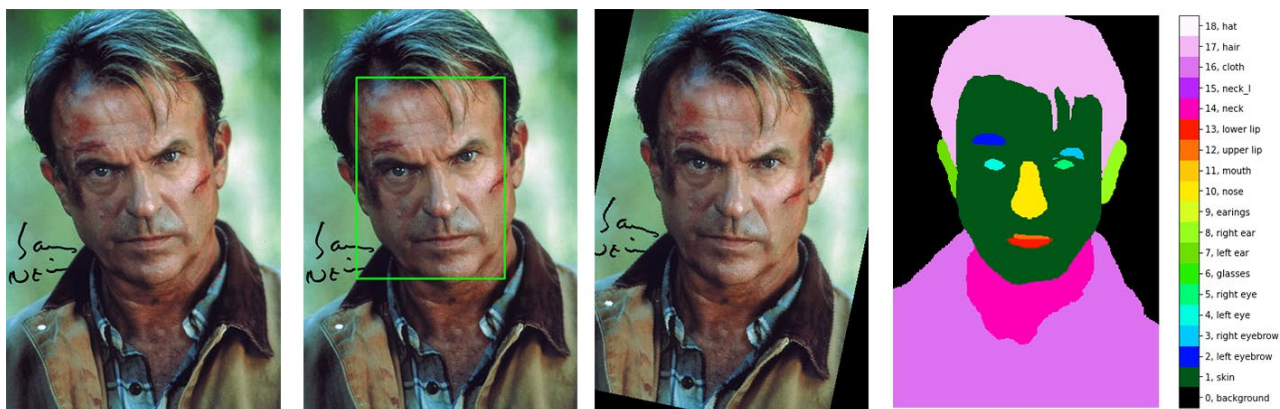


Рис. 2.6. Приклад сегментації обличчя

Отримавши сегментоване зображення, можна переходити до останнього (заключного) етапу визначення параметрів зовнішності – розрахунку пропорцій обличчя. Для цього потрібно на результуючій сегментованій фотографії особи визначати координати: найнижчу та найвищу точки обличчя; правий та лівий край роту; верхню та нижню точки губ; праву, ліву, верхню та нижню точки носу; верхню та нижню точки брів; а також ліві та праві точки кожного ока. Потрібні точки вираховуються аналізом кожної потрібної області по горизонтальній та вертикальній осях. Отримані координати точок дозволяють простими операціями визначити піксельні розміри кожної необхідної зони обличчя. Оскільки пропорції зовнішності розраховуються відношеннями розмірів різних частин обличчя, то не потрібно вираховувати значення розмірів в міліметрах. Визначені розміри областей обличчя в пікселях, використовуються для розрахунку пропорцій зовнішності.

В даному випадку будуть розраховуватись сім параметрів обличчя людини, які наведено нижче.

- Висота лобу. Розраховується як відношення відстаней $\frac{\text{ліва або права брова} - \text{верхня точка обличчя}}{\text{нижня точка обличчя} - \text{верхня точка обличчя}}$. Звичайним вважається значення $1/3 = 0,3$.
Значення словесних оцінок: низький – від 0,1 до 0,25; звичайний – від 0,25 до 0,35; високий – від 0,35 і більше.

- Висота носу. Розраховується як відношення відстаней $\frac{\text{нижня точка носу} - \text{верхня точка носу}}{\text{нижня точка обличчя} - \text{верхня точка обличчя}}$. Звичайним вважається значення $1/3 = 0,3$.
Значення словесних оцінок: низький – від 0,1 до 0,2; звичайний – від 0,2 до 0,35; високий – від 0,35 і більше.

- Ширина носу. Розраховується як відношення відстаней $\frac{\text{права точка носу} - \text{ліва точка носу}}{\text{права точка обличчя} - \text{ліва точка обличчя}}$. Звичайним вважається значення $1/5 = 0,2$.
Значення словесних оцінок: вузький – від 0,1 до 0,15; звичайний – від 0,15 до 0,25; широкий – від 0,25 і більше.

- **Ширина роту.** Розраховується як відношення відстаней $\frac{\text{права точка роту} - \text{ліва точка роту}}{\text{середня точка правого ока} - \text{середня точка лівого ока}}$. Звичайним вважається значення 0,85.

Значення словесних оцінок: вузький – від 0,7 до 0,8; звичайний – від 0,8 до 0,9; широкий – від 0,9 і більше.

- **Товщина губ.** Розраховується як відношення відстаней $\frac{\text{нижня точка роту} - \text{верхня точка роту}}{\text{нижня точка обличчя} - \text{нижня точка носу}}$. Звичайним вважається значення 0,3. Значення

словесних оцінок: тонкі – від 0,15 до 0,25; звичайні – від 0,25 до 0,35; товсті – від 0,35 і більше.

- **Близькість очей.** Розраховується як відношення відстаней $\frac{\text{ліва точка правого ока} - \text{права точка лівого ока}}{\text{права точка правого ока} - \text{ліва точка лівого ока}}$. Звичайним вважається значення 0,5.

Значення словесних оцінок: близько – від 0,3 до 0,4; звичайні – від 0,4 до 0,6; далеко – від 0,6 і більше.

- **Товщина брів.** Розраховується як відношення відстаней $\frac{\text{нижня точка брови} - \text{верхня точка брови}}{\text{нижня точка брови} - \text{верхня точка обличчя}}$. Звичайним вважається значення 0,2. Значення

словесних оцінок: тонкі – від 0,05 до 0,15; звичайні – від 0,15 до 0,25; товсті – від 0,25 і більше.

Розраховані характеристики зовнішності людини в подальшому використовуються для виконання пошуку в базі даних описів.

2.5 Висновки до розділу 2

Описано технологію пошуку інформації. При отриманні запиту у формі зображення, для нього будується вектор ознак та порівнюється з векторами з бази. Якщо користувач бажає виконати пошук зображеної особи в базі даних параметрів, для отриманої фотографії автоматично розраховуються пропорції зовнішності. Розраховані пропорції використовуються для пошуку записів в базі

параметрів. У разі надходження текстового запиту, його синтаксис аналізується та визначаються параметри обличчя і їх числові інтервали. Використовуючи числові інтервали параметрів з запиту, формується запит до бази пропорцій зовнішності.

Визначено алгоритм ідентифікації особи на основі порівняння зображень. Першим кроком проведення ідентифікації людини на фотографії є визначення розташування обличчя, використовуючи дескриптори HOG. Знайдена область обличчя вилучається з зображення та виконується побудова вектору ознак. Для побудови вектору ознак застосовується нейронна мережа Resnet-34 зі зменшеною розмірністю шарів. Отримані вектори ознак порівнюються з використанням функції косинусу подібності та встановлюється їх відповідність. Якщо кут між векторами ознак для вхідної фотографії та зображення з бази менше порогового значення, людина на зображенні вважається ідентифікованою.

Визначено синтаксис текстових запитів та спосіб їх обробки. Запит формулюється на природній мові та описує параметри зовнішності людини. В отриманому тексті запиту визначаються параметри та їх словесні оцінки. Для визначених параметрів за допомогою таблиці співвідношень проводиться розрахунок числових інтервалів та формується запит до бази даних.

Запропоновано алгоритм автоматичного розрахунку параметрів зовнішності людини. Він складається з трьох етапів: перевірка напрямку погляду людини; сегментація обличчя; обрахунок параметрів зовнішності. Для перевірки напрямку погляду, знаходяться контрольні точки носу і очей та перевіряється їх розташування на одній прямій. Сегментація обличчя виконується з використанням мережі ViSeNet. Отриманий результат сегментації обробляється та розраховуються піксельні розміри частин обличчя.

РОЗДІЛ 3

РОЗРОБКА СИСТЕМИ ПОШУКУ ІНФОРМАЦІЇ

3.1 Розробка архітектури системи пошуку інформації

Ефективність роботи системи підвищується, якщо її внутрішня структура добре організована та поділена елементи. Кожен з елементів відповідає за виконання конкретного завдання та взаємодіє з іншими, для досягнення спільної мети роботи всієї системи. В даному випадку систему поділено на три елементи (рис.3.1): модуль аналізу та обробки зображень; модуль обробки текстового запиту; загальний модуль системи, який буде відповідати на запити користувача. Останній зазначений модуль слугує для організації взаємодії всієї системи.

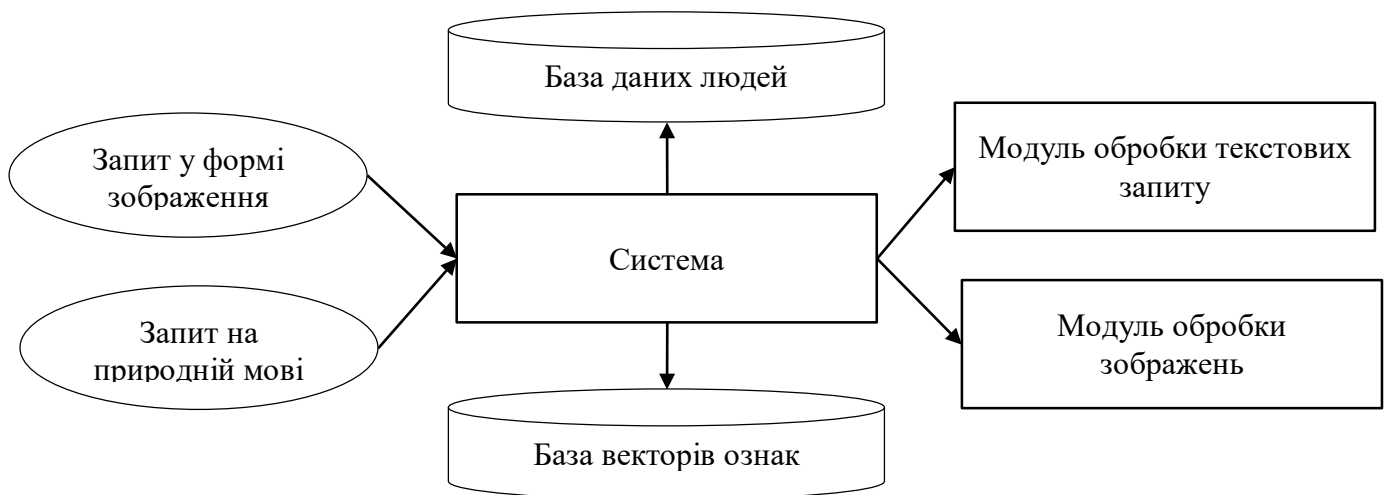


Рис. 3.1. Організація елементів системи пошуку інформації

Модуль аналізу та обробки зображень відповідає за всі операції обробки зображень, які передбачені в даній системі. Він виконує завдання вирівнювання зображення, визначення місця розташування людського обличчя, побудову вектору ознак для вхідного зображення, визначає контрольні точки обличчя людини та автоматично розраховує пропорції зовнішності зображеної особи.

Модуль обробки текстових запитів аналізує отриманий від користувача запит на природній мові, перевіряє його відповідність заданій структурі запиту

та буде відповідний запит на мові SQL для пошуку в реляційній базі даних. Отриманий запит розбивається на частини – оцінки параметрів зовнішності. Кожний зазначений параметр та його оцінка аналізується: за допомогою таблиці співставлення виконується визначення числового інтервалу для заданого слова-оцінки деякого параметру. Отримані числові інтервали підставляються в запит на мові SQL за відповідним синтаксисом. Оскільки даний модуль відповідає за переведення слів-оцінок в числові інтервали на основі використання таблиці співставлення, він також займається визначенням числових інтервалів для конкретного числа. Це необхідно для формування запиту до бази даних, після автоматичного обрахунку пропорцій обличчя, якщо користувач бажає виконати пошук людини на зображенні в базі параметрів зовнішності.

Модуль “система”, як вже було зазначено, виконує обробку запитів користувача та повертає результат пошуку. Для виконання свого завдання даний модуль використовує два інших модулі, взаємодіючи з ними шляхом виклику відповідних функцій.

При отриманні текстового запиту, модуль “система” звертається до модулю обробки текстового запиту та отримує від нього у відповідь сформований на SQL запит до бази даних. Сформований запит надсилається до системи управління базою даних. З отриманої від СУБД відповіді, модуль вилучає ім'я знайденої особи та повертає його користувачу в якості результату пошуку.

У випадку отримання від користувача, в якості запиту, файлу зображення, модуль “система” передає його модулю аналізу та обробки зображень. Якщо користувачу потрібно провести пошук фотографії по базі зображень, модуль обробки зображень знаходить на картинці обличчя та лише виконує побудову вектору ознак для нього. Побудований вектор ознак повертається модулю “система”, яка порівнює його зі всіма векторами в базі векторів. В разі виявлення достатньо близького вектору ознак, система повертає пов'язане з ним ім'я людини користувачу.

За потреби пошуку особи на зображенні в базі параметрів зовнішності, модуль обробки зображень проводить вирівнювання знайденого обличчя, його сегментацію та розраховує параметри зовнішності зображеної людини. Для розрахованих параметрів зовнішності, за допомогою таблиці співставлення, визначається інтервал пошуку, який підставляється в запит до бази пропорцій зовнішності, що формується на мові SQL. У відповідь на запит, система отримує від бази даних всі записи людей, котрі приблизно підходять під задані пропорції. Знайдені записи людей повертаються користувачу.

При реалізації системи засобами мови програмування, необхідно створити відповідні класи елементів. Схему організації класів системи наведено на рис. 3.2.

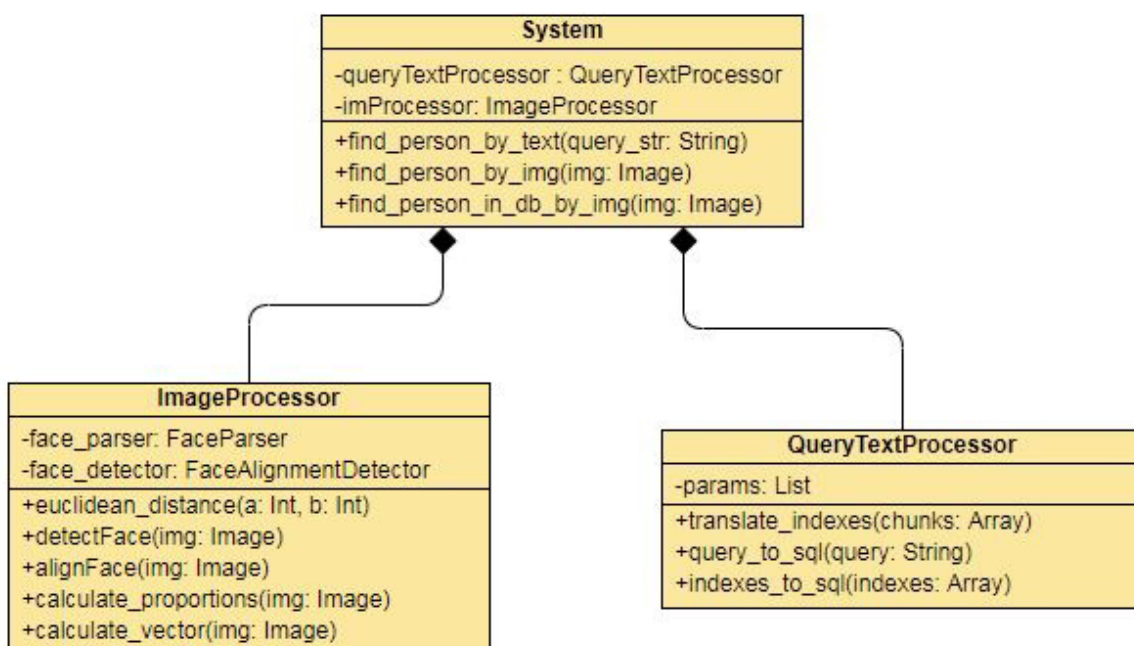


Рис. 3.2. Схеми організації класів системи пошуку

Розроблена система складається з трьох основних класів, які взаємодіють для досягнення мети роботи всієї системи пошуку інформації та залежать один

від одного. Елементами структури системи є класи “System”, “ImageProcessor” та “QueryTextProcessor”.

Клас “System” відповідає за обробку вхідних запитів користувача та повертає результат пошуку. Він створює екземпляри двох інших класів, як свої поля, та використовує кожен з них для обробки відповідного запиту, в залежності від формату пошукового запиту. В ньому присутні три методи:

- `find_person_by_text` – метод, що використовується, коли користувачу необхідно виконати пошук особи в базі описів за текстом запитом. На вхід приймає текстовий рядок – запит, та повертає результат пошуку. Використовує клас “QueryTextProcessor” для формування запиту до бази даних.

- `find_person_by_img` – виконує пошук та ідентифікацію людини, зображеної на фотографії. В якості параметру, отримує від користувача зображення, на якому потрібно визначити особу та. В результаті виклику даного методу, повертається ім’я, ідентифікованої на зображенні, особи та вхідне фото з відміченим обличчям. Використовує клас “ImageProcessor ” для обробки вхідного зображення та побудови вектору ознак.

- `find_person_in_db_by_img` – відповідає за пошук людини на зображенні в базі даних описів. Як і попередній метод, приймає фотографію з людиною, але повертає результат пошуку людей в реляційній базі даних за розрахованими оцінками. Додатково виводить числові значення оцінок параметрів зображеної особи, оцінки параметрів у словесній формі та сегментоване вхідне зображення. Ці дані необхідні для того, щоб пересвідчитись, що подане зображення підходить для коректної обробки та адекватного розрахунку оцінок параметрів. Використовує клас “ImageProcessor ” для обробки вхідного зображення та розрахунку числових значень параметрів зовнішності людини на вхідному зображенні, після чого передає отримані значення параметрів класу “QueryTextProcessor” для формування запиту до бази даних на мові SQL.

Клас “ImageProcessor” використовується класом “System” для обробки зображень, що подаються користувачем. Він використовує дескриптори HOG під

час пошуку обличчя на зображенні, нейронну мережу ResNet-34 для перетворення зображень в вектор ознак та нейромережу BiSeNet для сегментації зображень знайдених облич. Даний клас має такі методи:

- `detectFace` – метод, який виконує знаходження обличчя на вхідному зображенні використовуючи дескриптори HOG. На вхід отримує зображення на якому потрібно знайти обличчя людини та повертає координати ящика обмежень, в якому воно розташоване.

- `alignFace` – метод, необхідний для аналізу та повороту зображення обличчя людини. Приймає вхідне зображення з обличчям людини, визначає чи необхідно виконувати поворот та, за необхідності, повертає зображення. Як результат роботи, повертає виправлене зображення обличчя людини.

- `euclidean_distance` – функція, що використовується методом `alignFace` для визначення куту повороту, якщо необхідно вирівняти обличчя людини на вхідному зображенні.

- `calculate_vector` – метод, який викликається класом “System”. Використовуючи нейронну мережу ResNet-34, обраховує вектор ознак для вхідного зображення людини. Попередньо викликає метод `detectFace` для вилучення з зображення лише області розташування обличчя. Повертає результатом роботи розрахований вектор ознак. Якщо обличчя не було виявлено на вхідному зображенні, повідомляє про це.

- `calculate_proportions` – метод, який також використовується класом “System”. Необхідний для розрахунку параметрів зовнішності людини. Під час своєї роботи, викликає метод `detectFace` для визначення розташування обличчя на фото. Визначає, застосовуючи метод ансамблю дерев регресії, ключові точки обличчя та робить висновок наскільки обличчя на зображенні повернуте в напрямку камери. Якщо обличчя знаходиться під недостатньо прямим кутом до камери, метод закінчує свою роботу. Використовує нейронну мережу BiSeNet для сегментації отриманого обличчя. Викликає функцію `alignFace`, якщо обличчя на фотографії повернуте в бік камери, та передає їй сегментоване зображення. Після знаходження та повороту обличчя, розраховує значення семи його

параметрів (висота лобу, висота носу, ширина носу, ширина роту, товщина губ, близькість очей та густота брів). В результаті роботи даного методу, повертається список параметрів обличчя та їх числові значення.

Клас “QueryTextProcessor” використовується класом “System” для обробки текстових запитів користувача та формування запиту до реляційної бази даних на мові SQL. Даний клас використовує таблицю співставлення для визначення числових інтервалів, зазначених в текстовому запиті в словесній формі, параметрів оцінок зовнішності людини. Він зберігає слова-оцінки пропорцій обличчя в полі `params`. До нього належать такі методи:

- `translate_indexes` – метод, який виконує розрахунок числових інтервалів для поданих частин текстового запиту, що виражають параметри зовнішності особи та їх словесні оцінки. На вхід отримує масив текстових рядків. Використовує набір ключових слів, що зберігаються в полі `params`, для аналізу кожного елементу масиву та, за допомогою таблиці співставлення, розраховує числовий інтервал для кожного зазначеного слова-параметру та його слова-оцінки. В результаті своєї роботи повертає масив назв параметрів, з тією назвою, з якою вони зберігаються в базі описів зовнішності, та їх числові інтервали.

- `query_to_sql` – метод, який викликається класом “System”, коли необхідно проаналізувати введений користувачем текстовий запит та побудувати для нього відповідний запит до бази даних описів зовнішності на мові SQL. На вхід отримує рядок запиту, який був введений користувачем в формі “<оцінка параметру | параметр>, <оцінка параметру | параметр>...”. Виконує аналіз текстового запиту на відповідність формату. В разі, якщо параметр або його оцінка, відсутні в словнику визначених слів, або допущена друкарська помилка, відповідна частина запиту буде проігнорована. Даний метод використовує функцію `translate_indexes` для аналізу введених описових параметрів та визначення їх числових інтервалів. Виконує побудову запиту до бази описів зовнішності людей на мові SQL, використовуючи отримані числові інтервали, які підставляються в рядок запиту, при його її формуванні.

- `indexes_to_sql` – метод, який використовується при пошуку в базі даних описів зовнішності за запитом у вигляді зображення. Метод викликається класом “System” після того, як попередньо було розраховано параметри обличчя людини методом `calculate_proportions` класу `ImageProcessor`. Він виконує визначення слова оцінки для кожного числового значення параметру зовнішності та вираховує його інтервал. В якості вхідного параметру приймає масив параметрів зовнішності та їх значень в форматі {назва параметру: числове значення}. Після обробки кожного параметру та визначення його інтервалу, даний метод виконує побудову запиту до бази описів зовнішності людей на мові SQL з відповідними назвами параметрів та значеннями їх інтервалів. У результаті роботи даного методу, клас “System” отримує сформований запит до бази даних на мові SQL та надсилає його.

3.2 Визначення структури баз даних

Перед початком роботи, до бази даних зображень системи потрібно додати, фотографії людей, яких планується розпізнавати в подальшому. Бажано мати по декілька фотографій кожної людини, на яких вона знаходиться в різних фізичних умовах та має різний кут повороту до камери. Чим більше фотографій однієї особи наявно в базі даних, тим більш вірогідна правильна ідентифікація даної особи в подальшому. Всі зображення однієї особи мають знаходитися в одному місці – директорії. Відповідно, для кожної людини повинна бути створена окрема директорія з її фотографіями.

До початку виконання процесу пошуку, в офлайн режимі, всі зображення в базі даних обробляються – для кожного з них, за допомогою нейронної мережі, будується вектор ознак. Створенні вектори ознак додаються в окремо створену базу даних векторів, яка має структуру: <повне ім'я людини, побудований вектор ознак окремого зображення з нею>. Оскільки, для кожної особи в базі даних зображень присутні одразу декілька фотографій з нею, тому в базі векторів буде

декілька входжень записів для однієї людини. Кількість записів однієї особи відповідає кількості фотографій з нею в базі даних зображень.

Попередньо необхідно заповнити базу даних описів людей. В даному випадку база буде складатися з однієї таблиці, яка зберігає описи пропорцій обличчя внесених людей. Структуру таблиці наведено на рис. 3.3.

#	Name	Datatype	Length/Set	Unsigned	Allow NU...
1	ID	INT	11	<input type="checkbox"/>	<input checked="" type="checkbox"/>
2	FullName	VARCHAR	120	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3	state	INT	11	<input type="checkbox"/>	<input type="checkbox"/>
4	control_image_path	VARCHAR	256	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
5	images_path	VARCHAR	256	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
6	forehead_index	FLOAT		<input type="checkbox"/>	<input checked="" type="checkbox"/>
7	nose_height_index	FLOAT		<input type="checkbox"/>	<input checked="" type="checkbox"/>
8	nose_width_index	FLOAT		<input type="checkbox"/>	<input checked="" type="checkbox"/>
9	mouth_width_index	FLOAT		<input type="checkbox"/>	<input checked="" type="checkbox"/>
10	lips_thickness_index	FLOAT		<input type="checkbox"/>	<input checked="" type="checkbox"/>
11	eyes_closeness_index	FLOAT		<input type="checkbox"/>	<input checked="" type="checkbox"/>
12	eyebrows_thickness_index	FLOAT		<input type="checkbox"/>	<input checked="" type="checkbox"/>

Рис. 3.3. Структура таблиці параметрів зовнішності

Ця таблиця має такі поля:

- “ID” – ціле число, необхідне для гарантії унікальності кожного запису. Присвоюється інкрементивно, при додаванні нового запису в таблицю. Не може мати пустих значень.
- “FullName” – рядок тексту розміром до 120 символів, зберігає повне ім’я даної особи та використовується при виведенні результатів пошуку. Не може мати пустих значень.
- “state” – ціле число, яке позначає стать особи: “1” – чоловік, “2” – жінка. Не може мати пустих значень.
- “control_image_path” – рядок тексту розміром до 256 символів, зберігає шлях до контрольного зображення даної особи. Дане зображення може

бути використане для аналізу пропорцій та виведене в якості результату пошуку. Може бути не заповненим.

- “images_path” – рядок тексту розміром до 256 символів, зберігає шлях до папки з набором зображень даної особи. Виведення інших зображень людини користувачу може допомогти визначити правильність результату пошуку. Може бути не заповненим.

- “forehead_index” – число з комою, зберігає значення розміру лобу. Може бути не заповненим.

- “nose_height_index” – число з комою, значення висоти носу. Може бути не заповненим.

- “nose_width_index” – число з комою, значення ширини носу. Може бути не заповненим.

- “mouth_width_index” – число з комою, значення ширини роту. Може бути не заповненим.

- “lips_thickness_index” – число з комою, значення товщини губ. Може бути не заповненим.

- “eyes_closeness_index” – число з комою, значення близькості очей. Може бути не заповненим.

- “eyebrows_thickness_index” – число з комою, значення густоти брів. Може бути не заповненим.

За необхідності, до даної таблиці можливо додавати нові поля для збереження опису інших параметрів обличчя людини.

3.3 Висновки до розділу 3

Розроблено систему пошуку інформації, яка складається з трьох основних елементів – модуль “система”, модуль аналізу та обробки зображень та модуль обробки текстових запитів. Модуль обробки зображень виконує пошук облич на вхідних фотографіях, побудову векторів ознак для зображень та автоматичний розрахунок параметрів зовнішності. Модуль обробки текстових запитів аналізує синтаксис вхідних запитів та виконує побудову запиту до бази даних. Модуль “система” отримує і відповідає на пошукові запити користувача, відправляє запити до бази даних параметрів зовнішності та відповідає за порівняння побудованих векторів ознак для зображень.

Наведено структуру організації системи. Вона складається з трьох класів – “System”, “ImageProcessor” та “QueryTextProcessor”. Для кожного класу пояснено його основні функції роботи.

Сформовано структуру бази параметрів зовнішності та бази векторів. Бази параметрів зовнішності складається з однієї таблиці, яка зберігає значення пропорцій обличчя, виражені в числовій формі. База векторів зберігає розраховані вектори ознак для кожного зображення в базі зображень та ім'я відповідної людини. Також встановлено, що перед початком роботи системи, до бази даних потрібно додати деяку кількість зображень кожної особи та провести побудову векторів для них.

РОЗДІЛ 4

ЕКСПЕРИМЕНТАЛЬНА ВЕРИФІКАЦІЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ СИСТЕМИ ПОШУКУ ІНФОРМАЦІЇ

4.1 Експериментальна верифікація роботи системи на основі контрольного прикладу

Спроектowana система в процесі роботи використовує визначені методи для пошуку та ідентифікації людей на зображеннях. Для цього потрібно попередньо створити та натренувати необхідні для даних методів файли, а саме:

- для пошуку обличчя людини на зображенні: п'ять дескрипторів HOG – вигляд обличчя людини спереду, голова повернута вліво, голова повернута вправо, голова трохи повернута вліво та трохи повернута право;
- для побудови вектору ознак знайденого обличчя – нейронну мережу ResNet-34;
- для сегментації знайденого обличчя – нейромережу ViSeNet.

Щоб тренувати та перевіряти всі зазначені методи потрібно мати достатній об'єм навчальних даних. В даному випадку бажано мати набір фотографій з камер відеоспостереження закріплених над входом в потрібний об'єкт інфраструктури, який потрібно контролювати. Такі камери являються приватними, тому вільний доступ до даних з них відсутній, а отже наборів даних зображень з людьми в необхідній позі дуже мало. Для розробки системи будуть використані вибірки зображень людей без фіксованого кута зйомки та заднього фону.

Для стабільно працюючого детектора обличчя потрібно мати 3000-4000 позитивних прикладів і стільки ж негативних. Такі фотографії повинні бути розмічені – до кожного файлу зображення повинен бути описуючий файл, який містить координати обмежувального ящика обличчя. У вільному доступі є набори розмічених даних фотографій. До таких наборів належить VGG Face Dataset, який налічує більше 9 тис. фото різних людей; датасет FaceScrub в якому є 100 тис. фотографій 530 людей.

Розташований у вільному доступі датасет ImageNet - це база даних зображень, організована відповідно до ієрархії WordNet, тобто у вигляді графу, в якому кожен вузол налічує тисячі зображень конкретного об'єкту. Загальна кількість зображень в даному наборі складає більше 14 мільйонів файлів. Проект сприяв розвитку комп'ютерного зору та дослідженням глибинного навчання. Дані доступні для дослідників безкоштовно для некомерційного використання.

Оскільки системі потрібно не тільки знаходили обличчя, а також проводити його сегментацію та визначення контрольних точок, потрібен набір відповідно розмічених даних. Для тренування нейронної мережі ViSeNet для сегментації, використано набір даних CelebAMask-HQ – широкомасштабний набір зображень облич, який містить 30000 зображень облич з високою роздільною здатністю, вибраних із набору даних CelebA. Кожне зображення має сегментаційну маску атрибутів обличчя. Маски CelebAMask-HQ були анотовані вручну та мають розмір 512 на 512 пікселів та 19 класів, включаючи всі компоненти обличчя та аксесуари, такі як шкіра, ніс, очі, брови, вуха, рот, губи, волосся, шапка, окуляри, сережки, намисто, шия та одяга.

Тренування дескрипторів HOG виконано на наборі даних ibug 300-W – набір даних облич, який складається з 300 зображень у приміщенні та 300 на відкритому повітрі, отже, база даних складається з 600 анотованих екземплярів обличчя. Вона охоплює різні варіації ідентичності, виразу, умов освітлення, пози, оклюзії та розміру обличчя. Зображення прокоментовані із 68-бальною націнкою за напівавтоматичною методологією. Зображення бази даних були ретельно відібрані, щоб вони представляли характерний зразок складних, але природних виразів обличчя в різних умовах. Деякі зображення бази даних містять більше одного анотованого обличчя (293 зображення з 1 обличчям, 53 зображення з 2 обличчями та 53 зображення з від 3 до 7 облич). Файли зображень мають різні розміри, зокрема, половина файлів мають розмір від 48,6 тис. до 2 мільйонів пікселів, а загальний середній розмір становить 85 тис. пікселів (приблизно 292×292 пікселів).

Визначену систему пошуку інформації імплементовано на мові програмування Python в середовищі Jupyter. Дана система має структуру організації, що була визначена в підрозділі 3.1 (див. рис. 3.2).

Імплементована система використовує нейронні мережі ResNet-34 та ViSeNet, а також дескриптори HOG, в процесі своєї роботи. Методи було побудовано та натреновано на розмічених базах зображень, зазначених вище. Результуючі файли, необхідні для роботи методів, займають місце на фізичному носії, тому через необхідність завантаження їх в оперативну пам'ять, загальний час завантаження системи може складати 3 хвилини. Проте обробка одного пошукового запиту, в середньому, займає 0.3 секунди.

В якості системи управління базою даних було обрано технологію SQLite. СУБД займається організацією реляційної бази даних описів зовнішності людей та відповідає на запити, що надсилаються системою пошуку.

Пошуком та управлінням базою даних векторів займається розроблена система. Завантаження бази даних та пошук за векторами ознак виконує клас "System". Додавання нових векторів до даної бази виконується викликом окремої функції з передаванням, в якості параметру, шляху до набору підібраних фотографій окремої людини. Всі набори зображень доданих до системи осіб, зберігаються в спільній папці.

Розроблена система має три режими роботи – визначення людини на зображенні при подачі фото з нею в якості пошукового запиту; пошук людей у вигляді текстового запиту в базі даних описуючих їх пропорцій; пошук людей в базі описів, за поданим файлом зображення. Кожен з режимів потрібно перевіряти та аналізувати окремо.

Пошук за зображенням.

Контрольна перевірка пошуку людей за зображенням обличчя буде проведена на п'яти фото двох різних людей. Контрольні фотографії наведено на рис. 4.1. Для перевірки стійкості розпізнавання при зміні повороту обличчя людини на фото, було обрано 2 фото з поглядом осіб майже в камеру (Alan Grant (a) та Owen Grady (b)) , 2 фото з обличчям повернутим вбік (Alan Grant (a) та

Owen Grady (в)) та одне фото без людей (д). При цьому, обрані фотографії мають відносно невелику розмірність – в середньому 450 на 450 пікселів.



Рис. 4.1 Контрольні зображення людей

Щоб мати можливість пошуку за зображенням, для початку потрібно завантажити декілька фотографій потрібної людини в папку, запустити скрипт та передати йому шлях до папки з фотографіями. До бази даних зображень додано шість людей, при цьому в кожній папці розміщено мінімум по 20 фотографій відповідної особи. Додані файли зображень мають різні розмірності. Найменше зображення в базі має розмір 500 на 500 пікселів. На кожному доданому в базу зображенні, має бути лише обличчя потрібної людини, без присутності інших осіб. Скрипт по чергово обробить всі зображення конкретної людини за папки. На кожному зображенні виконується пошук обличчя, після чого, знайдена область обробляється нейронною мережею Resnet-34, яка формує вектор ознак для кожного зображення. Після побудови векторів зображень, вони додаються до бази векторів, в якій кожному вектору ставиться у відповідність ім'я людини.

В результаті пошуку за запитом – зображенням, система виводить імена людей які присутні на фото, а також вхідне зображення з ящиками обмеження, які вказують де саме розташовані знайдені обличчя та іменем людини біля кожного ящика. Результат ідентифікації людей на контрольних фото наведено на рис. 4.2.



Рис. 4.2. Результати пошуку за контрольними фотографіями

Після обробки контрольних зображень система знайшла та вірно ідентифікувала обличчя на всіх чотирьох зображеннях та не знайшла облич на пустій фотографії. Для знайдених облич було правильно визначено особу, а також обмежувальний ящик позначає місцезнаходження облич досить точно. Таким чином, система показала чотири правильно позитивних (true positive) результати та один правильно негативний (true negative). Даний результат свідчить про можливість системи виконувати пошук та ідентифікацію людини на фото, за умови наявності зображень з нею в базі даних.

Пошук в базі описів за текстом.

Переконаємося в можливості системи пошуку особи в базі описів людей за текстовим запитом. Для цього було оцінено зовнішні параметри шести осіб та внесено отримані дані до відповідної таблиці бази даних. Кожен параметр оцінений по шкалі від 0 до 1. Процес оцінювання проводився системою на контрольних зображеннях (обличчя осіб повернуті до камери, відсутні зайві об'єкти) та отримані значення перевірені людиною. Значення, що було додано до бази даних наведено в таблиці 4.1. У даній таблиці наведено лише числові значення параметрів.

Таблиця 4.1. Внесені записи до бази даних

Ім'я	Стать	Висота лобу	Висота носу	Ширина носу	Ширина роту	Товщина губ	Близькість очей	Ширина брів
Alan Grant	ч. (1)	0.52	0.2	0.2	0.8	0.25	0.45	0.2
John Hammond	ч. (1)	0.42	0.2	0.2	0.9	0.25	0.45	0.22
Ian Malcolm	ч. (1)	0.51	0.2	0.2	0.8	0.35	0.45	0.2
Claire Dearing	ж. (2)	0.49	0.2	0.17	0.8	0.25	0.45	0.15
Ellie Sattler	ж. (2)	0.51	0.2	0.2	0.8	0.35	0.45	0.2
Owen Grady	ч. (1)	0.52	0.2	0.2	0.8	0.25	0.45	0.2

Текстовий запит до системи сформовано на природній мові, відповідно за синтаксису запиту. Він виражає оцінку параметрів зовнішності людини у формі слів. Кожен введений параметр аналізується системою та для нього визначається числовий інтервал, за допомогою таблиці співставлень.

До системи пошуку введено три текстові запити "звичайний ніс, чоловік", "широкий рот" та "широкі губи". Вони дозволяють проілюструвати сортування за статтю та випадки, в яких параметр обличчя не є "звичайним". Результатом пошуку за першим запитом, відповідно до таблиці, мають бути всі чотири чоловіки з бази. Другий запит має повернути лише John Hammond, третій – Ian

Malcolm та Ellie Sattler. Скріншот отриманого результату пошуку наведено на рис. 4.3.

Після отримання текстового пошукового запиту, система змогла правильно зрозуміти його зміст, віднайти слова-оцінки та визначати числові інтервали пошуку для них. Використовуючи отримані інтервали було вірно побудовано запит до бази даних на SQL. В результаті отримано очікуваний результат пошуку.

```
▶ print("Запит: звичайний ніс, чоловік")
result = system.find_person_by_text("звичайний ніс, чоловік")
for r in result: print(r[1])
print()
print("Запит: широкий рот")
result = system.find_person_by_text("широкий рот")
for r in result: print(r[1])
print()
print("Запит: широкі губи")
result = system.find_person_by_text("широкі губи")
for r in result: print(r[1])

Запит: звичайний ніс, чоловік
John Hammond
Ian Malcolm
Alan Grant
Owen Grady

Запит: широкий рот
John Hammond

Запит: широкі губи
Ian Malcolm
Ellie Sattler
```

Рис. 4.3. Результати пошуку за текстовим запитом

Пошук в базі описів за зображенням.

Розроблена система підтримує функцію пошуку в базі описів зовнішності за зображенням. Дана можливість досягається завдяки алгоритму автоматичної оцінки зовнішності людини на фотографії, визначеному в попередньому розділі.

Для виконання такого пошуку, необхідно подати системі зображення, на якому погляд людини спрямований майже в камеру. Якщо людина на зображенні буде дивитись в камеру не під достатньо прямим кутом, система повідомить про неможливість правильного розрахунку пропорцій обличчя та припинить процес пошуку.

В якості пошукового запиту для перевірки даної функції, буде подано зображення Alan Grant та John Hammond. Alan Grant було обрано для перевірки пошуку, оскільки, як наведено в табл. 3.1, він має оцінку “звичайний” по всім параметрам зовнішності. Такі ж параметри зовнішності має Owen Grady. Це означає, в якості результату пошуку за зображенням, очікується отримати записи Alan Grant та Owen Grady. John Hammond єдиний в базі описів, де в якості значення параметру “ширина роту” присутнє значення 0.9, що відповідає словесній оцінці “широкий”. В даному випадку, очікується отримання результату лише John Hammond.

Підібрані зображення завантажено та подано до системи, зі вказівкою шукати в базі даних описів людей. Результати даної перевірки процесу пошуку наведено на рис. 4.4 (зліва – результат Alan Grant, справа – результат John Hammond).



Рис. 4.4. Результат пошуку за зображенням в базі описів

В якості результату пошуку, система виводить знайдені, в базі описів, імена людей, які підходять за зовнішністю до особи, зображеної на поданому фото. Додатково, система виводить результати оцінки параметрів зовнішності з їх назвою в тому форматі, в якому вони зберігаються в базі описів, а також їх словесну оцінку. Окрім вищезгаданого, виводиться також результат попередньої обробки зображення, а саме напрямок і кут повороту початкового зображення для вирівнювання знайденого обличчя на ньому, та схему результату сегментації самого обличчя з підписами легенди відповідних зон.

Як видно на рис. 4.4, система провела обробку вхідних зображень, повернула їх, вирівнявши зображене обличчя, визначила пропорції параметрів зовнішності та їх словесну оцінку, виконала відповідний пошук в базі описів людей. Результат пошуку за зображенням Alan Grant, як і очікувалось, має значення одразу двох людей, які підходять під такі параметри. Пошук за

фотографією John Hammond, повернув ім'я лише однієї особи, що відповідає очікуваному результату.

Перевіряючи всі три пошукові режими роботи розробленої системи, було отримано очікуваний результат у всіх випадках, що свідчить про адекватність системи та її здатність вирішувати поставлену задачу.

4.2 Аналіз ефективності розробленої системи пошуку інформації

Основним завданням розробленої системи є оперативний пошук особи за поданим зображенням або словесним портретом, тому важливо визначити та оцінити швидкість і точність обробки запитів.

В процесі пошуку система виконує різні дії в залежності від типу пошукового запиту. Це потребує аналізу результатів, у відповідь на конкретну форму запиту, загалом, а також перевірку кожного з етапів обробки пошукового запиту окремо.

Для оцінки точності визначення людей при пошуку за запитом у формі зображення в базі параметрів зовнішності, потрібно оцінити ступінь точності знаходження обличчя на зображенні та кількість правильних визначень людини.

Система виконує виявлення обличчя на фотографії використовуючи дескриптори HOG та метод ковзного вікна. Від точності визначення розташування обличчя на фотографії, залежить подальша точність порівняння векторів ознак знайдених облич. При вимірюванні точності знаходження регіону розташування об'єктів на зображенні використовують метрику Intersection over Union (IoU). Вона розраховується як відношення кількості пікселів на перетині знайденого детектором регіону та справжнього регіону знаходження об'єкта, до сумарної кількості пікселів знайденого та справжнього регіону.

$$IoU = \frac{\text{регіон співпадіння}}{\text{сумарний регіон}} \quad (4.1)$$

Проводити оцінку знаходження об'єкту на зображенні потрібно на заздалегідь розмічених даних. З набору фотографій різних людей взято 500 екземплярів зображень, на яких буде проводитись оцінка. Всі тестові зображення проаналізовано дескрипторами НОГ та вираховано значення IoU для кожного результату пошуку обличчя.

Для прикладу, візьмемо одну розмічену фотографію. Для неї справжні координати обмежуючої рамки обличчя це дві точки $A_1(154; 30)$ – лівий верхній кут та $A_2(446; 207)$ – правий нижній кут. Використовуючи дескриптори НОГ, автоматично визначено розташування обличчя на зображенні та отримано результат – точки $B_1(138; 27)$ та $B_2(457; 212)$. Тоді значення оцінки IoU для даного прикладу розраховується так:

$$x_{max} = \max(x_{A1}, x_{B1}) = \max(154, 138) = 154$$

$$y_{max} = \max(y_{A1}, y_{B1}) = \max(30, 27) = 30$$

$$x_{min} = \min(x_{A2}, x_{B2}) = \min(446, 457) = 446$$

$$y_{min} = \min(y_{A2}, y_{B2}) = \min(207, 212) = 207$$

$$\begin{aligned} S_{\text{перетину}} &= (x_{\min} - x_{\max})(y_{\min} - y_{\max}) = (446 - 154)(207 - 30) \\ &= 292 * 177 = 51684 \end{aligned}$$

$$S_{\text{сп.рег}} = (x_{A2} - x_{A1})(y_{A2} - y_{A1}) = (446 - 154)(207 - 30) = 292 * 177 = 51684 \text{ – справжній регіон}$$

$$S_{\text{перед.рег}} = (x_{B2} - x_{B1})(y_{B2} - y_{B1}) = (457 - 138)(212 - 27) = 319 * 185 = 59015 \text{ – передбачений регіон}$$

$$IoU = \frac{S_{\text{перетину}}}{(S_{\text{сп.рег.}} + S_{\text{перед.рег.}} - S_{\text{перетину}})} = \frac{51684}{59015} = 0,87$$

Таким чином, для даного прикладу, значення IoU становить 0.87, що означає 87% точності автоматичного виявлення обличчя, використовуючи НОГ дескриптори.

Для тестової вибірки середнє значення $IoU_{cp} = 0,85$. Також тестову вибірку було поділено на фотографії, на яких обличчя людини повернуте в різні напрямки та до кожного класу було застосовано відповідний дескриптор. Результати середнього значення IoU для кожного випадку наведено в таблиці 4.2.

Таблиця 4.2. Середні значення IoU

Обличчя повернуте вліво	Обличчя повернуте вліво на незначний кут	Обличчя повернуте до камери	Обличчя повернуте вправо на незначний кут	Обличчя повернуте вправо	Середнє значення
0,82	0,84	0,92	0,85	0,79	0,85

Можна зробити висновок, що застосування дескрипторів HOG для пошуку обличчя на фотографіях показує достатньо високий результат точності, тобто його доцільно використовувати в системі пошуку. Найнижчий результат отримано у випадку повного повороту обличчя, що свідчить про можливе зменшення точності ідентифікації обличчя на фото, при його повороті.

Щоб виконати оцінку точності пошуку людини в базі фотографій, необхідно до бази даних додати декілька наборів зображень людей з підписами їх імен. Потім створити та завантажити в базу вектори ознак для кожного зображення. Після чого провести перевірку точності пошуку на нових, не присутніх в базі зображень, фотографіях.

Для оцінки точності визначення людини на зображенні використано метрику точності, яка розраховується за формулою:

$$\text{Точність} = \frac{TP}{TP + FP} \quad (4.2)$$

Де TP – true positive, кількість зображень людини, на яких система правильно визначила її особу;

FP — false positive, кількість зображень людини, на яких система не правильно визначила, що дана особа не являється потрібною.

До бази даних було завантажено зображення 30 людей (по 20 фотографій) та проведено перевірку роботи системи на 10 зображеннях кожної людини.

Розрахунок точності проведено для кожної людини окремо. Визначено кількість випадків правильного розпізнавання людини, та випадків, коли потрібну людину було неправильно ідентифіковано. Отриманні значення використано для прорахунку точності ідентифікації даної людини. З розрахованих оцінок точності було виведено середнє значення, яке становить 85,31.

Також проведено оцінку точності ідентифікації осіб при різних кількостях зображень людей в базі даних. Графік зміни точності ідентифікації наведено на рис. 4.6.

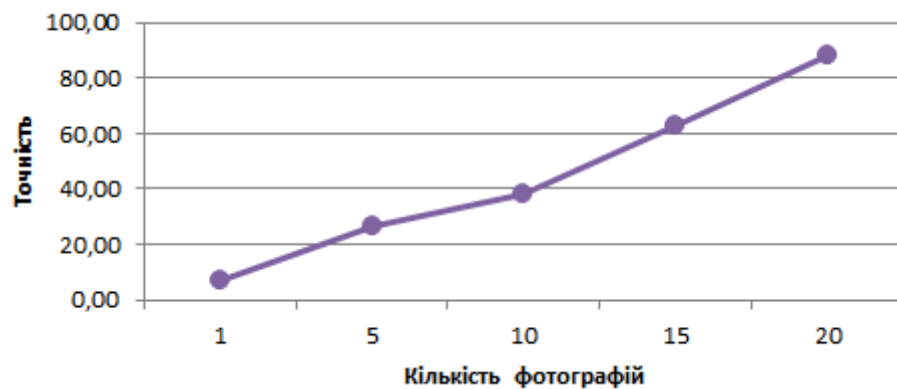


Рис. 4.6. Точність ідентифікації при зміні кількості фотографій

Точність ідентифікації зростає зі збільшенням кількості фотографій людини в базі даних. При цьому для досягнення задовільної точності роботи системи, необхідно мати принаймні 20 зображень кожної людини.

Одне зображення обробляється системою в середньому за 0,37 секунди, при розрахунку на графічному процесорі. При цьому час знаходження обличчя на фотографії займає 0,011 секунди. Даний показник затрати часу є достатнім для виконання оперативного пошуку людей при аналізі відеозаписів. Такий результат досягається лише при роботі системи на графічних процесорах. Якщо ж використовувати центральний процесор, то середній час обробки одного зображення займає 72 секунди. Також потрібно враховувати, що зі збільшенням розміру бази фотографій, середній час ідентифікації особи на зображенні буде зростати.

Отриманий результат перевірки пошуку людини за фотографією в базі зображень вказує на те, що дана система може правильно ідентифікувати осіб на фото та виконувати оперативний пошук людей. Це свідчить, що розроблена система пошуку може вирішувати поставлену задачу ідентифікації особи на фотографії, порівнюючи його з наявними зображеннями в базі даних.

Основним завданням, яке має вирішувати розроблена система, є проведення пошуку можливих людей в базі даних параметрів зовнішності за зображенням. Для цього система має автоматично вилучати пропорції обличчя людини з фотографії та порівнювати їх зі значеннями, збереженими в базі даних.

Для аналізу можливості пошуку людини на зображенні в базі даних параметрів, до неї було внесено 20 записів людей, які мають різні пропорції обличчя. Майже всі вони мають деякий параметр, який має значення не “звичайний”. Так в базі налічуються записи з таким параметрами: лоб – 3 високий та 1 низький; ширина носу – 2 широкий та 3 вузький; висота носу – 2 високий та 4 низький; товщина губ – 5 широкі та 8 вузькі. Ширина брів та близькість очей у всіх осіб в базі має значення “звичайна”. Саме унікальність параметрів обличчя дозволяє зменшити результат пошуку та визначення особи.

Відібрано по 5 фотографій кожної особи, яка присутня в базі даних – в сумі 100 фотографій. На всіх підібраних фотографіях особа була повернута в бік камери та середній розмір зображення становить 700 на 550 пікселів. Після проведення тесту було визначено, що у 73% випадків в результаті пошуку особи на фотографій була присутня потрібна людина. Час обробки одного зображення в середньому займає 0,14 секунди, що свідчить про можливість оперативного пошуку людей на фотографії, у випадку наявності бази параметрів. При цьому люди з тонкими губами були неправильно визначені найчастіше (на 23 з 40 фотографій осіб з тонкими губами даний параметр було оцінено неправильно), через те, що система оцінювала пропорції їх губ як “звичайні”. Це пояснюється тим, що товщина губ за фотографії займає найменшу кількість пікселів.

При автоматичній оцінці параметрів зовнішності найчастіше неправильно визначався параметр висоти лобу, адже на багатьох фотографіях у людей може

бути присутній головний убір або зачіска може прикривати зону лобу. Тому автоматичне визначення даного параметру потребує подальшого аналізу та розробки методу покращення точності розрахунку, наприклад враховуючи відстань від найвищої точки голови людини.

Проведена перевірка можливості пошуку особи в базі параметрів за її фотографією, показує, що розроблена система може автоматично виконувати оцінку параметрів зовнішності людини на зображенні та використовувати отримані значення для встановлення її особистості.

Також результат перевірки, свідчить про те, що система може автоматично виконувати розрахунок параметрів зовнішності людини. Дана здатність дає змогу в подальшому автоматично виконувати аналіз фотографій обличчя з бази даних зображень та формувати базу параметрів зовнішності без втручання особи, яка буде контролювати систему.

Отримані результати перевірки розробленої системи вказують на адекватність її роботи та свідчать про те, що система здатна вирішувати поставлені задачі. За результатами тестування можна зробити висновок, що запропонована система оперативного пошуку з використанням нейронних мереж, достатньо добре справляється з поставленою задачею пошуку людей з фотографій в базах даних зображень та параметрів зовнішності. До переваг розробленої системи пошуку можна віднести:

- достатньо високу точність ідентифікації людини на фотографії, при порівнянні її із зображеннями в базі;
- високу швидкість обробки одного зображення – в середньому 0,37 секунди при пошуку в базі даних фотографій та 0,14 секунди при пошуку в базі параметрів зовнішності;
- можливість виконання пошуку людини на фотографії або за словесним портретом у випадку наявності лише бази даних пропорцій обличчя;
- можливість автоматичної оцінки параметрів зовнішності людини на фотографії для формування баз даних параметрів зовнішності.

Недоліками розробленої системи, є:

- автоматична оцінка параметрів обличчя неможлива, якщо погляд особи на зображенні не направлено в бік камери;
- автоматичний розрахунок пропорцій дає неправильні значення деяких параметрів, якщо на зображенні присутні предмети, які перекривають частину обличчя. Через даний недолік, автоматичний розрахунок пропорцій обличчя потребує контролю результатів розрахунку людиною;
- необхідність роботи на графічному процесорі для досягнення достатньої швидкості обробки зображень;
- невисока точність ідентифікації особи при низькій роздільній здатності зображення.

Має місце подальше покращення алгоритму автоматичного розрахунку пропорцій обличчя на фотографії. Це збільшить точність пошуку людини із зображення в базі параметрів зовнішності та дозволить автоматичне формування такої бази, аналізуючи фотографії людей, без втручання особи, яка буде контролювати отримані розрахункові значення.

Подальшим розвитком розробленої системи оперативного пошуку особи може служити створення додаткової підсистеми вибору з відеоряду, який формується, наприклад, шляхом автоматичного зчитування з відеокамер спостереження за навколишнім оточенням, зображень людських облич з подальшим використанням цієї інформації в якості вхідної для системи ідентифікації особистості.

4.3 Висновки до розділу 4

Проведено експериментальну верифікацію роботи системи та продемонстровано виконання основних функцій на основі контрольного прикладу. В результаті перевірки обробки всіх пошукових запитів було отримано очікувані значення.

Розраховано середню точність виявлення обличчя людини на фотографії. Дане значення становить 85%. Такий результат свідчить про доцільність застосування дескрипторів НОГ для виконання поставлених задач. Проведено

оцінку точності розпізнавання особи системою. Середня точність ідентифікації особи становить 85.31%, що є високим результатом. На основі даного результату можна зробити висновок про здатність системи виконувати поставлену задачу ідентифікації людей на фотографії, порівнюючи його зображеннями в базі даних.

Перевірено здатність системи автоматично визначати параметри зовнішності людини на фотографії та порівнювати їх з записами в базі даних параметрів зовнішності. В 73% запитів в результаті пошуку була присутня потрібна людина. Середній час обробки одного зображення склав 0,14 секунди. Зроблено висновок, про здатність системи виконувати оперативний пошук людей на фотографії в базі параметрів. Визначено, що найменш точним є розрахунок параметрів губ та лобу.

Наведено основні переваги та недоліки розробленої системи. Перевагами системи є: висока точність ідентифікації людини на фотографії; висока швидкість обробки одного зображення; можливість автоматичної оцінки параметрів зовнішності людини на фотографії. До недоліків віднесено: неможливість пошуку людини на зображенні в базі параметрів, якщо її погляд не направлено в бік камери; потреба контролю результатів автоматичного розрахунку пропорцій обличчя людиною; невисоку точність ідентифікації особи при низькій роздільній здатності зображення.

На основі результатів проведених перевірок, зроблено висновок про здатність розробленої системи вирішувати поставлені задачі.

ВИСНОВКИ

В роботі було розглянуто задачу ідентифікації людей, сучасний стан її досліджень, роль і місце процесу пошуку інформації в ній. Для розпізнавання осіб найбільше використовують біометричні ознаки, зокрема параметри обличчя. Ще наприкінці минулого століття вчені проводили розробку методів аналізу обличчя на фотографіях. Пошук інформації при визначенні особистості людини необхідний для швидкого доступу до даних, які підходять за деякими критеріями, для їх подальшого аналізу.

Вивчено процес пошуку інформації в задачах ідентифікації осіб та основні етапи роботи системи пошуку. В процесі пошуку інформації її спочатку індексують та поміщують в базу даних. В разі отримання пошукового запиту, його порівнюють з індексованими даними, сортують за критеріями пошуку та релевантністю, після чого повертають знайдену інформацію в якості відповіді.

Визначено, що основним типом інформації, якою оперують в процесі ідентифікації є мультимедійні файли та реляційні бази даних. Для проведення процесу пошуку таких файлів до них застосовують анотування, класифікацію або виконують побудову вектору ознак. Результат обробки інформації в подальшому використовується при пошуку файлів.

Проведено аналіз сучасних систем ідентифікації особи за обличчям та визначено, що в них відсутня можливість порівнювати фотографію людини з параметрами зовнішності, які виражені в словесній формі або зберігаються у формі числових значень. Було зроблено висновок про необхідність розробки інформаційно-пошукової системи, яка зможе виконувати ідентифікацію особи в реальному часі за фотографією її обличчя або словесним портретом.

Описано технологію пошуку інформації. Система може приймати в якості запиту зображення або текст. При отриманні запиту у формі зображення, для нього будується вектор ознак, який порівнюється з векторами з бази. Підтримується можливість порівняння зображеної особи з записами в базі даних пропорцій обличчя, яка досягається використанням розробленого алгоритму автоматичного розрахунку параметрів зовнішності. У випадку надходження текстового запиту, його синтаксис аналізується, після чого формується запит до бази даних описів зовнішності на мові SQL.

Визначено алгоритм ідентифікації особи на основі порівняння зображень. Для цього вирішено застосовувати дескриптори HOG та модифіковану нейронну мережу Resnet-34. Дескриптори HOG використовуються для визначення розташування обличчя на фотографії. Нейромережа виконує побудову вектору ознак для кожного знайденого обличчя, після чого отримані вектори ознак порівнюються для встановлення особи зображеної людини.

Створено синтаксис текстових пошукових запитів та визначено спосіб їх обробки системою. Він складається з назв параметрів і їх значень, виражених в словесній формі, та відображає словесний портрет описуваної людини. Для введених параметрів, за допомогою таблиці співвідношень, визначаються числові інтервали, після чого формується запит до бази даних на мові SQL.

Розроблено алгоритм автоматичного розрахунку параметрів зовнішності людини на фотографії. Спочатку проводиться перевірка напрямку погляду зображеної особи. Для цього, методом каскаду дерев регресії, визначаються контрольні точки обличчя. Потім, в разі, якщо погляд людини направлено в бік камери, зображення сегментується використовуючи нейронну мережу ViSeNet. Отриманий результат сегментації аналізується та розраховуються піксельні розміри частин обличчя.

Розроблено архітектуру інформаційно-пошукової системи. Система складається з трьох основних елементів: модулю обробки зображень, який виконує пошук облич на фотографіях, побудову векторів ознак та автоматичний розрахунок параметрів зовнішності; модулю обробки текстових запитів, що відповідає за аналіз вхідних текстових запитів та формування запитів до бази даних; модулю “система”, який відповідає на запити користувача, надсилає запити до бази параметрів зовнішності та виконує порівняння векторів ознак для зображень.

Сформовано структуру бази параметрів зовнішності. Вона складається з однієї таблиці, яка зберігає значення пропорцій обличчя, виражені в числовій формі.

Реалізовано розроблену систему пошуку та проведено експериментальну верифікацію її роботи. В результаті перевірки здатності системи обробляти різні типи запитів було отримано очікувані значення.

Визначено середню точність ідентифікації особи на зображенні. Вона становить 85.31%. На основі отриманого значення точності було зроблено висновок про здатність системи вирішувати поставлену задачу розпізнавання людей на фотографії.

Проведено перевірку здатності системи автоматично визначати параметри зовнішності людини на фотографії та порівнювати їх з записами в базі даних. В 73% випадків, система змогла точно розрахувати пропорції зовнішності та правильно визначила всіх людей, які їм відповідають.

Визначено основні переваги та недоліки розробленої системи. До переваг створеної системи віднесено: високу точність ідентифікації людини на зображенні; високу швидкість обробки одного запиту; здатність автоматичної оцінки параметрів зовнішності людини на фотографії. Недоліками системи є: відсутність можливості розрахунку параметрів зовнішності людини, якщо її погляд не направлено в бік камери; необхідність часткового контролю результатів автоматичного визначення пропорцій обличчя; невисоку точність ідентифікації особи при низькій роздільній здатності зображення.

Зроблено висновок про здатність розробленої системи вирішувати поставлені задачі.

На основі отриманих результатів роботи можна зробити висновок, що всі завдання дослідження було виконано, а мету дипломної роботи досягнуто.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. S. Liao. Partial face recognition: Alignment-free approach / S. Liao, A. K. Jain, S. Z. Li // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 2012. – Vol. 35, No. 5. – p. 1193–1205.
2. M. Jridi. One lens optical correlation: Application to face recognition / M. Jridi, T. Napoleon, A. Alfalou // *Applied Optics*. – 2018. – Vol. 57. – p. 2087–2095.
3. T. Napoleon. Pose invariant face recognition: 3D model from single photo / T. Napoleon, A. Alfalou // *Optics and Lasers in Engineering*. – 2017. – Vol. 89. – p. 150–161.
4. Y. Ouerhani. Fast face recognition approach using a graphical processing unit “GPU” / Y. Ouerhani, M. Jridi, A. Alfalou // *2010 IEEE International Conference on Imaging Systems and Techniques*. – 2010. – p. 80–84.
5. W. Yang. A fingerprint and finger-vein based cancelable multi-biometric system / W. Yang, S. Wang, J. Hu, G. Zheng, C. Valli // *Pattern Recognition*. – 2018. – Vol. 78. – p. 242–251.
6. N. P. Patel. Optimize Approach to Voice Recognition Using IoT. / N. P. Patel, A. Kale // *2018 International Conference on Advances in Communication and Computing Technology (ICACCT)*. – 2018. – p. 251–256.
7. M. S. Nixon. Human Identification Based on Gait / M. S. Nixon, Tieniu Tan, Rama Chellappa // *Springer Science & Business Media*. – 2010. – p. 188
8. Q. Wang. New perspectives in face correlation research: A tutorial / Q. Wang, A. Alfalou, C. Brosseau. // *Advances in Optics and Photonics*. – 2017. – Vol. 9. – p. 1–78.
9. A. Alfalou. Optimization of decision making for face recognition based on nonlinear correlation plane / A. Alfalou, C. Brosseau, W. Kaddah // *Optics Communications 2015*. – 2015. – Vol. 343. – p. 22–27.
10. C. Zhao. Bisecting k-means clustering based face recognition using block-based bag of words model / C. Zhao, X. Li, Y. Cang // *Optik - International Journal for Light and Electron Optics*. – 2015. – Vol. 126, No. 19. – p. 1761–1766.

11. A. HajiRassouliha. FPGA implementation of 2D cross-correlation for real-time 3D tracking of deformable surfaces / A. HajiRassouliha, T.P.B. Gamage, M.D. Parker, M.P. Nash, A.J. Taberner, P.M. Nielsen // 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013). – 2013. –p. 352–357.
12. Y. Kortli. A comparative study of CFs, LBP, HOG, SIFT, SURF, and BRIEF techniques for face recognition / Y. Kortli, M. Jridi, A. Al Falou, M. Atri // Pattern Recognition and Tracking XXIX. – 2018. – Vol. 10649. – p. 106490M.
13. Z. Dehai. A pca-based face recognition method by applying fast fourier transform in pre-processing / Z. Dehai, D. Da, L. Jin, L. Qing // 3rd International Conference on Multimedia Technology (ICMT-13). – 2013. – Vol. 226. – p. 1148–1155.
14. Y. Ouerhani. Road mark recognition using HOG-SVM and correlation / Y. Ouerhani, A. Alfalou, C. Brosseau // Optics and Photonics for Information Processing XI. – 2017. – Vol. 10395. – p. 103950Q.
15. P. Khoi. Face Retrieval Based on Local Binary Pattern and Its Variants: A Comprehensive Study / P. Khoi, L.H. Thien, V.H. Viet // International Journal of Advanced Computer Science and Applications. – 2016. – Vol. 7. – p. 249–258.
16. Y. Ouerhani. Optimized pre-processing input plane GPU implementation of an optical face recognition technique using a segmented phase only composite filter / Y. Ouerhani, M. Jridi, A. Alfalou, C. Brosseau // Optics Communications 2013. – 2013. – Vol. 289. – p. 33–44.
17. D.N. Parmar. Face recognition methods and applications / D.N. Parmar, B.B. Mehta // International Journal of Computer Technology and Applications. – 2014. – Vol. 4. – p. 84-86.
18. H.D.S. Gowda. Multimodal Biometric Recognition System Based on Nonparametric Classifiers / H.D.S. Gowda, G.H. Kumar, M. Imran // Data Analysis Learning 2018. – 2018. – Vol. 43. – p. 269–278.
19. R.J. Jr. Uhl. A framework for recognizing a facial image from a police sketch / R.J. Jr. Uhl, N.V. Lobo // IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 1996. – p. 586–593.

20. W. Hu. A survey on visual surveillance of object motion and behaviors / W. Hu, T. Tan, L. Wang, S. Maybank // IEEE Transactions on Systems, Man, and Cybernetics. – 2004. – Vol. 343. – p. 334–352.
21. J. Sivic. Person spotting: video shot retrieval for face sets / J. Sivic, M. Everingham, A. Zisserman // International Conference on Image and Video Retrieval. – 2005. – p. 226-236.
22. A. Bialkowski. A Database for Person Re-Identification in Multi-Camera Surveillance Networks / A. Bialkowski, S. Denman, S. Sridharan, C. Fookes, P. Lucey // 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA). – 2012. – p. 1-8.
23. S. Bak. Retrieval system for person re-identification / S. Bak, F. Bremond, V. Bathrinarayanan // 2014 International Conference on Computer Vision Theory and Applications (VISAPP). – 2014. – p. 732-738.
24. D. A. Vaquero. Attribute-based people search in surveillance environments / D. A. Vaquero, R. S. Feris, D. Tran, L. Brown, A. Hampapur M. Turk // 2009 Workshop on Applications of Computer Vision (WACV). – 2009. – p. 1-8.
25. Информационные ресурсы и поисковые системы: учебное пособие / Максимов Н.В., Голицина О.Л., Тихомиров Г.В та інші. – Москва: Московський інженерно-фізичний інститут, 2008. – 400 с.
26. Борисов Д.Н. Корпоративные информационные системы. – Воронеж: Издавнично-поліграфічний центр Воронізького державного університету, 2007. – 99 с.
27. A. Bouguettaya. Interconnecting Heterogeneous Information Systems / A. Bouguettaya, B. Benatallah, A. K. Elmagarmid // Springer Science+Business Media. – 1998. – 228 p.
28. G. J. Kowalski. Information Storage and Retrieval Systems / G. J. Kowalski, M. T. Maybury // Kluwer Academic Publishers. – 2000. – 333 p.
29. S. Deb. Multimedia Systems and Content-Based Image Retrieval / S. Deb // IGI Global. – 2003. – 406 p.

30. Z. Ma. Artificial Intelligence for Maximizing Content Based Image Retrieval / Z. Ma // Information Science Reference. – 2009. – 429 p.
31. Z. Yu-Jin. Semantic-Based Visual Information Retrieval / Z. Yu-Jin // IRM Press. – 2006. – 385 p.
32. M. Mahmood. Using Artificial Neural Network for Multimedia Information Retrieval / M. Mahmood, W. J. Al-Kubaisy, B. Al-Khateeb // Journal of Southwest Jiaotong University. – 2019. – Vol. 54, No. 3.
33. N. Mehendale. Facial emotion recognition using convolutional neural networks / N. Mehendale // SN Applied Sciences. – 2020. 0 Vol. 446, No. 2.
34. N. Dadal. Histogram of Oriented Gradients for Human Detection / N. Dadal, B. Triggs // IEEE Conference Computer Vision and Pattern Recognition. – 2005. – p. 886-893.
35. P. Viola. Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M. Jones // IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2001. – Vol. 1. – p. 511.
36. J. Uijlings. Selective Search for Object Recognition / J. Uijlings, K. Sande, T. Gevers, A.W.M. Smeulders // International Journal of Computer Vision. – 2013. – Vol. 104. – p. 154-171.
37. P. F. Felzenszwalb. Efficient graph-based image segmentation / P. F. Felzenszwalb, D. P. Huttenlocher // International Journal of Computer Vision. – 2004. – Vol. 59, No. 2. – p. 167-181.
38. K. He. Deep Residual Learning for Image Recognition / K. He, X. Zhang, S. Ren, J. Sun // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – p. 770-778.
39. I. Elezi. The Group Loss for Deep Metric Learning / I. Elezi, S. Vascon, A. Torcinovich, M. Pelillo, L. Leal-Taixe // European Conference on Computer Vision 2019. – 2019. – p. 277-294.
40. Рогачёв Н. Е. Проблема коллапса нейронных сетей при использовании Triplet Loss. – Мінськ: Білоруський державний університет, 2018. – с. 92.

41. C. Yu. BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation / C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, N. Sang // European Conference on Computer Vision (ECCV). – 2018. – p. 325-341.

42. V. Kazemi. One Millisecond Face Alignment with an Ensemble of Regression Trees / V. Kazemi, J. Sullivan // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2014. – p. 1867-1874.

P. Dollar. Cascaded pose regression / P. Dollar, P. Welinder, P. Perona // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2010. – p. 1078-1085