

І. В. Курта, А. Е. Лагун

Національний університет “Львівська політехніка”,
кафедра інформаційних систем і технологій

РОЗРОБЛЕННЯ СИСТЕМИ РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ ДЛЯ ВІДЕОСПОСТЕРЕЖЕННЯ

<https://doi.org/10.23939/amm2020.01.057>

© Курта І. В., Лагун А. Е., 2020

Досліджено принципи побудови систем спостереження та розпізнавання об’єктів. Наведено класифікацію способів розпізнавання людських облич. Проаналізовано роботу мережі прогресивного калібрування (ПКМ) для розпізнавання людських облич. Розроблено алгоритм розпізнавання облич, створено програмну систему розпізнавання облич і проведено її тестування.

Ключові слова: зображення, алгоритм, розпізнавання облич, база даних, нейронна мережа, штучний інтелект, клієнт, сервер.

In the article were researched the principles of building systems for observation and recognition of objects. Also we have given the classification of human faces recognition methods. Authors have analized the features of operation for the progressive calibration network (PCN) for human face recognition. And finally has been created and tested the developed face recognition algorithm as the realized software system.

Key words: image, algorithm, face recognition, database, neural network, artificial intelligence, client, server.

Вступ

Сьогодні різні цифрові технології є надзвичайно поширеними через розвиток комп’ютерної техніки. Зокрема, особливо актуальними є системи комп’ютерного зору та аналізу відеопотоків. Такі системи можна застосовувати для захисту від несанкціонованого доступу на приватні території та організації, пошуку вільного місця для паркування автомобіля, розпізнавання обличчя людини, наприклад у відеодомофонах, визначення номерів автомобіля у разі порушення правил дорожнього руху. Отже, при використанні систем комп’ютерного зору основним завданням є розпізнавання об’єктів. Якщо для людини розпізнавання образів є простим завданням, то для комп’ютерної техніки все не так просто. Одним із важливих завдань є пошук оптимальних зображень стосовно критеріїв якості розпізнавання і швидкості обробки. Також під час обробки відеопослідовностей потрібно знайти цифрову відповідність між двома сусідніми зображеннями. Важливим є використання інтелектуального відеоспостереження, яке використовує штучний інтелект та нейронні мережі для вирішення багатьох завдань розпізнавання образів. Для інтелектуальних відеосистем можна використовувати додаткову камеру слідкування, яка може налаштовуватися при зникненні об’єкта із зони спостереження. Ще одним із застосувань систем відеоспостереження є аналіз відеопотоків під час контролю дорожньо-транспортного руху, безпеки на приватних територіях та в приміщеннях з обмеженим доступом, використання автоматизованого управління на промислових підприємствах.

Оскільки системи аналізу відеопотоку є надзвичайно актуальними, то в цій статті буде розглянуто методи та алгоритми розпізнавання об'єктів, що використовують штучний інтелект, нейронні мережі та машинне навчання, а також досліджено системи відеоспостереження та розпізнавання об'єктів у реальному часі.

Огляд систем відеоспостереження

Сьогодні питання захисту від несанкціонованого доступу на приватні території є особливо актуальними, причому встановлення систем відеоспостереження дає змогу вирішити проблеми організаційно-технічного захисту. В системах відеоспостереження використовують різні прилади та устаткування для організації відеонагляду та відеоконтролю у приміщенні або на деякому іншому захищеному об'єкті. Ці системи можна використовувати на територіях різного масштабу як зовні, так і всередині будівель [1].

Наведемо тепер перелік функцій, які виконують системи відеоспостереження.

1. Запис подій, які відбуваються на об'єкті на пристрої пам'яті, які містяться стаціонарно в системі або у хмарних сховищах.

2. Використання постійного знімання території для систем захисту, що дає змогу оперативно врегулювати певні непередбачені ситуації.

3. Можливість спостереження за територією в нічний час з використанням інфрачервоного світла, що дозволяє показати і записати зображення краще, ніж у денний час.

Розвиток інтелектуальної відеоаналітики відбувається за двома основними технологіями – ідентифікація і трекінг. На основі правил алгоритмів відеоаналізу формуються функціональні можливості системи, які необхідні для побудови сучасних систем відеоспостереження.

Процес ідентифікації дозволяє розпізнати образ за відеозображенням, згрупувати його за класами або конкретними шаблонами і порівняти із заздалегідь підготовленою базою еталонних зображень. Найпопулярнішими системами ідентифікації є розпізнавання обличчя і автомобільних номерів.

Технологія трекінгу в кадрі передбачає пошук рухомих об'єктів, після чого ці об'єкти визначаються і класифікуються. При цьому робиться акцент на характеристиках рухомого об'єкта – розмір, колір і швидкість. При обробці рухомих зображень варіантів трекінгу може бути доволі багато.

У випадку трекінгу використовують різні відеодетектори, які бувають ситуаційними і сервісними.

Ситуаційні детектори є апаратним забезпеченням, яке фіксує місцезнаходження об'єкта відеоспостереження. Якщо відбувається перехід у заборонені зони, то видається сигнал про порушення [2]. Можливі такі порушення, які характерні для ситуаційних детекторів:

- перетин об'єктом визначеної лінії;
- переміщення в контрольованій зоні;
- вихід чи зупинка об'єкта на контрольованій території;
- залишений на контрольованій території предмет.

Сервісні детектори є програмним забезпеченням, яке використовується в IP-камерах для захисту територій. У цих детекторах використовують такі характеристики, як перекриття об'єктиву, переміщення або відхилення камери, відсутність фокусування, зміна фону або засвічування камери.

Зазначимо, що до трекінгу належить також інтелектуальний пошук в архівах. Це пошук, який допомагає оператору швидко знаходити потрібний матеріал за фактом спрацювання детектора, коли точний час події невідомий.

Мета роботи

Метою роботи є аналіз та дослідження систем відеоспостереження та розпізнавання об'єктів в реальному часі з виявленням облич, які можуть бути розташовані як вертикально, так і повернуті на різні кути.

Виклад основного матеріалу

Класифікація алгоритмів розпізнавання людських облич

Система виявлення облич здатна ідентифікувати або перевірити особу на цифровому зображенні або відеокадрі. Системи виявлення облич, як правило, ґрунтуються на порівнянні особливостей зображення шуканого обличчя з обличчями, які зберігаються в базі даних. Також у системах розпізнавання облич можна використати біометричний додаток на основі штучного інтелекту. Після аналізу моделей на основі текстур і форми обличчя цей біометричний додаток може розпізнати людину.

Відома класифікація методів розпізнавання людських облич, яка представлена на рис. 1. Ці методи поділяють на чотири категорії (рис. 1).



Рис. 1. Класифікація методів виявлення обличчя

1 категорія. Метод на основі знань використовує для розпізнавання облич накопичені знання. Зрозуміло, що кожне людське обличчя має ніс, очі та рот, які розташовані на певній відстані між собою. Найважливішим у цьому методі є вибір відповідного набору правил. За таким підходом не можна знайти багато облич у кількох зображеннях.

2 категорія. Метод з використанням функцій полягає у знаходженні облич, використовуючи різні структурні особливості обличчя. Насамперед створюють класифікатор структурних особливостей, а потім використовують його для розмежування частин обличчя та не обличчя. Основна ідея полягає у тому, щоб перейти межу інстинктивного знання облич. Цей підхід розділений на кілька частин і є успішним для більшості фотографій з великою кількістю облич.

3 категорія. В методі порівняння шаблонів використовуються наперед визначені і параметризовані шаблони для знаходження облич. В цьому випадку основним є співвідношення між вхідними зображеннями і існуючими шаблонами облич. Взагалі кажучи, людське обличчя поділяється на чотири частини: очі, ніс, рот і контур обличчя. Інший спосіб визначення обличчя використовує технологію виявлення ребер. Такий підхід є простим, проте недостатнім для розпізнавання обличчя. Однак для вирішення таких проблем можна використати шаблони, які деформуються.

4 категорія. Метод, що орієнтований на зовнішній вигляд і залежить від набору делегатів зображень обличчя, що навчаються для з'ясування моделі обличчя. Для пошуку відповідних характеристик зображень обличчя використовуються машинне навчання і статистичний аналіз. Метод на основі зовнішнього вигляду, своєю чергою, поділяється на такі похідні методи:

- на основі власних поверхонь, що використовується для розпізнавання облич за допомогою аналізу основних компонентів;
- на основі розподілу, які можуть бути використані для визначення підпростору, що представляє схеми обличчя;
- з використанням нейронних мереж, які дозволяють виявити об'єкти, обличчя, емоції та розпізнати обличчя;
- підтримки векторних машин, що використовують лінійні класифікатори, які максимально підвищують вірогідність отримання правильного рішення на основі навчальних наборів;
- використання байєвих класифікаторів, які обчислюють ймовірність наявності обличчя на фотографії, обчислюючи частоту серії рисунків у навчальних зображеннях;
- використання прихованої моделі Маркова та марківських випадкових полів, у яких є варіанти рис обличчя, які зазвичай характеризують смужками пікселів;
- на основі індуктивного навчання, що використовують алгоритми Квінлана C4.5 або Мітчелла FIND-S.

Базова схема алгоритму прогресивних мереж калібрування

Для розроблення системи розпізнавання використовують алгоритм PCN [3], в якому використано мережу прогресивного калібрування і який є детектором обличчя в режимі реального часу. Згідно з цим алгоритмом поступово калібрується кут повороту в площині кожного обличчя кандидата вертикально в триступеневу багатозадачну мережу, яка згортається. Процес калібрування необхідного кута повороту поділяється на кілька кроків і лише наближено прогнозує орієнтацію на кожному етапі. До вхідного зображення застосовується принцип піраміди зображень для отримання та виявлення всіх наявних зображень різних розмірів. Кожне обличчя передається через детектор поетапно. Після проходження всього зображення детектор обирає обличчя кандидатів та їхні обмежувальні координатні поля і створює список рівнів довіри для кожного обмежувального поля. Потім більшість кандидатів із низькою ймовірністю наявності потрібного обличчя відхиляються.

За цим зображенням усі кандидати на обличчя, яке шукають, отримують, використовуючи принципи піраміди зображення і розсувного вікна. Кожне вікно-кандидат проходить через детектор поетапно. На кожному етапі PCN детектор відхиляє більшість кандидатів із невеликою подібністю обличчя, регресує обмежувальні поля інших кандидатів та калібрує орієнтації обернутих у площині (RIP) кандидатів на обличчя.

1 етап.

Для кожного вхідного вікна x PCN виконує три завдання: класифікація: обличчя чи “не обличчя”, регресія обмежувального поля та калібрування:

$$[p, q, s] = Y(x), \quad (1)$$

де Y – це детектор, структурований як невелика конволюційна нейронна мережа (КНМ). p – коефіцієнт правильності обличчя, q – вектор прогнозу регресії обмежувального поля, s – оцінка орієнтації.

Перше завдання має мету розрізнити обличчя від “не обличчя” з малими втратами за формулою:

$$L_1 = k \log(p) + (1 - k) \log(1 - p), \quad (2)$$

де $k = 1$, якщо x є обличчям, інакше 0.

Наступним завданням є спроба регресувати тонке обмежувальне поле [4], як показано нижче:

$$L_2(q, q') = S(q - q'), \quad (3)$$

де q і q' є передбаченими та справжніми результатами регресії, відповідно, S – це надійне плавне визначення неавторизованих людей. Регресії обмежувального поля складаються з характеристик:

$$t_w = w' / w,$$

$$t_a = (a' - a + 0.5(w' - w)) / w',$$

$$t_b = (b' - b + 0.5(w' - w)) / w', \quad (4)$$

де a, b – це координати лівого верхнього кутка вікна, w – ширина вікна. Змінні a і a' містять значення прогнозованого і реального вікон (аналогічно b і b' , w і w').

Третє завдання визначає орієнтацію кандидата у бінарній класифікації так:

$$L_3 = k \log(s) + (1 - k) \log(1 - s), \quad (5)$$

де $k = 1$, при повороті x догори і 0 – донизу.

Загалом завдання першого етапу можна визначити за формулою:

$$\min_Y L = L_1 + \lambda_2 \cdot L_2 + \lambda_3 \cdot L_3, \quad (6)$$

де λ_2, λ_3 – параметри для збалансування різних втрат.

Після оптимізації виразу (6) отримано PCN, що може використовуватися для усування непотрібних вікон, для отримання невеликої кількості кандидатів. Для інших кандидатів спочатку вони оновлюються до нових обмежувальних полів, які потім регресуються. Насамкінець оновлені кандидати обертаються відповідно до прогнозованих RIP кутів.

Прогнозований RIP кут на першому етапі можна обчислити за формулою:

$$\theta_1 = \begin{cases} 0^\circ, & \text{якщо } s \geq 0,5 \\ 180^\circ, & \text{якщо } s < 0,5 \end{cases} \quad (7)$$

Зокрема, $\theta_1 = 0$ означає, що кандидат направлений догори, тому обертання не потрібне, інакше $\theta_1 = 180$, кандидат направлений вниз, і потрібно повернути його на 180 градусів. Діапазон RIP кутів змінюється від $[-180, 180]$ до $[-90, 90]$. На основі даних вертикальних облич обертаємо прогнозовані зображення з різними RIP кутами, утворюючи новий набір, що містить обличчя з RIP кутами до 360 градусів.

Протягом навчання використовують три види даних: позитивні, негативні та підозрювані зразки. Позитивними зразками є вікна, для яких обчислений коефіцієнт більший за 0,7; негативні – у яких коефіцієнт менший за 0,3; а підозрювані – для коефіцієнтів від 0,4 до 0,7. Перші два типи зразків визначають навчання для розрізнення облич та “не облич”. Позитивні та підозрювані зразки сприяють навчанню регресії та калібруванню обмежувального поля.

2 етап

На цьому етапі ще чіткіше відрізняють обличчя від “не обличчя”. У випадку грубої орієнтації використовують потрібну класифікацію діапазону RIP кутів, – $[-90, -45]$, $[-45, 45]$, $[45, 90]$. Калібрування обертання проводять із прогнозованим RIP кутом:

$$d = \arg \max_i s_i$$

$$\theta_2 = \begin{cases} -90^\circ, & \text{якщо } idn = 0, \\ 0^\circ, & \text{якщо } idn = 1, \\ 90^\circ, & \text{якщо } idn = 2 \end{cases} \quad (8)$$

Кандидати $v_j; enp$ бути обернені на $-90, 0$ або 90 градусів, відповідно. Після другого етапу діапазон RIP кутів зменшується від діапазону $[-90, 90]$ до $[-45, 45]$. При навчанні на другому етапі рівномірно обертаються початкові зображення в діапазоні $[-90, 90]$, а жорсткі негативні зразки відкидаються за допомогою навченої на першому етапі PCN. Позитивні та підозрювані зразки в діапазоні RIP кутів $[-90, -60]$, $[-30, 30]$, $[60, 90]$ відповідають ідентифікаторам 0, 1, 2 для калібрування.

3 етап

Після другого етапу всі кандидати калібруються до діапазону RIP кутів від -45 до 45 градусів. Тому на третьому етапі приймається остаточне рішення. Оскільки RIP кут зменшений до малого діапазону на попередніх етапах, то на третьому етапі намагаються спрогнозувати точні кути кандидатів обличчя на заміну грубих орієнтацій.

Наведемо декілька прикладів для обчислення RIP кутів. Кутову регресію RIP виконують від грубого до точного каскаду регресії. Під час обчислення рівномірно обертаємо початкові

зображення в діапазоні $[-45, 45]$ і фільтруємо жорсткі негативні зразки за допомогою навченого на другому етапі PCN.

Насамкінець, RIP кут кандидата можна отримати, об'єднавши прогнози на всіх трьох етапах:

$$\theta_{\text{RIP}} = \theta_1 + \theta_2 + \theta_3. \quad (9)$$

Створення та дослідження системи розпізнавання облич

Система розпізнавання (рис. 2) складається з таких об'єктів:

- клієнт;
- сервер;
- мобільний додаток.

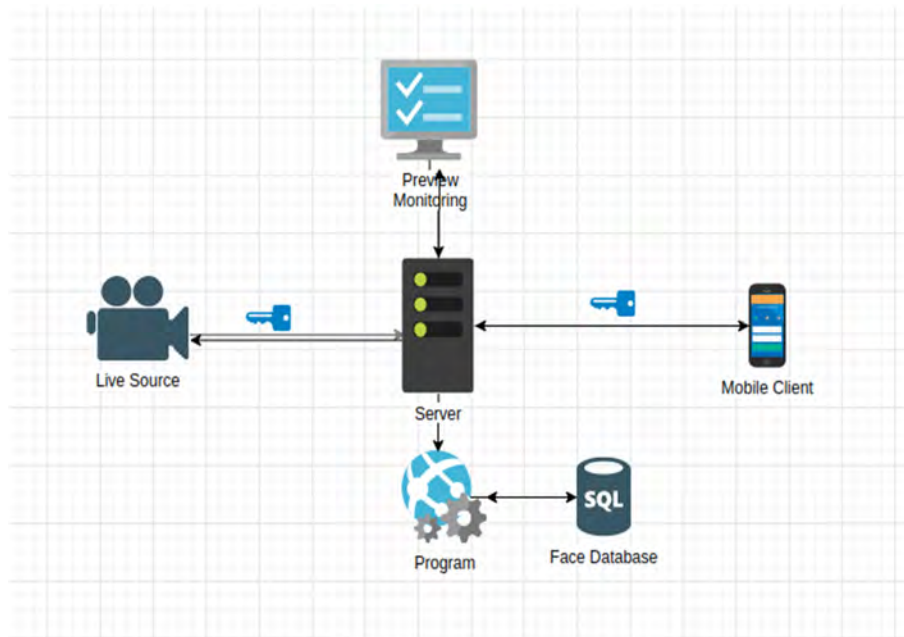


Рис. 2. Загальна структура системи

Клієнт – це програмне забезпечення. Клієнт може взаємодіяти з камерою, записувати відео і відправляти його на сервер в реальному часі. Основним призначенням сервера є обробка зображень з використанням алгоритмів виявлення облич, які містяться в базі даних. Також сервер підтверджує вибрані обличчя, взаємодіючи з базою даних. У випадку виявлення незареєстрованої особи надсилається сповіщення додатка, який запущений у мобільному клієнті.

На рис. 3 зображено діаграму використання розробленої системи, а на рис. 4 – діаграму використаної послідовності.

На діаграмі послідовності зображено чотири основні компоненти. Джерело запису клієнта формує відео і надсилає зображення серверу. Своєю чергою, сервер обробляє зображення і взаємодіє з базою даних для перевірки та виявлення відомих облич. У випадку наявності незареєстрованого обличчя сервер відправляє повідомлення мобільному пристрою. Насамкінець у випадку реєстрації нового обличчя отримана за допомогою алгоритму інформація надсилається серверу і зберігається в базі даних.

Ще одним важливим завданням розробленої системи виявлення зареєстрованих облич є необхідність раціонального зберігання даних. У розробленій системі використано реляційну модель даних, що використовує систему управління базами даних SQLite [5].

Загальна структура бази даних містить чотири таблиці. Перша з них є таблицею-словником, а інші три – таблицями-сутностями. Такий спосіб побудови таблиць дає змогу структурувати дані і забезпечити високу швидкість обробки даних для зареєстрованих зображень.

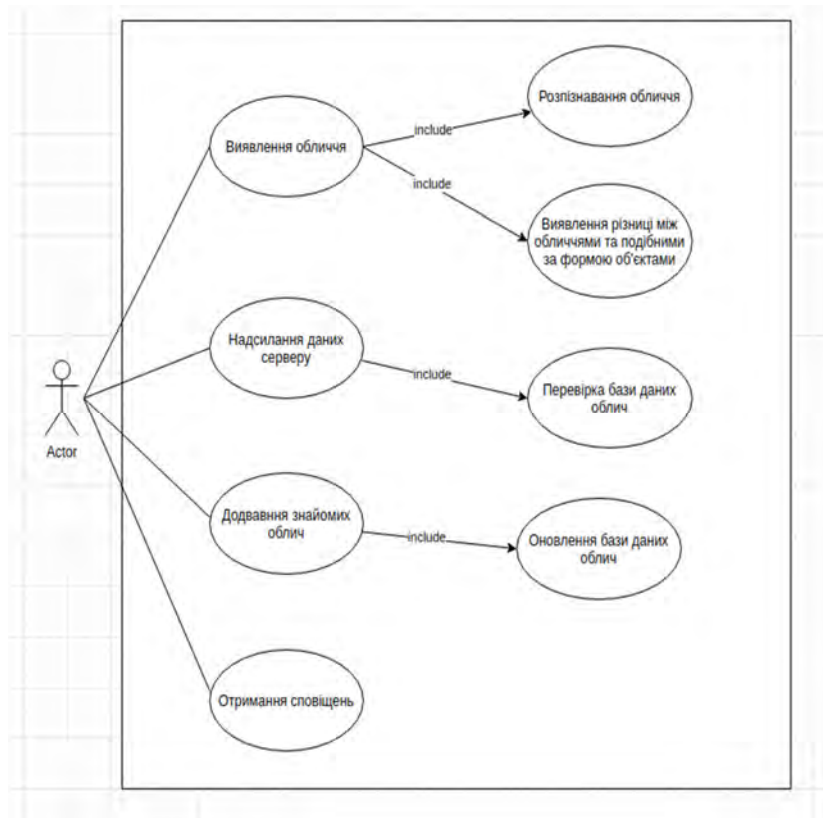


Рис. 3. Діаграма варіантів використання

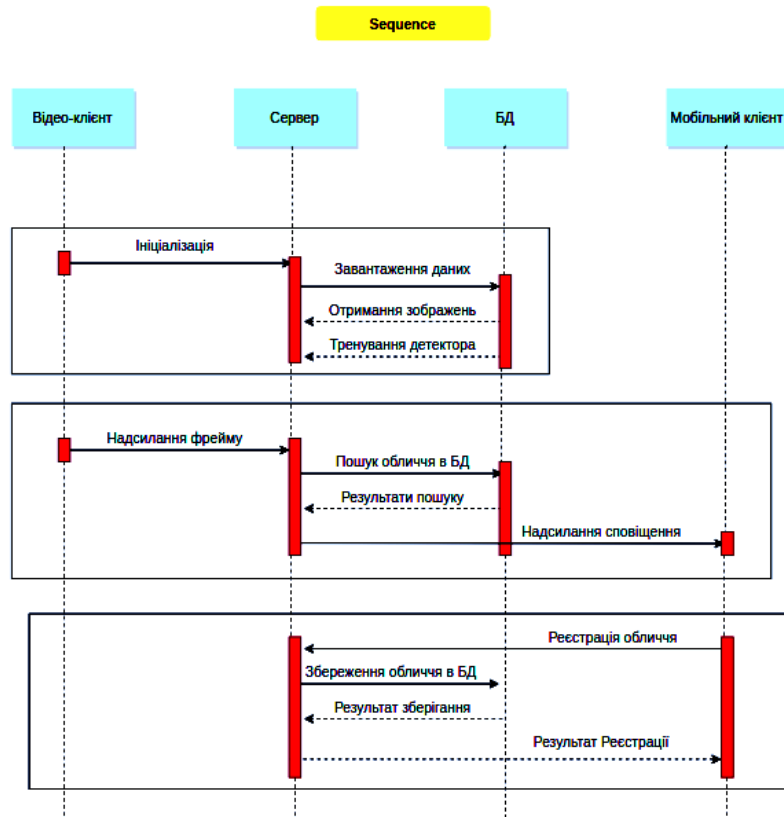


Рис. 4. Діаграма послідовності

На рис. 5 зображено таблиці бази даних “Images”, “Persons” та “RegisteredFaces”. У стовпцях таблиці міститься така інформація:

- Id – ідентифікатор унікального номера елемента в таблиці;
- path – шлях до файла зображення;
- data – дані зображення;
- name – оригінальне ім'я елемента, що однозначно ідентифікує назву певного елемента;
- imageId – Id зображення людини, що вказує на елемент таблиці Images;
- personId – Id людини, що вказує на елемент таблиці Persons.

Images		
PK	Id	(int)
	path	(text)
	data	(blob)

Persons		
PK	Id	(int)
	name	(text)
FK	imageId	(int)

RegisteredFaces		
PK	Id	(int)
FK	personId	(int)

Рис. 5. Таблиці бази даних системи розпізнавання облич

Для безпечного збереження даних та забезпечення їх швидкої обробки та взаємодії з базою даних використовують зберігання та обробка даних на сервері. Для системи розпізнавання створено ТСП сервер для ОС Linux з використанням мови програмування C++ [6, 7]. Нами використано системні засоби доступу до мережеских інтерфейсів (сокетів портів тощо). Для встановлення мережевого з'єднання та передавання даних мережею було розроблено класи TCPServer та TCPClient.

Для розроблення мобільного додатка використано середовище Android Studio, зокрема для інтерфейсу мобільного додатка – форми Android Studio (рис. 6).

Тестування основних компонентів системи

Для перевірки працездатності системи розпізнавання облич було проведено тестування основних частин системи. На рис. 7 наведено основні результати тестування, які показують, що проблем при виконанні досліджень немає.

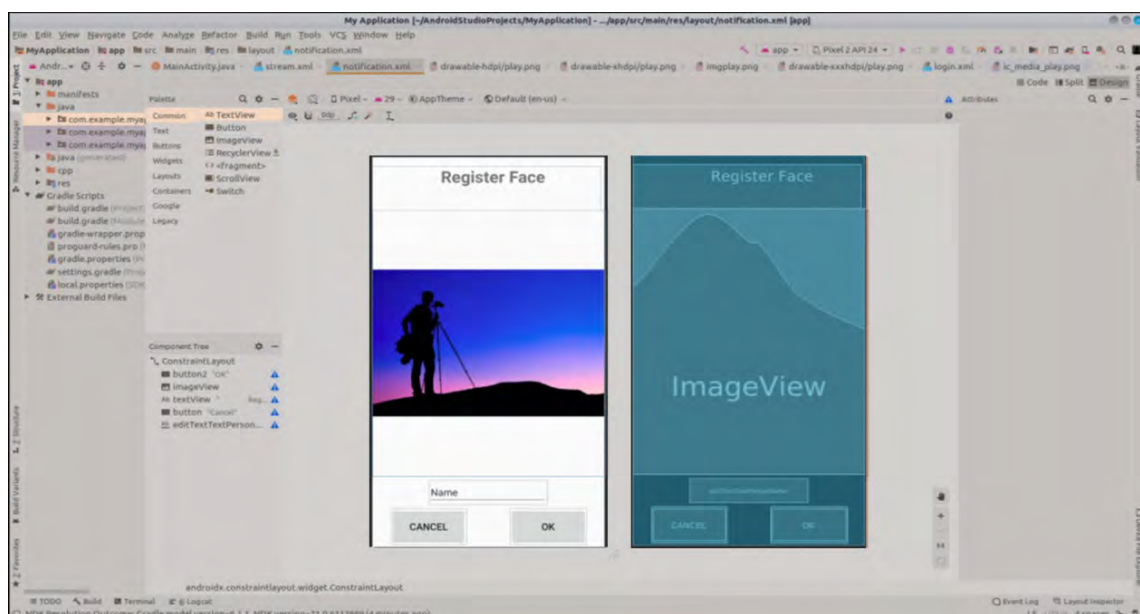


Рис. 6. Вікно редагування форм

Id	Summary	Pre-conditions	Description	Steps to reproduce	Expected result	Actual result
1	Тестування розпізнавання обличчя	Необхідна веб-камера	Тестування розпізнавання обличчя та позначення його на інтерфейсі	1) Запустити систему 2) Знаходитись перед камерою для розпізнавання	Обличчя виявлено та обведено рамкою на інтерфейсі користувача	Обличчя виявлено та обведено рамкою на інтерфейсі користувача
2	Тестування клієнт-серверного з'єднання	Необхідно мати інтернет з'єднання	Тестування клієнт-серверного з'єднання	1) Запустити сервер 2) Запустити клієнт 3) Перевірити результат з'єднання	З'єднання встановлено без помилок	З'єднання встановлено без помилок
3	Тестування відправки повідомлень між клієнтом та сервером	Необхідно мати інтернет з'єднання	Тестування відправки повідомлень між клієнтом та сервером	1) Запустити сервер 2) Запустити клієнт 3) Відправити повідомлення клієнту 4) Відправити повідомлення серверу	З'єднання встановлено, повідомлення доставлені без пошкодження даних	З'єднання встановлено, повідомлення доставлені без пошкодження даних
4	Тестування реєстрації обличчя	Необхідна веб-камера, та смартфон на Android	Тестування реєстрації обличчя	1) Запустити систему 2) Запустити мобільний додаток 3) Натиснути кнопку "Register Face" 4) Зробити фото 5) Натиснути кнопку "Ready"	Реєстрація виконана успішно, дані збережені на сервері	Обличчя виявлено та обведено рамкою на інтерфейсі користувача

Рис. 7. Test cases

Висновки

У результаті проведених досліджень систем штучного інтелекту, методів машинного навчання, каскадних моделей розпізнавання об'єктів та нейронних мереж встановлено, що найкраще використати для створення інформаційної системи розпізнавання людських облич метод на основі прогресивних мереж калібрування. Цей метод протягом незначної кількості кроків дозволяє виявити обличчя, які можуть бути розташовані як вертикально, так і повернуті на різні кути. Розроблено архітектуру системи та продемонстровано діаграми, що показують результати роботи. Розроблено програмне забезпечення системи, а саме базу даних на основі системи управління базами даних SQLite, сервер та десктопний клієнт на операційній системі Linux і мобільний додаток, що дозволило дослідити характеристики інформаційної системи. Також було протестовано систему розпізнавання людських облич та виправлено виявлені помилки.

Список літератури

1. Кашикарров А. П. Системы видеонаблюдения. Практикум / А. П. Кашикарров. К.: Феникс, 2014. 128 с.
2. Дамьяновски В. CCTV. Библия видеонаблюдения. Цифровые и сетевые технологии. пер. с англ. / В. Дамьяновски. М.: Ай-Эс-Эс Пресс, 2006. 480 с.
3. Xuepeng Shi, Shiguang Shan, Meina Kan, Shuzhe Wu, Xilin Chen. Real-Time Rotation-Invariant Face Detection With Progressive Calibration Networks / S. Xuepeng, S. Shiguang, K. Meina, W Shuzhe, C. Xilin // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018, pp. 2295–2303.
4. Girshick R. Fast R-CNN / R. Girshick R. // The IEEE International Conference on Computer Vision (ICCV). December 2015.
5. Кузин А. В. Базы данных / Кузин А. В., Левонисова С. В. 5-е изд. К.: Академия, 2012. 317 с.
6. Скотт Мейерс. Эффективное использование STL / М. Скотт. СПб.: Питер, 2002. 224 с.
7. Nuruzzaman Faruqi. Open Source Computer Vision for Beginners: Learn OpenCV using C++ in fastest possible way (2nd Edition) / F. Nuruzzaman. Kindle direct publishing, 2017.

Reference

1. Kashkarov A. P. *Sistemy videonablyudeniya. Praktikum* / A. P. Kashkarov. K.: Feniks, 2014. 128 s.
2. Damyanovski V. *CCTV. Bibliya videonablyudeniya. Tsifrovyye i setevyye tehnologii. Per. s angl.* / V. Damyanovski. M.: Ay-Es-Es Press, 2006. – 480 s.
3. Xuepeng Shi, Shiguang Shan, Meina Kan, Shuzhe Wu, Xilin Chen. *Real-Time Rotation-Invariant Face Detection With Progressive Calibration Networks* / S. Xuepeng, S. Shiguang, K. Meina, W Shuzhe, C. Xilin // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018, pp. 2295–2303.
4. Girshick R. *Fast R-CNN* / R. Girshick R. // *The IEEE International Conference on Computer Vision (ICCV)*. December 2015. <https://doi.org/10.1109/ICCV.2015.169>
5. Kuzin A. V. *Bazyi dannyih* / Kuzin A. V., Levonisova S. V. 5-e izdanie. K.: Akademiya, 2012. 317 s.
6. Skott Meyers. *Effektivnoe ispolzovanie STL* / M. Skott. SPb.: Piter, 2002. 224 s.
7. Nuruzzaman Faruqui. *Open Source Computer Vision for Beginners: Learn OpenCV using C++ in fastest possible way (2nd Edition)* / F. Nuruzzaman. – Kindle direct publishing, 2017.