


**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

Факультет комп'ютерних наук та кібернетики
Кафедра інтелектуальних програмних систем

Кваліфікаційна робота
на здобуття ступеня бакалавра
за спеціальністю 121 Інженерія програмного забезпечення
на тему:

**РОЗРОБКА ЗАСТОСУНКУ ДЛЯ РОЗПІЗНАВАННЯ ЛЮДСЬКОГО
ОБЛИЧЧЯ НА ЗОБРАЖЕННІ ПРИ СКЛАДНИХ УМОВАХ ВИДИМОСТІ**

Виконала студентка 4-го курсу
Василина-Катерина БОРИСЮК  (підпис)

Науковий керівник:
Кандидат фізико-математичних наук
Костянтин ЖЕРЕБ _____ (підпис)

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

Студент  (підпис)

Роботу розглянуто й допущено до захисту на
засіданні кафедри інтелектуальних програмних
систем

«25» травня 2022 р.,
протокол № 10
Завідувач кафедри
Олександр ПРОВОТАР _____ (підпис)

РЕФЕРАТ

Обсяг роботи: 20 сторінок

Кількість ілюстрацій: 12

Використані джерела: 16

РОЗРОБКА ЗАСТОСУНКУ ДЛЯ РОЗПІЗНАВАННЯ ЛЮДСЬКОГО ОБЛИЧЧЯ НА ЗОБРАЖЕННІ ПРИ СКЛАДНИХ УМОВАХ ВИДИМОСТІ

Метою роботи є дослідження існуючих засобів пошуку та ідентифікації облич, створення програмного засобу, який використовує найбільш влучні з них для фотографій поганої якості.

Методи розроблення: тестування існуючих програмних рішень, комбінування найбільш вдалих.

Об'єкт роботи: програмне забезпечення для виявлення і розпізнавання облич, спеціальні набори даних.

Інструменти розробки: мова python, бібліотеки TensorFlow, Albumentations, Labelme, середовище розробки PyCharm

Результати роботи: програмний модуль, здатний розпізнавати обличчя при гіршій якості зображень.

ЗМІСТ

ВСТУП	4
РОЗДІЛ 1. ВИЯВЛЕННЯ ОБЛИЧ НА ЗОБРАЖЕННІ	6
1.1 Швидкісне виявлення за допомогою каскадів простих ознак	6
1.2 Виявлення за допомогою гістограм орієнтованих градієнтів	7
1.3 Виявлення облич методами глибокого навчання	10
1.4 Виявлення орієнтирів на обличчі	11
1.4 Вибір кращого алгоритма для поставленої задачі	12
РОЗДІЛ 2. ІДЕНТИФІКАЦІЯ ОСОБИ НА ЗОБРАЖЕННІ	13
2.1 Розпізнавання за геометричними ознаками	13
2.1.1 Перша відома автоматична система розпізнавання	13
2.1.2 Метод Eigenfaces	15
2.2 Методи штучного інтелекту для розпізнавання	16
РОЗДІЛ 3. СПОТВОРЕННЯ ЗОБРАЖЕНЬ ПРИ ПОГАНІЙ ВИДИМОСТІ	18
РОЗДІЛ 4. ВИБІР ТЕХНОЛОГІЙ ДЛЯ РЕАЛІЗАЦІЇ	21
ВИСНОВКИ	22
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	24

ВСТУП

Оцінка сучасного стану об'єкта дослідження або розробки.

Пошук обличчя та ідентифікація особи на фото - давно відомі і розглянуті багатьма задачами, багато різних підходів до їх вирішення були випробувані для вирішення практичних проблем. Очевидно перспективний напрям розвитку в цій сфері - отримання кращих результатів при гіршій якості вхідних даних.

Актуальність роботи та підстави для її виконання.

Фотографії, на яких треба розпізнати обличчя не завжди зняті в хорошій якості або людина, яка користується пристроєм, якому треба її розпізнати, знаходиться в умовах поганого освітлення. Чим гірші умови, тим краще під конкретну задачу повинен бути налаштований алгоритм і поступові покращення, пошук кращих комбінацій допоміжних засобів - частина глобального прогресу в задачі розпізнавання облич. Чим краще ця задача виконується комп'ютером, тим легша робота і життя людини.

Мета й завдання роботи.

Метою кваліфікаційної роботи є створення програмного засобу для кращого розпізнавання облич при поганій якості. Для досягнення поставленої мети сформульовано наступні завдання:

- Вивчити специфіку і деталі поставленої задачі.
- Розглянути існуючі алгоритми, вибравши оптимальні комбінації і модифікації.
- Об'єднати вибрані складові в програмний засіб.
- Відібрати оптимальні для відлагодження засобу дані.

Об'єкт, методи й засоби дослідження та розробки.

Ідентифікація особи складається з різних етапів, можна виділити виявлення обличчя і, власне, ідентифікацію, вони розглянуті в роботі окремо.

Для виявлення обличчя розглядаються алгоритми, що зустрічаються в ужитку останні 5 років, включаючи більш прості і більш складні, різні за швидкістю, потребою в ресурсах і попередній підготовці, більш та менш загальної дії.

Для ідентифікації розглядаються принципово різні підходи, включаючи як ті, що повністю покладаються на існуючі знання про людські обличчя, так і ті, що використовуються в цій задачі лише як частковий випадок, будучи засобом вирішення широкого кола задач. Також є загальні засоби, чиї модифікації пристосовані окремо до облич.

Для того, аби отримати дані зі специфічним запитом на погану якість, розглядаються як відомі датасети для конкурсів і освітніх програм, так і матеріали з власної роботи, пов'язаної з фотографією. Також використовується штучне розширення датасетів за допомогою спотворень і випадкової обробки програмним шляхом.

Зрештою, розглядаються рекомендовані фреймворки, бібліотеки і можливості конкретних бібліотек, рекомендовані для реалізації кожного із вищеописаних етапів.

Можливі сфери застосування.

Перш за все робота створена аби допомогти виявляти учасників великих спільних фотографій при зйомці публічних подій. Та сфера застосування потенційно набагато ширша і включає верифікацію при логіні, оплаті, підтвердженні дій, розблокуванні пристрою.

РОЗДІЛ 1. ВИЯВЛЕННЯ ОБЛИЧ НА ЗОБРАЖЕННІ

Першим етапом розпізнавання обличчя на зображенні є виявлення цього обличчя. Ми скористаємось методами машинного навчання, розглянувши найпопулярніші підходи і вибравши ті модифікації, які краще всього справляються з поганим освітленням і видимістю.

Узагальнено багато відомих алгоритмів виявлення об'єктів зображення працюють за наступною схемою: по зображенню ковзає т.зв. "вікно"(прямокутна область, виділена на зображенні), всередині якого вираховується вектор з певних ознак і на основі значень цього вектора робиться висновок: вікно є зображенням бажаного об'єкта чи ні. Далі розглянемо різні підходи до реалізації деталей такого алгоритма.

При пошуку бібліотек мови Python, які допомагають з виявлення об'єктів, в популярній літературі найчастіше зустрічаються `openCV cv2` і `facial_recognition`.

1.1 Швидкісне виявлення за допомогою каскадів простих ознак

`openCV cv2` використовує алгоритм Віоли-Джонса, винайдений і описаний в [8]. Цей алгоритм використовується в дуже багатьох побутових задачах, наприклад при автоналаштуванні камери, зокрема завдяки своїй простоті і швидкодії, яка при цьому не заважає хорошим результатам.

Для видобутку ознак в кожному пікселі використовуються фільтри, подібні до згорткових. Використовуються три види фільтрів(див. Рис.1). Фільтр двох прямокутників обчислює різницю між сумою пікселів у двох прямокутних областях однакового розміру, форми, суміжні горизонтально або вертикально. Фільтр трьох прямокутників обчислює суму в двох зовнішніх прямокутниках, що віднімається від суми в центральному прямокутнику. Нарешті, функція чотирьох прямокутників обчислює різницю між діагональними парами прямокутників.

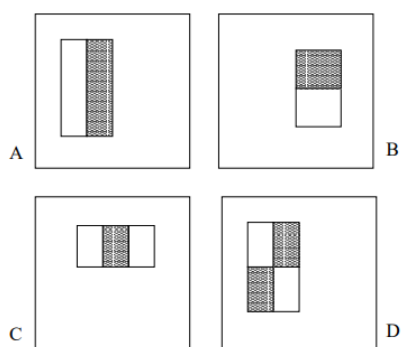


Рис.1 Наар-фільтри

Роботу, що описує алгоритм вирізняють три ключові внески. Перший – це впровадження нового представлення зображення під назвою «Інтегральне зображення», яке дозволяє дуже швидко обчислювати функції, які використовуються детектором. В такому представленні значення кожного пікселя обчислюється як $ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$. Другий — це алгоритм навчання, заснований на AdaBoost, який вибирає невелику кількість критичних візуальних ознак із більшого набору та дає надзвичайно ефективні класифікатори. Третій внесок — це метод об'єднання дедалі складніших класифікаторів у «каскад», який дозволяє швидко відкидати фонові області зображення, витрачаючи більше обчислень на перспективні об'єктоподібні області. Каскад можна розглядати як специфічний для об'єкта механізм фокусування уваги, який забезпечує статистичні гарантії того, що відкинуті області навряд чи містять об'єкт, що цікавить.

1.2 Виявлення за допомогою гістограм орієнтованих градієнтів

`facial_recognition` побудована на базі C++ `DLib library`. `DLib` використовує об'єкт `object_detector`, налаштований на знаходження людських облич, які дивляться більш-менш у бік камери. Він створюється за допомогою об'єкта `scan_fhog_pyramid`, який є інструментом для запуску класифікатора з плаваючим вікном фіксованого розміру над пірамідою зображення.

Піраміди зображень - це, просто кажучи, серія зображень з різною роздільною здатністю, що поєднуються разом [6]. Для отримання кожного

наступного “поверху” піраміди вихідне зображення щоразу зменшується за різним алгоритмом, залежно від задачі. Ці зображення збираються разом, утворюючи «піраміду». Кожне зображення в пірамідах зображень називається «октавою».

Що стосовно класифікатора, задачу знаходження об’єкта можна розглядати як задачу бінарної класифікації. Нам, наприклад, потрібно виявити, чи є поточне вікно зображенням обличчя людини, чи ні. При розгляді задач бінарної класифікації, можна часто натрапити на Support Vector Machine, вперше описану в [7]. Машина концептуально реалізує таку ідею: вхідні вектори нелінійно відображаються в просторі об’єктів дуже високої розмірності. У цьому просторі ознак будується лінійна поверхня прийняття рішення. Спеціальні властивості цієї поверхні забезпечують високу узагальнювальну здатність навчальної машини. В роботі такі мережі показали високу здатність до узагальнення.

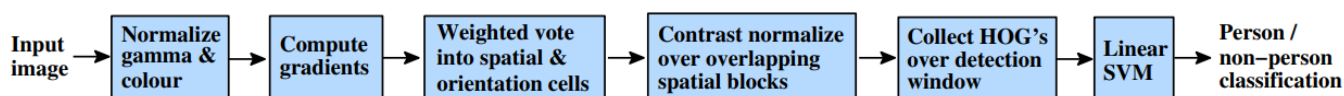


Рис.2 Схема обробки зображення методом, описаним в [4]

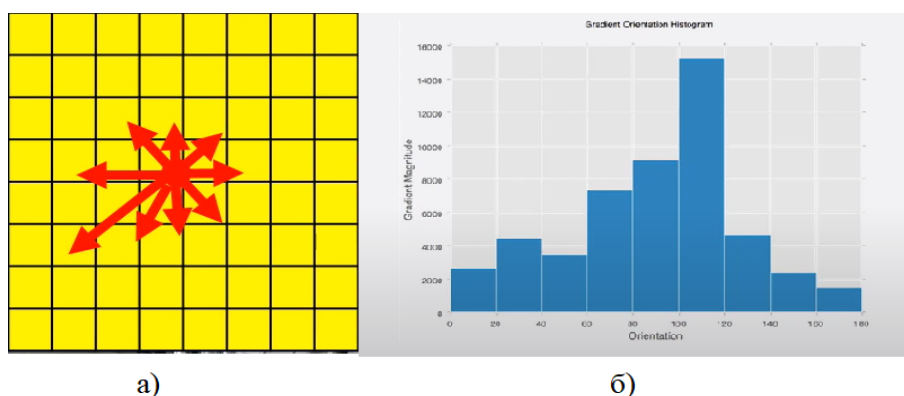


Рис.3 Приклад

НОГ для комірки 8x8 (а) - у вигляді просторової діаграми, б) - у вигляді стовпчикової діаграми)

В `facial_recognition`, класифікатор ковзає по піраміді НОГ, як описано в [4]. Схематично алгоритм показаний на Рис.2. Основна ідея полягає в тому, що зовнішній вигляд і форму об’єкта часто можна досить добре охарактеризувати

розподілом градієнтів інтенсивності або напрямків країв, навіть без точного знання відповідного градієнта або положень країв. На практиці це реалізується шляхом поділу вікна зображення на невеликі просторові області («комірки»), для кожної комірки створюється одновимірна гістограма напрямків градієнта або орієнтації країв від пікселів комірки. Схематичне зображення діаграми наведено на Рис.3. Далі значення усіх гістограм об'єднуються. Для кращої інваріантності до освітлення, затінення тощо корисно також контрастно нормалізувати локальні результати перед їх використанням. Це можна зробити шляхом накопичення міри «енергії» локальної гістограми на дещо більших просторових областях («блоках») і використання результатів для нормалізації всіх клітин у блоці. Ми називатимемо нормалізовані блоки дескрипторів гістограмами орієнтованого градієнта (HOG). Замощення вікна виявлення щільною (фактично, такою що перекривається) сіткою дескрипторів HOG і використання комбінації цих HOG дескрипторів у традиційному віконному класифікаторі на основі SVM дає ланцюжок виявлення людини на зображенні.

Переваги представлення за допомогою HOG зокрема в тому, що воно зчитує структуру країв, яка, в свою чергу, майже вичерпно описує локальні форми, і робить це в локальному представленні з легко контрольованим ступенем інваріантності до локальних геометричних і фотометричних перетворень. Наприклад, повороти не значно впливають на результат, якщо вони невеликі. Водночас автори підмічають, що для хороших результатів, зображення не повинно бути надто розмитим.

В бібліотеці DLіb гістограми HOG обраховуються в покращеному таким чином, як описано в [5] вигляді. Замість 36-вимірних гістограм орієнтованих градієнтів там представлено альтернативний 13-вимірний набір ознак, який фіксує по суті ту ж інформацію. Цей низькорозмірний набір функцій також розширений таким чином, щоб включити як контраст-чутливі, так і нечутливі до контрасту функції, що призводить до 31-вимірного вектора ознак і на тестових даних відпрацьовує продуктивніше. Автори також підмічають, що

простіші моделі можуть краще працювати на практиці, оскільки менше страждають від труднощів під час навчання.

1.3 Виявлення облич методами глибокого навчання

У вищевказаних алгоритмах за етап знаходження ознак і класифікацію відповідали різні, доволі прості і давно відомі моделі/алгоритми. Натомість зі збільшенням обчислювальних потужностей стало можливим ефективно видобувати ознаки через глибоку згорткову модель і класифікувати об'єкти по цих ознаках через глибоку щільну модель. На перший погляд показники, отримані найуспішнішими моделями, представленими зокрема на конкурсі ImageNet відрізняються не надто сильно. Також в багатьох статтях, де розглядаються різні підходи до тренування відповідних моделей, в основу цих моделей іноді покладена архітектура VGG16(див. Рис.4), дуже сильна в класифікації.

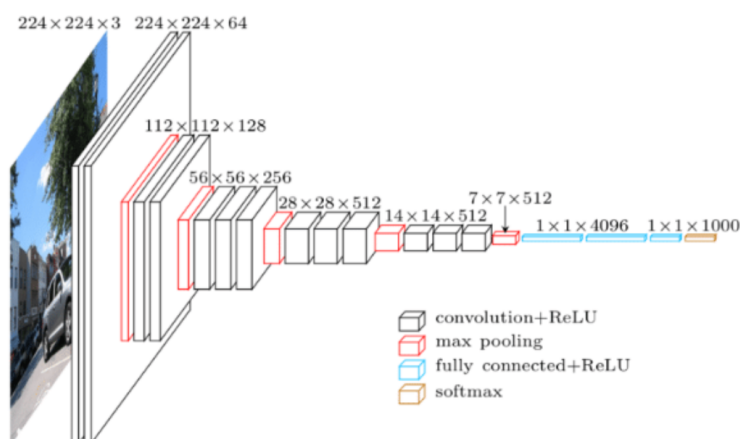


Рис.4 Архітектура VGG16

Автори цієї архітектури в [10] зазначають, що оцінили дуже глибокі згорткові мережі (до 19 вагових шарів) для класифікації(яка є частиною процесу виявлення облич) зображень і продемонстрували, що глибина представлення позитивно впливає на точність класифікації, і що зразкова продуктивність на датасетах ImageNet може бути досягнута за допомогою звичайної архітектури ConvNet із значно збільшеною глибиною. Такі моделі також добре узагальнюються для широкого кола завдань і наборів даних, збігаючи або перевершуючи складніші конвеєри розпізнавання, побудовані на менш глибоких представленнях зображень.

Автори [12] називають ключовим відкриттям те, що можна об'єднати згорткові нейронні мережі високої ємності (CNN, таку як, VGG16) та оцінку “знизу-вгору” регіону знаходження потенційного об’єкта, для якісного виявлення та сегментування об’єктів. Представлена там система виявлення R-CNN об’єктів складається з трьох модулів. Перший генерує незалежні від категорії пропозиції регіонів. Ці пропозиції визначають набір потенційних об’єктів, доступних нашому детектору. Другий модуль — це велика згорткова нейронна мережа, яка витягує вектор ознак фіксованої довжини з кожної області. Третій модуль являє собою набір лінійних SVM, специфічних для класу. Згодом у [13] було запропоновано фреймворк Fast R-CNN, який модернізує R-CNN, роблячи навчання і обрахунок швидшими, а результати точнішими.

Застосувавши ще одну модифікацію, Faster R-CNN для виявлення облич, автори [11] роблять висновок: хоч її і розроблено для загального виявлення об’єктів, вона демонструє вражаючу продуктивність для розпізнавання обличчя після перенавчання на відповідному датасеті для виявлення обличчя.

1.4 Виявлення орієнтирів на обличчі

Так як деякі алгоритми чутливі до поворотів обличчя в різних площинах на фото, в багатьох з них зустрічається етап геометричного перетворення зображення з метою розміщення конкретних деталей обличчя в конкретних позиціях зображення. Подекуди використовуються складні нелінійні перетворення, але часто достатньо лише афінних. Частиною цього перетворення є, власне, виявлення деталей обличчя, наприклад, точок, запропонованих на Рис.5, яке, до речі саме по собі є важливою задачею, зокрема дуже поширеною для використання в розважальних цілях. Питання поганій освітленості в цьому випадку не менш важливе.

Так як методи, перераховані вище, в цілому добре узагальнюються на пошук не лише облич, а об’єктів загалом, вони також використовуються для пошуку деталей обличчя.

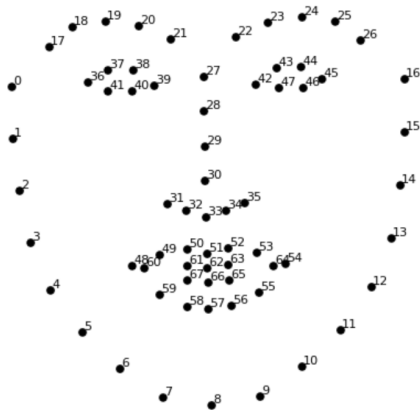


Рис.5 Варіант опорних точок на обличчі,

запропонований в [16]. Автор зображення - Brandon Amos.

1.4 Вибір кращого алгоритма для поставленої задачі

Класифікація в усіх розглянутих алгоритмах відбувається методами машинного навчання і відносно схожим чином. Перші два із розглянутих алгоритмів видобувають ключові ознаки із зображень більш контрольованим способом, ніж глибока нейронна мережа, видобуток ознак в цих випадках описаний детально і структуровано. З іншого боку, глибока нейронна мережа має кращу здатність до узагальнення завдяки великій кількості параметрів і багатьом шарам нелінійних перетворень.

Ресурси для тренування під конкретну задачу на конкретних даних, яких вимагає глибока нейронна мережа значно перевищують ресурси, яких потребує виявлення за допомогою каскадів пристих ознак і гістограми направлених градієнтів. Отримання відповіді для кожного конкретного випадку при цьому відрізняється по часу не настільки ж значно, але метод Нааг-каскадів особливо швидкий, тому досі використовується в додатках, особливо тих, де пошук обличчя потрібен в реальному часі без затримок.

Поставлена задача пріоретизує якість отриманого результату, а не швидкодію, тому в практичній частині першим ділом розглядається пошук облич за допомогою глибокої нейронної мережі.

РОЗДІЛ 2. ІДЕНТИФІКАЦІЯ ОСОБИ НА ЗОБРАЖЕННІ

Розпізнавання облич є легким завданням для людей. Експерименти в [2] показали, що навіть одно-триденні немовлята здатні розрізнити відомі обличчя. Але так як процес розпізнавання облич людьми вивчений поки що недостатньо, важко сказати, наскільки складною аналогічна задача є для комп'ютера. Чи використовуються внутрішні риси (очі, ніс, рот) чи зовнішні риси (форма голови, лінія волосся) для успішного розпізнавання обличчя? Як ми аналізуємо зображення і як мозок його кодує? Девід Х'юбел і Торстен Візель показали, що наш мозок має спеціалізовані нервові клітини, які реагують на конкретні місцеві особливості сцени, такі як лінії, краї, кути або рух. Оскільки ми не бачимо світ як розрізнені шматки, наша зорова кора повинна якимось чином об'єднати різні джерела інформації в корисні шаблони. Автоматичне розпізнавання обличчя полягає в тому, щоб витягти ці значущі риси із зображення, помістити їх у корисне представлення та виконати певну класифікацію для них.

2.1 Розпізнавання за геометричними ознаками

Розпізнавання обличчя на основі геометричних особливостей, мабуть, є найбільш інтуїтивним підходом.

2.1.1 Перша відома автоматична система розпізнавання

Одна з перших автоматизованих систем розпізнавання обличчя була описана в [1]: маркерні точки (положення очей, вух, носа, ...) використовувалися для побудови вектора ознак (відстань між точками, кут між ними, ...). Розпізнавання проводилося шляхом розрахунку евклідової відстані між векторами ознак прикладу та еталонного зображення. Під час тестування системи 15 з 20 людей були правильно ідентифіковані, спираючись виключно на вимірювання, отримані за допомогою комп'ютера. Це названо автором як перший успіх машинної ідентифікації обличчя людини на фото.

Такий метод за своєю природою стійкий до змін освітленості, що могло б зробити його дуже вдалим для цієї роботи, але має величезний недолік: точна реєстрація маркерних точок складна навіть з допомогою найсучасніших

алгоритмів. Деякі з останніх робіт з геометричного розпізнавання обличчя були виконані в [3]. Люди на зображеннях в тестовому датасеті були зняті обличчям до камери, при комбінації денного та штучного світла, знімки зроблені без жодних значних спотворень. В роботі згадано, що для успішного розпізнавання людини на фото за геометричними даними набір ознак повинен задовольняти наступним вимогам:

- оцінка повинна бути максимально простою;
- залежність від умов освітлення повинна бути якомога меншою;
- залежність від невеликих змін виразу обличчя повинна бути невеликою;
- інформаційний зміст повинен бути якомога вищим.

Для виявлення ознак було використано вертикальні та горизонтальні інтегральні проекції, формули яких наступні:

$$V(x) = \sum_{y=y_1}^{y_2} I(x, y) \quad H(y) = \sum_{x=x_1}^{x_2} I(x, y)$$

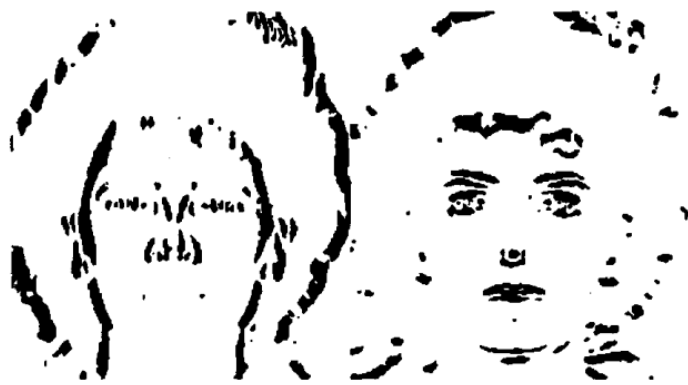


Рис.6. Проекції для виявлення

ознак

В результаті отримувались зображення вигляду Рис.6, на яких чітко виділені геометричні ознаки завдяки перепадам яскравості в областях цих ознак, що не залежить від освітлення, а лише від різниці між яскравістю в областях конкретних контурів. Далі зображення порівнювали з антропометричним шаблоном, завдяки чому отримувались конкретні виміри.

Був використаний 22-вимірний вектор ознак, і експерименти з великими наборами даних показали, що самі геометричні об'єкти можуть не нести

достатньо інформації для розпізнавання обличчя. Авторами зроблено висновок, що розпізнавання за допомогою вектора геометричних об'єктів може бути корисним для невеликих баз даних або як етап відбору для більш складних стратегій розпізнавання. До того ж результативність використаних методів надто сильно залежить від якості як опорного, так тестового зображення і умов, в яких воно зроблене.

2.1.2 Метод Eigenfaces

Пропонуючи Eigenfaces - власний метод для розпізнавання, автори [15] стверджують, що більша частина попередніх робіт ігнорувала питання про те, які саме аспекти обличчя важливі для ідентифікації. Такі риси можуть бути або не бути безпосередньо пов'язані з нашим інтуїтивним уявленням про риси обличчя, такі як очі, ніс, губи та волосся.

У метод Eigenfaces зображення обличчя представляється як точка з багатовимірною простору зображень, і шукається представлення меншої розмірності за допомогою т.зв. "власних облич" (Eigenfaces, приклад набору яких з оригінальної роботи зображено на Рис.12) Завдяки цьому класифікація стає легкою. Підпростір нижчої розмірності шукається за допомогою аналізу головних компонентів (Principal component analysis), який ідентифікує осі з максимальною дисперсією.

Розпізнавання здійснюється шляхом проектування нового зображення в підпростір власних облич, а потім класифікація обличчя шляхом порівняння його положення в просторі обличчя з положеннями відомих осіб. Автоматичне навчання та згодом розпізнавання нових облич є практичним у цьому контексті. Розпізнавання в дуже різноманітних умовах досягається шляхом навчання обмеженій кількості характерних зображень (наприклад, «прямий» вид, вид в півоберта і вид в профіль).



Рис.12 Набір Eigenfaces

Цей підхід має переваги перед іншими схемами розпізнавання облич у швидкості та простоті, здатності до навчання та нечутливості до невеликих або поступових змін у зображенні обличчя. Але такий вид перетворення не враховує жодних міток класів, наприклад в ситуації, коли дисперсія генерується із зовнішніх джерел, наприклад світла. Осі з максимальною дисперсією взагалі не обов'язково містять визначальну інформацію, тому класифікація ускладнюється, хоча для боротьби з цим знайдені допоміжні перетворення.

2.2 Методи штучного інтелекту для розпізнавання

У випадку облич, які втратили певні деталі через недоосвітленість і дивляться не прямо в камеру дуже важливою стає здатність алгоритма розпізнавання до узагальнення, адже далеко не на всіх прикладах конкретні знання про людське обличчя допомагають з його розпізнаванням.

З цією метою ми розглянемо модель, запропоновану в [16], адже її завдання формулюється достатньо узагальнено - виявлення, чи належать два зображення одному класу, чи ні, тобто верифікація. Схема моделі представлена на Рис.7

Для цього використовуються дві згорткові моделі з однаковими вагами, що в результаті дає однаковий набір ознак для двох зображень, які тестуються. Цей набір ознак далі попарно віднімається і кінцевий результат вирішується на основі нелінійно перетвореної зваженої суми цих різниць.

Згорткові моделі в процесі навчання знаходять найбільш інформативні ознаки, які в наявному датасеті вказують на різницю між обличчями двох різних людей.

Для тренування і тестування такої мережі потрібна підмножина декартового квадрата множини з багатьма (в т.ч. темними і неякісними) фотографіями багатьох різних людей.

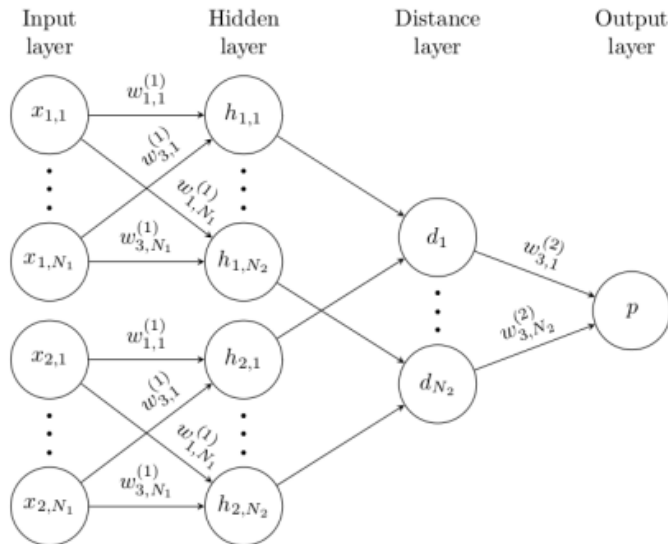


Рис.7 Примітивна Siamese

Neural Net з одним щільним шаром замість згортки

РОЗДІЛ 3. СПОТВОРЕННЯ ЗОБРАЖЕНЬ ПРИ ПОГАНІЙ ВИДИМОСТІ

При недосконалому обладнанні чи надто поганих умовах зйомки деталі на фотографіях спотворюються, втрачається точність. Сучасні засоби постобробки дозволяють, тим не менше, уникнути багатьох з них, так як зображено на Рис.8. Наприклад, аберації які виникають через фізичні характеристики лінз: бочкоподібні аберації на краях(витягнення об'єктів, близьких до краю) чи хроматичні аберації(різниця в заломленні різних кольорів, яка викликає появу кольорових обідків навколо об'єктів). Також, наприклад, в [14] запропонований алгоритм відновлення витягнутого зображення завдяки існуючим знанням про симетрію обличчя і положення частин обличчя один відносно одного.

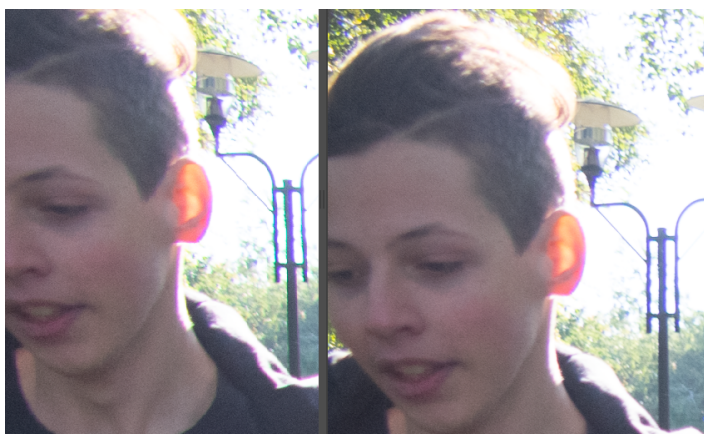


Рис.8 Корекція аберацій

засобами від компанії Adobe

Крім цього, при збереженні зображень в достатньо інформативних форматах типу RAW, де записуються значення яскравості кожного пікселя без стиснення і багато важливих метаданих, при постобробці можливо відобути деталі, непомітні на оригінальному зображенні, бо, наприклад, різниця яскравості сусідніх пікселів надто мала, але стає помітнішою при її масштабуванні.

Далі наведені фактори, які погіршують якість зображень і які варто враховувати, розширюючи датасет для тренування і валідації. Ці фактори впливають на якість зображення в момент зйомки, їх часто складно, або і неможливо виправити при пост-обробці, тому є сенс шукати способи

працювати з ними вже на етапі видобутку корисної інформації з зображення, використання його для якихось цілей.

Перший і найбільш очевидний - недостатня освітленість. По-перше, значення яскравості пікселів дискретні. При недостатньому освітленні різні області об'єкта менше відрізняються по яскравості, ніж могли б при нормальному освітленні. Якщо різниця яскравості двох сусідніх регіонів зображення менша, ніж "ціна поділки" шкали яскравості пікселів, ці регіони будуть відображені однаково і будуть втрачені деталі об'єкта, як на Рис.9. Ця ситуація виникає в темних приміщеннях, в темну пору дня і, як підмічено авторами [14] при зйомці з дуже великою фокусною відстанню. Не вдаючись до методів штучного інтелекту, яскравість зображення можна збільшити лише збільшивши отвір діафрагми, витримку чи чутливість матриці. При великих фокусних відстанях отвір діафрагми збільшити надто складно, адже це залежить від фізичних розмірів об'єктива.

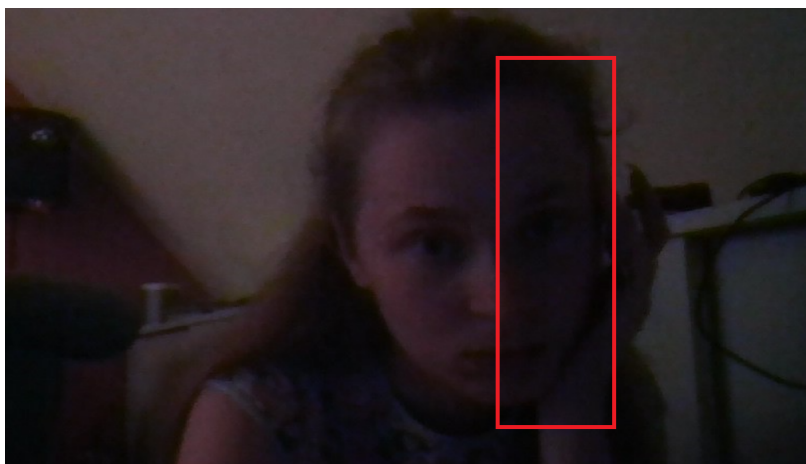


Рис.9 Втрата деталей в

темряві

При збільшенні витримки на рухомих об'єктах з'являється т.зв motion blur(розмиття в русі), через який можуть спотворюватись деталі об'єктів. Гіпертрофований і виділений кольором приклад цього ефекту наведений на Рис.10. Так як за час експозиції об'єкт встигає зміститись, світло від його деталей потрапляє на різні місця сенсора одночасно. Також в деяких камерах шторки, що закривають сенсор, під час відкриваються поступово, спочатку експонуючи верх кадра, потім низ.



Рис.10 Гіпертрофований

motion blur

Зрештою, при підвищенні чутливості сенсора, на нього можуть впливати сторонні чинники, внесок яких при великій чутливості був незначним. Таким чином деякі пікселі сенсора передаватимуть не той колір і яскравість, який отримали. При достатньо малій роздільній здатності це може спотворити деталі зображення, як очі героїв фото на Рис.11.



Рис.11 Спотворення

деталей обличчя через шум та затіненість.

РОЗДІЛ 4. ВИБІР ТЕХНОЛОГІЙ ДЛЯ РЕАЛІЗАЦІЇ

Як мову для реалізації вибрано Python, адже для неї створено багато бібліотек, що максимально спрощують процес розробки і реалізують відомі алгоритми.

Так як велику частину роботи складає побудова і тренування нейронних мереж, в роботі використана бібліотека TensorFlow. Вона популярна зокрема завдяки наявності великої кількості засобів для вирішення суміжних задач.

Окрім функціоналу для побудови нейронних мереж, бібліотека надає інструменти для завантаження та підготовки даних в пакеті TensorFlow Data, в т.ч. у вигляді pipeline, що дозволяє обробити надто великі об'єми даних, завантажуючи їх по черзі і за потреби кешуючи. В пакеті є досить багато функціональних інструментів для модифікації даних.

Для побудови і тренування нейронних мереж використовуємо TensorFlow Keras Functional API. В рамках нього можна побудувати будь-яку мережу, як ациклічний граф, в т.ч. мережі з кількома різними входами і виходами(наприклад, коли потрібні два зображення на вхід замість одного). Підхід Keras полягає в тому, щоб давати розробнику можливість налаштовувати деталі роботи функцій, поступово відкриваючи нові рівні складності, тому в окремих аспектах роботи можна мати дуже сильний контроль і гнучкість.

Для формування тренувального датасету залучені відомі датасети як Labeled Faces in The Wild, а також фотографії з власної роботи. Для позначення обмежуючих прямокутників з людськими обличчями на фото використовується бібліотека Labelme, яка пропонує дуже зручний інтерфейс з подальшим збереженням усіх анотацій для зображення в запропонований файл. Для кращої точності роботи мереж і більшої пристосованості до погано освітлених кадрів, було як вибрано такі кадри з готових датасетів, так і проведено розширення наявних за допомогою бібліотеки Albumentations, яка пропонує потокові інструменти для розширення набору даних.

ВИСНОВКИ

Розпізнавання людей складається з ідентифікації обличчя на фото і його розпізнавання. Більшу частину роботи - огляд існуючих статей на обидві теми і порівняння ефективності підходів в умовах поганої якості і освітлення, вибір надавався на користь якості, а не швидкодії.

Зображення може мати погану якість внаслідок недоліків камери чи певним чином підібраних налаштувань. В багатьох випадках цього неможливо уникнути і спотворення неможливо виправити в пост-обробці, тому на них треба звертати увагу, як на невід'ємну ознаку вхідних даних. Ці спотворення або стирають деталі зображення, або додають нових, природньо невластивих людському обличчю. Їх можна відтворити при підготовці датасета для тренування моделей.

Виявлення обличчя складається з отримання важливих ознак із зображення і класифікації набору цих ознак як обличчя/не обличчя. Існують методи, які отримують ці ознаки, спираючись на існуючі знання про об'єкти, що шукаються або просто обличчя. Методи ж глибокого навчання не опираються на заздалегідь відомі знання, а натомість отримують закономірності з існуючих даних. Тому обрані саме вони, адже при випадкових спотвореннях деякі тривіальні закономірності в даних можуть порушитись.

Для ідентифікації вибрано дуже узагальнену стратегію верифікації, яка винайдена для максимально широкого кола класів зображень, починаючи з рукописних символів і закінчуючи фотографіями реальних об'єктів. Це реалізовано за допомогою глибокої Siamese Neural Network, яка має дві однакові згорткові мережі в основі і щільні шари, які рахують різницю ознак, отриманих цими мережами.

Із фреймворків так бібліотек обрані ті, що стають в нагоді при роботі з глибоким навчанням: Tensorflow Keras Functional API завдяки можливості працювати зі складними і розгалуженими архітектурами нейронних мереж, TensorFlow.Data завдяки можливості ефективно і легко потоково обробляти величезні об'єми даних. Для підготовки даних обрано labelme для анотацій

завдяки зручному інтерфейсу, що особливо важливо при потребі вручну підписати великі об'єми даних. Також Albumentations для розширення датасету шляхом постобробки зображень, аби вміло використати здатність глибоких мереж до узагальнення на основі отриманих даних.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Picture processing system by computer complex and recognition of human faces. /Takeo Kanade - Японія, Kyoto University, 1974.
2. Newborns' face recognition: Role of inner and outer facial features. Child development/Chiara Turati, Viola Macchi Cassia, Francesca Simion, and Irene Leo/ 77(2):297–311, 2006
3. Face recognition through geometrical features. In Computer Vision/Roberto Brunelli and Tomaso Poggio/ECCV'92, pages 792–800. Springer, 1992.
4. Histograms of Oriented Gradients for Human Detection/Navneet Dalal and Bill Triggs/ CVPR 2005
5. Object Detection with Discriminatively Trained Part Based Models / P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan / IEEE Transactions on Pattern Analysis and Machine Intelligence, /Vol. 32, No. 9, Sep. 2010
6. Pyramid methods in image processing/E.H. Andelson and C.H. Anderson and J.R. Bergen and P.J. Burt and J.M. Ogden/ 1984.
7. Support-Vector Networks /Corinna Cortes, Vladimir Vapnik, - Сполучені Штати Америки, AT&T Bell Labs., Holmdel, NJ 07733, USA, 1995
8. Rapid Object Detection using a Boosted Cascade of Simple Features/Paul Viola, Michael Jones/2001
9. Object Detection with Deep Learning: A Review/Zhong-Qiu Zhao, Member, IEEE, Peng Zheng, Shou-tao Xu, and Xindong Wu, Fellow, IEEE/2019
10. Very deep convolutional networks for large-scale image recognition /Karen Simonyan, Andrew Zisserman, - Сполучені Штати Америки, University of Oxford, 2015

11. Face Detection with the Faster R-CNN /Huaizu Jiang, Erik Learned-Miller, -
Сполучені Штати Америки, University of Massachusetts Amherst, 2016
12. Rich feature hierarchies for accurate object detection and semantic
segmentation/ Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik -
Сполучені Штати Америки, UC Berkeley, 2014
13. Fast R-CNN/ Ross Girshick - Сполучені Штати Америки, Microsoft
Research, 2015
14. Face Recognition: Long-Range and Surveillance/Walter Scheirer, Archana
Sarkota - Сполучені Штати Америки, University of Colorado at Colorado
Springs and Securics, Inc. ,2011
15. Eigenfaces for Recognition/ Matthew Turk, Alex Pentland, Сполучені Штати
Америци, Vision and Modeling Group, The Media Laboratory, Massachusetts
Institute of Technology, 1991
16. Siamese Neural Networks for One-shot Image Recognition/Gregory Koch,
Richard, Ruslan Salakhutdinov - Канада, University of Toronto, 2015