

УДК 004.94

ДОСЛІДЖЕННЯ ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ В СИСТЕМАХ РОЗПІЗНАВАННЯ ОБЛИЧЧЯ

Горбан Олександр Юрійович

канд. техн. наук, доцент Бабілонга Оксана Юріївна

Національний університет «Одеська політехніка», УКРАЇНА

АНОТАЦІЯ. Проаналізовано особливості реалізації основних складових систем розпізнавання обличчя: знаходження обличчя на фотографії (*face detection*), пошуку елементів зображення обличчя (*landmarks detection*), кодування обличчя (*face encoding*), порівняння зображення обличчя з шаблонами (*face matching*). Розглянуто використання методів машинного навчання на основі згорткових нейронних мереж для розпізнавання обличчя. Визначена актуальність розробки адаптивних алгоритмів розпізнавання з використанням нейронних мереж, які зможуть автоматично налаштуватися на варіативні умови формування зображень.

Вступ. Розпізнавання обличчя – це область, яка знаходить все більше застосувань в сучасному світі. Нейронні мережі стали ключовим інструментом в цій області, оскільки вони мають виняткову здатність виявляти та розпізнавати шаблони в даних. Ця доповідь розглядає, як нейронні мережі використовуються для розпізнавання обличчя, які виклики виникають в процесі їх використання і які можливі шляхи їх подолання.

Розпізнавання обличчя є актуальною темою, що знаходить застосування в безлічі сфер - від безпеки до соціальних медіа. Нейронні мережі стали ключовим інструментом в цій області завдяки своїм винятковим можливостям у розпізнаванні шаблонів.

Мета роботи. Метою роботи є дослідження підходів, алгоритмів та технологій використання нейронних мереж в сучасних системах розпізнавання обличчя для розуміння, які типи нейронних мереж можуть бути використані на різних етапах в системах розпізнавання обличчя, а також для виявлення напрямків подальшого дослідження.

Основна частина роботи. Нейронні мережі – це потужний інструмент машинного навчання, який набув широкого розповсюдження в останні десятиріччя. Одним з актуальних застосувань нейронних мереж є завдання розпізнавання обличчя, що використовується в різноманітних сферах.

Під розпізнаванням обличчя може розумітися одне з двох завдань [1]:

1. **Верифікація** – порівняння двох зображень для визначення, що на них одна і та ж людина чи ні. Верифікація частіше реалізується в біометричних системах контролю доступу та відбувається за таким алгоритмом:
 - система формує зображення обличчя користувача та автоматично витягує із зображення певний вектор ознак;
 - з бази даних системи береться фотографія людини з минулого;
 - два зображення автоматично звіряються;
 - якщо міра близькості між двома векторами менша за певний показник, то робиться висновок, що на зображеннях одна і та ж людина, і підтверджується операція верифікації;
2. **Ідентифікація** – пошук людини за зображенням обличчя в існуючій базі чи картотеці зображень.

Задача розпізнавання обличчя може бути поділена на кілька під задач [2], включаючи: виявлення обличчя (локалізація обличчя на зображенні), витягнення особливостей (визначення ключових характеристик обличчя, таких як відстані між очима, форма носа, розмір рота тощо), а також власне розпізнавання та ідентифікацію обличчя (визначення, кому належить обличчя, або перевірка того, чи належить воно певній особі). Таким чином, в системі розпізнавання обличчя можна виділити кілька складових:

– знаходження обличчя на фотографії (*face detection*). Навчається незалежно від інших частин послідовності. Для навчання та тестування детектора осіб необхідний певний набір даних, у якому є розмічені прямокутниками обличчя (*bounding boxes*) і, бажано, розмічені ключові точки обличчя: очі, ніс, куточки рота. Якщо немає власного набору даних, можна використовувати

вільно доступні набори даних, наприклад *WIDER face*, в якому понад 350 тисяч розмічених зображень осіб. Зазвичай ніхто не вигадує нову архітектуру детектора і просто використовують доступні, вже створені, наприклад, *MTCNN*, *Retina Face*, *SCRFD*, *Yolov5Face*. Залежно від варіантів застосування, детектори можуть додатково до навчатись на даних користувача, за рахунок чого забезпечується необхідна швидкість та точність детектування області обличчя на зображенні;

– пошук елементів зображення обличчя (*landmarks detection*). Ця частина послідовності найменш важлива, і зазвичай її немає потреби навчати. Спочатку детектор обличчя визначає прямокутник і ключові точки обличчя, зазвичай п'ять точок. Потім вирівнювач за допомогою афінного перетворення повертає і зрушує ключові точки, а разом з ними і зображення обличчя до еталонного положення;

– кодування обличчя (*face encoding*) – розпізнавання та витягування дескрипторів зображення обличчя. Це найважливіша частина програми: навчена модель найбільше важить і найдовше працює за часом. На цьому етапі потрібні три складові. По-перше, хороша архітектура нейронної мережі. Архітектуру можна вибрати в *Imagenet* – одному з найважливіших ресурсів у комп'ютерному зорі, або зробити поєднання з кількох архітектур у зв'язці з дистиляцією знань (*knowledge distillation*). По друге, набір даних, який зазвичай складається з сотень тисяч фото реальних людей у різному ракурсі, різного віку, статей та рас, який може бути вільно доступним. Серед популярних датасетів наступні: *VGG2*, *faces*, *MegaFace*, *UMD*, *MSCeleb1M* [3]. По-третє, відмінна функція втрат (*loss function*). Наприклад, функція втрат *ArcFace* або її поліпшення. Навчання моделі відбувається як завдання класифікації;

– порівняння зображення обличчя з шаблонами (*face matching*). При розпізнаванні осіб на зображенні зазвичай вектори ознак витягуються з кожного зображення окремо, а потім осереднюються в один агрегований набір векторів. Для того щоб кадри з погано видимим обличчям давали менший внесок, часто навчають невелику нейронну мережу, яка привласнює ваги кожному кадру: чим краще видно обличчя, тим вища вага. Ця модель потрібна для відстеження траєкторій руху людей та використовується не в кожній системі. Популярні трекери базуються, наприклад, на фільтрі Калмана.

Незважаючи на свою потужність, нейронні мережі також стикаються з декількома проблемами і викликами. Одним з них є потреба в великій кількості даних для навчання. Також, нейронні мережі можуть мати тенденцію до перенавчання, коли вони «надто добре» вчаться на тренувальних даних і погано справляються з новими даними. Ця проблема може бути розв'язана за допомогою різних методів, таких як урегулювання або аугментація наборів зображень, що використовуються для перевірки.

Крім того, є етичні та юридичні проблеми, пов'язані з використанням технологій розпізнавання обличчя. Це включає питання приватності, зберігання та використання біометричних даних, а також можливість зловживання цією технологією.

В сучасних системах, завдання розпізнавання обличчя частіше вирішується на основі методів машинного навчання [4], оскільки вони допомагають системі «навчитися» розрізняти різні обличчя на основі великої кількості прикладів, представлених у вигляді даних для навчання. Так, з появою глибокого навчання та нейронних мереж, зокрема згорткових нейронних мереж (CNN), процес розпізнавання обличчя значно покращився. Вони були розроблені спеціально для обробки зображень та виявлення ключових особливостей на них, таких як кольори, текстури та форми.

CNN складаються з багатьох шарів, які відповідають за виявлення різних характеристик обличчя. Перші шари зазвичай фокусуються на виявленні низькорівневих особливостей, таких як краї та кольори, тоді як глибші шари можуть виявляти високорівневі особливості, такі як форма та структура елементів обличчя.

Для навчання нейронних мереж використовуються великі набори даних з зображеннями обличчя, які містять різні вікові групи, етнічні групи, вирази обличчя та освітлення. Ці дані розділяються на тренувальний, валідаційний та тестовий набори.

Процес навчання полягає в тому, що мережа опрацьовує тренувальні зображення, поступово оптимізуючи свої параметри таким чином, щоб мінімізувати помилку між передбаченими та дійсними мітками. Валідаційний набір даних використовується для перевірки якості моделі під час навчання, допомагаючи запобігти перенавчанню. Нарешті, тестовий набір даних використовується для оцінки загальної точності моделі після закінчення навчання.

CNN можуть використовуватися в поєднанні з іншими методами машинного навчання. Наприклад, одним з популярних підходів є використання CNN для витягнення особливостей, а потім використання іншого класифікатора, наприклад, машини опорних векторів (SVM), для розпізнавання обличчя на основі цих особливостей.

Висновки. Використання нейронних мереж у системах розпізнаванні обличчя допомагає створювати сучасні системи, які знаходять застосування в різних сферах, включаючи безпеку, біометричну ідентифікацію, контроль доступу та соціальні медіа. Завдяки своїй здатності вчитися на основі великої кількості даних, нейронні мережі стали ключовим інструментом у сучасних системах розпізнавання обличчя.

Незважаючи на значний прогрес в області розпізнавання обличчя, наприклад, застосування CNN, ця область все ще стикається з рядом викликів. Одним з них є вплив на точність розпізнавання варіативності умов зйомки, а саме – освітлення, пози голови, виразу обличчя, віку тощо. Тому подальші дослідження в цій області повинні бути спрямовані на розробку нових методів та алгоритмів, які можуть подолати ці виклики. Зокрема, корисним буде дослідити створення адаптивних алгоритмів розпізнавання, які зможуть автоматично налаштуватися на варіативні умови формування зображень.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Apoorva Chaudhary (2019) Assisting blind using facial recognition techniques (pp. 17-20).
2. Philip M. Parker Ph.D (2019) The 2020-2025 World Outlook for Facial Recognition (pp. 34-42).
3. Gerardus Blokdyk (2021) Facial Recognition Technology A Clear and Concise Reference (pp.191-198).
4. A Review on Automatic Facial Expression Recognition Systems Assisted by Multimodal Sensor Data (2019) / [N. Samadiani, G. Huang, B. Cai and etc.], Sensors, 19.8: 1863.

INVESTIGATION OF NEURAL NETWORKS FOR FACE RECOGNITION

Oleksandr Gorban

PhD, Associate professor, Oksana Babilunha

Odessa Polytechnic National University, UKRAINE

ANNOTATION. The peculiarities of the implementation of the main component systems of face recognition are analyzed: finding a face in a photo (face detection), searching for elements of a face image (landmarks detection), face encoding (face encoding), comparing a face image with templates (face matching). The use of machine learning methods based on convolutional neural networks for face recognition is considered. The relevance of the development of adaptive recognition algorithms using neural networks, which can automatically adjust to variable conditions of image formation, is determined.