

МОДИФІКАЦІЯ АРХІТЕКТУРИ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ OPENFACE ДЛЯ СИСТЕМИ РОЗПІЗНАВАННЯ ОБЛИЧ

Мякишин О. Д., магістр, oleksandr.miakshyn.kita@donntu.edu.ua

Ануфрієв П. О., аспірант, pavlo.anufriiev@donntu.edu.ua

*Донецький національний технічний університет,
м. Покровськ, Україна*

З розвитком технологій і істотним збільшенням потужності обчислювальних систем, стали актуальними завдання відтворення розумових процесів людини і, зокрема, задача розпізнавання обличч. Системи, що вирішують таке завдання, можуть використовуватися для автоматичної ідентифікації людей на зображеннях, для верифікації та авторизації користувачів додатків або співробітників підприємств, в сфері безпеки, тощо. Так як нейронні мережі показують найкращі результати якості і швидкості розпізнавання, для них пропонуються нові архітектури, алгоритми, модифікації, що підвищують ефективність їх функціонування, то для дослідження є доцільним використовувати саме їх. Найбільш придатними для цієї задачі є згорткові нейронні мережі, що з самого початку використовувалися для розпізнавання зображень [1, 2].

Задачами даного дослідження є проведення аналізу структури системи розпізнавання, аналізу моделей сучасних ефективних згорткових нейронних мереж з метою вилучення підходів та алгоритмів, які в них використовуються. Результат аналізу втілюється через модифіковану архітектуру для побудови власної системи розпізнавання. Розглянуто оцінки ефективності її роботи та якості розпізнавання.

Більш детально структуру технології розпізнавання проаналізовано в [3, 4]. Її етапи: детекція обличч на зображенні, виділення ознак кожного обличчя, безпосереднє розпізнавання, що складається з верифікації та ідентифікації. Детекція обличчя, як і детекція будь-якого іншого об'єкта на зображенні, полягає в знаходженні координат області, яка його обмежує (bounding box). Координати задаються в пікселях. За допомогою нейронної мережі, знаходження координат може вирішуватися як задача регресії [5]. Після детекції обличч на зображенні для кожного обличчя виділяються ознаки, що унікально характеризують його в стислому вигляді. Такі ознаки можуть представлятися у вигляді вектору чисел. При використанні глибокого навчання, ознаки виділяються самими нейронними мережами. Згорткові нейронні мережі виділяють ознаки у вигляді результуючих карт ознак зображення, а якщо в якості другої частини згорткової мережі додається перцептрон, то вектором ознак є вихідний вектор мережі.

Для аналізу обрано відому модель мережі для ідентифікації та верифікації осіб – FaceNet [6]. Ця мережа є сіамською мережею, яка кодує два вхідних зображення обличч в вектор ознак, а потім порівнює їх між собою по

відстані, щоб зрозуміти наскільки схожі ці два обличчя. При цьому головне завдання – навчити мережу таким чином, щоб для однієї і тієї ж людини нейронна мережа видавала найбільш схожий вектор ознак, а для різних людей – різні вектори. Для цього в даній мережі була розроблена спеціальна функція втрат, названа Triplet Loss (1):

$$\mathcal{L}(A, P, N) = \max(\|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2 + \alpha, 0) \quad (1)$$

де: A — вхід якорю (anchor);

P — позитивний вхід того ж класу, що і A;

N — негативний вхід того ж класу, що і A;

α — розрив між позитивними і негативними парами;

f — вбудовування.

Функція оперує трьома зображеннями обличчя. На двох із зображень знаходиться одна і та ж людина, а на третьому – інша. Одне з двох зображень тієї ж людини називається якорем (A – anchor), друге береться за позитивний (P – positive) приклад, а зображення з іншою людиною – за негативний (N – negative) (рис. 1). Суть використання Triplet Loss полягає в тому, щоб навчити нейронну мережу для тієї ж людини видавати близькі по відстані вектори ознак, а для різних людей – вектори, що відрізняються на задану константу, яка називається відступом. Відступ потрібен, щоб нейронна мережа не намагалася знаходити найбільш тривіальне рішення, роблячи всі вектори однаковими.

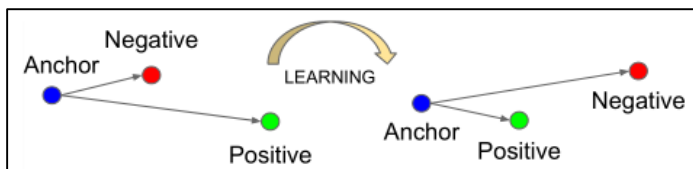


Рисунок 1. Навчання, застосовуючи Triplet Loss [6]

Подальшим розвитком мережі FaceNet стала мережа OpenFace [7]. Її розробка була акцентована на тому, щоб дослідити і довести можливість навчання мережі для розпізнавання облич на відкритому наборі

даних, на відміну від її попередників. Також ця мережа передбачалася для розпізнавання облич на мобільних пристроях, потужність яких не така велика, і при цьому бути все ще досить швидкою, для використання в режимі реального часу. Тому структура OpenFace є модифікованою структурою мережі FaceNet зі зменшеною кількістю параметрів для більш швидкого навчання і роботи мережі. Особливостями архітектур FaceNet та OpenFace є шари «поглинання» (inception). Шари inception утворюються з розгалуження шару, коли до нього додаються шари не всі послідовно, а паралельними групами, які надалі можуть складатися з послідовних шарів. Після цього inception шар поглинає всі розгалуження в один цілісний шар. Архітектуру OpenFace було обрано для модифікації і її подальшого навчання, як класифікатора (в кінець мережі додається шар, що класифікує).

При цьому важливо відзначити, що шар згортки у всіх випадках складається з трьох шарів – безпосередньо згортка, батч-нормалізація, а потім активація за допомогою функції активації ReLU.

Також, крім методу навчання через класифікатор, було вирішено модифікувати структуру мережі через її спрощення. Для цього в модифікованій структурі видаляються 4 і 5 групи inception шарів. Проте, мережа навчається на досить великому датасеті. За аналітичними оцінками таким чином можна домогтися більш швидкого навчання мережі, не погіршуючи якість розпізнавання. Архітектура оригінальної мережі та запропонована модифікація наведена на рис. 2.

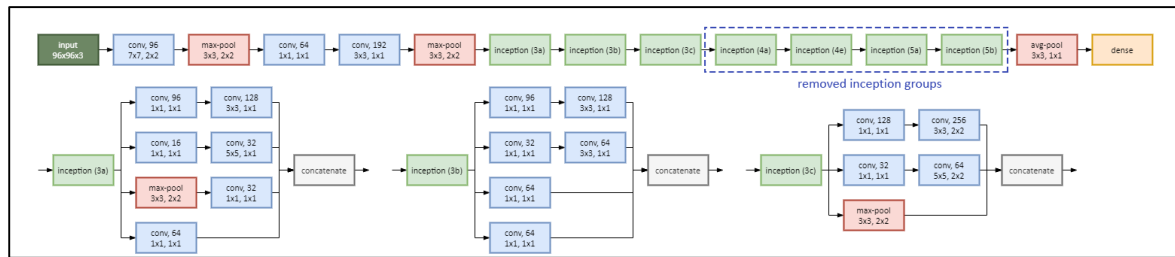


Рисунок 2. Архітектура мережі OpenFace та запропонована модифікація [7]

Для експериментів з модифікованою нейронною мережею обрано відповідні датасети: два датасети для навчання – Labeled Faces in the Wild (LFW) та Pin Faces, для тестування – ORL Faces.

Результати експериментів наведені на рис. 3 та 4. Рис. 3 демонструє графіки залежності акуратності розпізнавання від порогового значення для кожної з мереж, отримані при порівнянні векторів ознак за допомогою Евклідової відстані та косинуса кута. З отриманих графіків видно, що акуратність розпізнавання при найкращому порозі для модифікованої мережі вища, ніж для оригінальної. Різниця між двома модифікованими мережами невелика. Графік мережі, яка навчалася на датасеті LFW вужчий, коли графік для мережі, навченої на PinsFace, має ширшу форму. Це означає, що якість розпізнавання менше, якщо для навчання мережі застосовувати PinsFace, де кількість класів менше. Крім того, як можна побачити, результати майже ідентичні при застосуванні Евклідової відстані та косинуса кута. Звідси можна дійти висновку, що залежність якості розпізнавання від порога корелює з кількістю класів і для завдання розпізнавання осіб вибір методу порівняння векторів ознак немає значення.

Далі для кожної мережі було розраховано набір метрик у точці, яка відповідає максимальній акуратності. Отримані значення акуратності наведено на рис. 4. Зліва результати оригінальної мережі OpenFace, посередині – модифікованої мережі, навченої на датасеті LFW, а справа – модифікованої мережі, навченої на датасеті PinFaces.

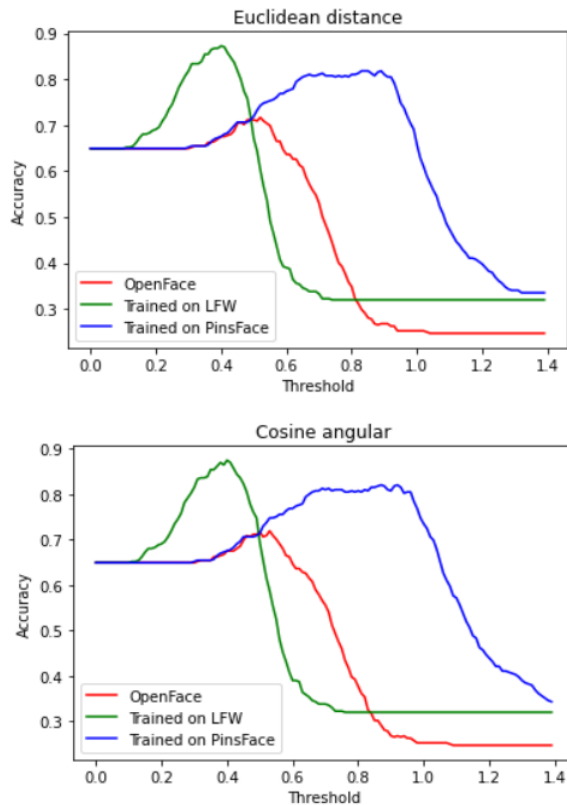


Рисунок 3. Залежність акуратності розпізнавання від порога при порівнянні векторів ознак по Евклідовій відстані та косинусу кута

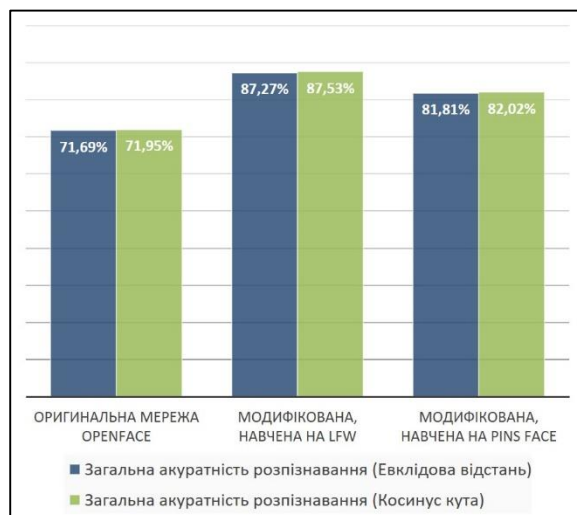


Рисунок 4. Порівняння найкращих результатів моделей

Як можна побачити при навчанні власних моделей вдалося підвищити загальну акуратність розпізнавання приблизно на 11-17%. При чому мережа, яка навчалася на датасеті LFW дала на 6% кращий результат, ніж при навчанні на PinFaces. Можливо, це пов'язано з тим, що LFW має більшу кількість класів (5749 класів), в порівнянні з PinFaces (105 класів). Це узгоджується з дослідженням [3]. Мережа навчилася відрізняти більш складні приклади один від одного, тобто зображення облич людей, схожих між собою.

Як результат, модифіковано архітектуру OpenFace через метод навчання класифікатором і спрощення архітектури мережі. Використані набори для навчання мережі – LFW та Pin Faces, для тестування застосовано ORL Faces. Поліпшення після модифікації на тестовому наборі досягло майже 17%. В результаті отримано працездатну модель згорткової нейронної мережі для розпізнавання облич, яка дає високі показники точності і може бути використовуватися для впровадження у майбутню систему розпізнавання облич, яка розроблюється.

Література

1. S. Ramakrishnan, "Introductory Chapter: Face Recognition - Overview, Dimensionality Reduction, and Evaluation Methods," Face Recognit. - Semisupervised Classif. Subsp. Proj. Eval. Methods, pp. 1–6, 2016, doi: 10.5772/63995.
2. Путь в Machine Learning и Deep Neural Networks [Електронний ресурс] – Режим доступу до ресурсу: <https://habr.com/ru/company/oleg-bunin/blog/470904/>
3. Ануфрієв П. О. Дослідження впливу тестових баз зображень на точність систем розпізнавання обличч з різними методами класифікації/ П. О. Ануфрієв, Э. О. Башков // Наукові праці ДонНТУ. Серія «Інформатика, кібернетика та обчислювальна техніка». – 2020. – №2 (31). – С. 74-83
4. Mala Sundaram and Ambika Mani, "Face Recognition: Demystification of Multifarious Aspect in Evaluation Metrics," Face Recognit. - Semisupervised Classif. Subsp. Proj. Eval. Methods, p. 18, 2016, doi: 10.5772/62825.
5. Z. H. Feng, J. Kittler, M. Awais, P. Huber, and X. J. Wu, "Face Detection, Bounding Box Aggregation and Pose Estimation for Robust Facial Landmark Localisation in the Wild," IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., vol. 2017-July, pp. 2106–2111, 2017, doi: 10.1109/CVPRW.2017.262.
6. F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 815–823, 2015, doi: 10.1109/CVPR.2015.7298682.
7. M. S. Brandon Amos, Bartosz Ludwiczuk, "Openface: A general-purpose face recognition library with mobile applications," C. Sch. ..., 2016 r., vol. 16, no. October 2015, pp. 1–18, 2016, [Електронний ресурс] – Режим доступу до ресурсу: <http://reports-archive.adm.cs.cmu.edu/anon/anon/usr0/ftp/2016/CMU-CS-16-118.pdf>.

Анотація

Представлено модифіковану архітектуру OpenFace через метод навчання класифікатором і спрощення архітектури мережі. Поліпшення після модифікації на тестовому наборі досягло майже 17%. В результаті отримано працездатну модель згорткової нейронної мережі, яка дає високі показники точності і може бути використана для впровадження у майбутню систему розпізнавання обличч.

Ключові слова: розпізнавання обличч, згорткова нейронна мережа, модифікація.

Аннотация

Представлена модифицированная архитектура OpenFace через метод обучения классификаторам и упрощение архитектуры сети. Улучшение после модификации на тестовом наборе достигло почти 17%. В результате получена работоспособная модель сверточной нейронной сети, которая дает высокие показатели точности и может быть использована для внедрения в будущую систему распознавания лиц.

Ключевые слова: распознавание лиц, сверточная нейронная сеть, модификация.

Abstract

A modified OpenFace architecture through the classifier learning method and network architecture simplification is presented. The improvement after modification on the test dataset reached almost 17%. The result is a workable model of a convolutional neural network, which gives high accuracy and can be used for implementation in the future face recognition system.

Keywords: face recognition, convolutional neural network, modification.