

Системы обнаружения распознавания лиц, основанные на различных архитектурах нейронных сетей: анализ и сравнение технологий

Гринин Игорь Леонидович

аспирант кафедры программного обеспечения автоматизированных систем, Волгоградский государственный технический университет, frederickbrown@yandex.ru

Кравченко Сергей Вячеславович

аспирант кафедры программного обеспечения автоматизированных систем, Волгоградский государственный технический университет frederickbrown@yandex.ru

В настоящей статье авторами исследуются структуры и принципы работы методов обнаружения, распознавания и выравнивания лиц. Методологиями исследования стали сравнительный анализ популярных нейросетевых технологий обработки изображений, а также пошаговый разбор анализ на примере таких моделей. Были разобраны принципы работы технологий распознавания, обнаружения и выравнивания лиц, основанных на нейронных сетях и технологий, работающих без участия нейронных сетей, в том числе работающих на адаптивных графах и собственных векторах. Итогом исследования стало описательный анализ работы систем и подсистем распознавания, обнаружения и выравнивания лиц. В процессе исследования были получены необходимые теоретические знания, для работы с изображениями лиц, которые могут стать полезны для различных возможных обработок и других видов работ с таким типом данных.

Ключевые слова: обнаружение лиц, распознавание лиц, графы, сверточные нейронные сети, компонентная нейронная сеть

Введение

В современном мире постоянно развивающихся информационных технологий. Постоянная работа с изображениями, в том числе с фотографиями и видеосъемками людей. Настоящая статья рассказывает о технологиях работы с изображениями, специализирующихся на обнаружении и опознавании лиц.

Авторами рассматривается работа моделей обработки изображений. Многие современные приложений для работы с камерами, камеры наблюдения на улицах - все они имеют встроенные системы распознавания лиц для лучшей и точной работы.

Прежде всего, как и любая сложная система, модели состоят из подсистем. В зависимости от типа модели и ее предназначения - обнаружение или опознавание лица, количество подсистем может быть различным, и, естественно различаться качественно.

На данный момент, в научной литературе имеется большое количество информации [1-4]. Анализ литературы показывает, что тема работы с распознаванием лиц исключительно востребована, всего опубликовано около 50 000 статей, при этом только за ару последних лет год опубликовано около 10000 статей. Это показывает возрастающий интерес к данной тематике. Однако, именно как составные части, рассматриваемой нами модели, чаще всего не рассматриваются. К тому же, к сожалению, данный вопрос исследуются не системно.

Другие методологии

Для быстрого введения в тему статьи, авторами было решено так же кратко ознакомить читателя с другими видами системами распознавания лиц, работающих не с помощью нейронных сетей.

Первой такой системой будет метод гибкого сравнения на графах.

Из самого названия метода уже понятно, что основной сутью его работы является сопоставление двух графов. Каждый из этих графов является описанием лица - то есть хранит информацию о нем в своих ребрах и вершинах. Первый из этих графов является образцом, так называемый «эталонный» граф. Он не меняется. Второй же, напротив - изменчивый, способный подстраиваться под образец. В этих вариациях графы собой могут являть как и обычные прямоугольные сетки, так и более сложные, построенные по форме лица.



Рис. 1 Прямоугольная сетка

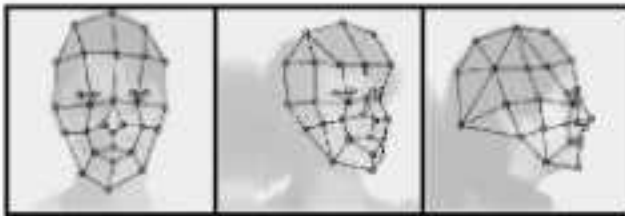


Рис. 2 Сетка по форме лица

Деформация графа происходит по следующему алгоритму.

Каждая из вершин графа смещается из своего первоначального местоположения. При этом она перемещается на такую позицию, чтобы в двух соответствующих вершинах каждого из графов разница признаков в вершинах старалась оставаться минимальной. Как только все вершины сместятся на максимально возможную дистанцию, при которой будет достигнуто минимальное различие между признаками графов, деформация считается законченной.



Рис. 3 Изменение сетки графа

Хоть данный метод и является крайне точным, (95-97% точности распознавания), он применяется редко, из-за его высокой вычислительной сложности и долгой времени работы при большом количестве лиц.

Вторым методом для ознакомления будет метод главных компонент.

Для правильной работы с данным методом создается матрица, куда добавляются все лица баз данных, по одному лицу на каждую строку. После этого все данные в матрице нормируются и высчитывается матрица ковариации по приведенным к нулевому среднему и первой дисперсии строкам. В получившейся матрице решается задача собственных значений, после чего формируются группы векторов с присвоенными им собственными лицами.



Рис. 4 Собственные вектора и лица

Данный метод много применяется на практике, в приложениях. Однако при плохом освещении или недостаточной точности изображения результативность метода сильно падает.

Методологии, основанные на нейронных сетях

Сейчас существует достаточно большое количество разновидностей нейронных сетей, пригодных для реализации систем обнаружения и распознавания лиц.

В этой статье мы проанализируем две популярных в каждом из направлений нейронных сети для распознавания лиц: MTCNN и FaceNet.

Нейронная сеть обнаружения лиц MTCNN

Эта сеть является составной, поскольку является каскадом из трех сверточных нейронных сетей. Каждая из них имеет собственное название - P-net, O-net и R-net, и, соответственно, отвечает за собственное действие.

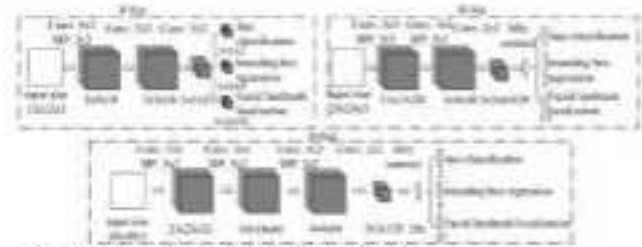


Рис. 5 Каскад трех сверточных сетей

Перед тем, как мы будем описывать работу каждой из под сетей, надо обязательно сказать, что данная сеть распознает лица при помощи так называемых пяти «якорей» - пяти точек-ориентиров на лице.



Рис. 6 Пять точек-«якорей»

Эти точки обозначают глаза, нос и рот. По этим точкам строится граф, рассчитывающий расстояния, и определяющийся, как отдельно взятое лицо. В общем, это имеет некое сходство с методом гибкого сравнения графов, однако далее в ход вступают силы нейронных сетей.

Первая из трех сетей - P-net. Она генерирует несколько так называемых кадров-кандидатов. Другими словами, с помощью этой сети находятся области с предполагаемыми лицами и отрезаются от остального изображения. Именно в этой сети задаются пять точек лица и записываются все вектора.

Следующим шагом является R-net. Смысл ее работы максимально противоположен предыдущей сети - теперь необходимо выбрать все, что точно не может быть лицами на изображении. Эти области также отсекаются от основного изображения.

Третья же сеть, O-net, выбирает из выданных ей первой итерацией кандидатов наиболее достоверных. Поскольку глубина нейронной сети уже достаточно глубока, то лица распознаются с высокой точностью.

Почему же используются сразу три сети, если все равно главный результат выдает только одна? Ответ достаточно прост - все дело в оптимизации процесса. Ведь

если изначально будет использовано только O-net, то производительность будет очень низкой, ведь там будет произведен очень большой набор работ. Поэтому, первые две итерации необходимы, для того, чтобы сделать изначальную фильтрацию областей на изображении, и уже после это передавать эту информацию для выделения областей лиц. Это позволяет сократить время работы основного процесса.

Нейронная сеть распознавания FaceNet

Также, как и предыдущая, эта сеть является компонентной. Однако, FaceNet состоит всего из двух подсетей. Эти две сети абсолютно одинаковые, поэтому такая система называется симметричной. Такие сети обучаются дифференцированию, то есть в данном случае, сеть учится понимать какие изображения похожи, а какие нет.

FaceNet использует особую функцию потерь называемую TripletLoss. Она минимизирует дистанцию между якорем и изображениями, которые содержат похожую внешность, и максимизирует дистанцию между разными.

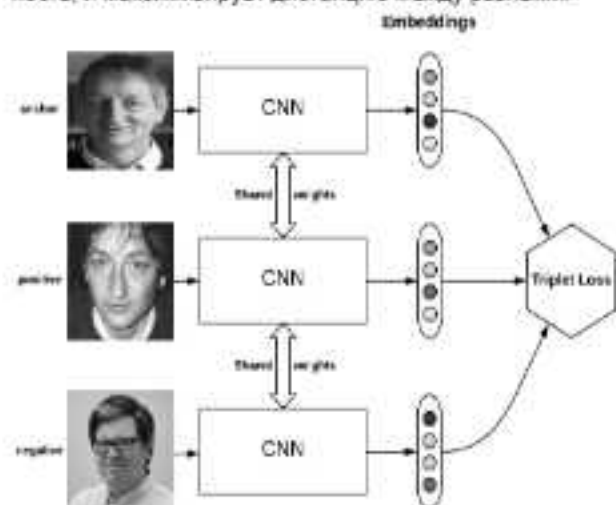


Рис. 7 Функция TripletLoss

Дистанцию между точками-якорями и изображениями TripletLoss определяет и оптимизирует напрямую. Однако, на тренировочных моделях, в процессе их обучения выборка троек изображений достаточно специфична. В зависимости от того, как она поставлена, результат работы модели будет разным. При постоянном случайном выборе, например, модель будет работать, однако максимальной производительности она не достигнет. Если изображения для сравнения изначально будут сложными для анализа, то модель может и не работать вовсе. Поэтому, для тренировки модели предлагается использовать, так называемые «полужесткие» данные, дабы модель смогла быть рабочей и производительной. При этом, база данных изображений для обучения должна быть очень объемной.

Натренированная сеть, завершив свое обучение, теперь может определять лица, проводя их сравнения со своей базой данных.

Выводы

Технологии обнаружения и распознавания лиц идут рука об руку, поскольку существование одной из них почти не имеет смысла без наличия другой. Даже две рассмотренных нами сегодня сети на самом деле являются

двумя составными частями одной сети Google FaceNet, и работают, обрабатывая общую информацию.



Рис. 8 Принцип работы общей сети

Каждая модель имеет собственные достоинства и недостатки, но вместе эти 2 сети образуют самую популярную систему обнаружения и распознавания лиц, с наиболее высокой точностью работы. Это показывает важность правильного и оптимизированного решения по организации работы нейронных сетей.

Литература

1. Аноприенко А.Я., Иванова С.В., Лычагин Р.В. Анализ современных подходов к процедурам узнавания и распознавания лиц для систем биометрического контроля // Информатика и кибернетика. 2016. № 1 (3). С. 25-31.
2. Shamova U. Face recognition in healthcare: general overview // Languages in professional communication. Сборник материалов международной научно-практической конференции преподавателей, аспирантов и студентов, ответственный редактор Л. И. Корнеева. 2020. С. 748-752.
3. Majekodunmi T.O., Idachaba F.E. A review of the fingerprint, speaker recognition, face recognition and iris recognition based biometric identification technologies // Proceedings of the World Congress on Engineering 2011, WCE 2011. 2011. С. 1681-1687.
4. Raveendra K., Ravi J. Performance evaluation of face recognition system by concatenation of spatial and transformation domain features // International Journal of Computer Network and Information Security. 2021. Т. 13. № 1. С. 47-60.
5. Minu M.S., Arun K., Tiwari A., Rampuria P. Face recognition system based on haar cascade classifier // International Journal of Advanced Science and Technology. 2020. Т. 29. № 5. С. 3799-3805.
6. Нейросети в большом городе. Разбираемся, как они помогают идентифицировать людей, и запускаем собственную нейросеть <https://habr.com/ru/company/skillfactory/blog/536410/>
7. Анализ существующих подходов к распознаванию лиц <https://habr.com/ru/company/synesis/blog/238129/>
8. Face Detection using MTCNN — a guide for face extraction with a focus on speed <https://towardsdatascience.com/face-detection-using-mtcnn-a-guide-for-face-extraction-with-a-focus-on-speed-c8d59f82d49>
9. Introduction to FaceNet: A Unified Embedding for Face Recognition and Clustering <https://medium.com/analytics-vidhya/introduction-to-facenet-a-unified-embedding-for-face-recognition-and-clustering-dbdac8e6f02>

Face recognition detection systems based on various neural network architectures: analysis and comparison of technologies

Grinin I.L., Kravchenko S.V.

Volgograd State Technical University

JEL classification: C10, C60, C61, C67, C68, C69

The aim of the research is to study the structures and principles of the methods of detection, recognition and alignment of faces. The research methods were a comparative analysis of popular neural network image processing technologies, as well as a step-by-step analysis of the analysis using the example of such models. The principles of face recognition, detection and alignment technologies based on neural networks and technologies that work without the participation of neural networks, including those that work on adaptive graphs and eigenvectors, were analyzed. The result of the study was a descriptive analysis of the work of systems and subsystems for face recognition, detection and alignment. In the course of the research, a number of theoretical knowledge was obtained for working with face images, which can be useful for various possible treatments and other types of work with this type of data.

Keywords: face detection, face recognition, graphs, convolutional neural networks, component neural network

References

1. Analysis of modern approaches to face recognition and recognition procedures for biometric control systems / Anoprienko A. Ya., Ivanitsa S. V., Lychagin R. V. *Informatics and Cybernetics*. 2016. NO. 1 (3). PP. 25-31.
2. Facial recognition in healthcare: a general overview / Shamova U. In the collection: *Languages in professional communication. Collection of materials of the international scientific and practical conference of teachers, postgraduates and students: the executive editor is L. I. Korneeva*. 2020. pp. 748-752.
3. Overview of fingerprint-based biometric identification, speaker recognition, face recognition, and its technologies / Majekodunmi T. O., Idachaba F. E. In the collection: *Proceedings of the World Congress on Technology 2011, WCT 2011*. 2011. pp. 1681-1687.
4. Evaluation of the effectiveness of the face recognition system by concatenating the spatial and transformational characteristics of the subject area / Raveendra K., Ravi J. *International Journal of Computer Network and Information Security*. 2021. Vol. 13. No. 1. pp. 47-60.
5. Face recognition system based on the haar cascade classifier / Minu M. S., Arun K., Tiwari A., Rampuria P. // *International Journal of Advanced Science and Technology*. 2020. Vol. 29. No. 5. pp. 3799-3805.
6. Neural networks in the big city. We understand how they help identify people, and we launch our own neural network <https://heor.com/ru/company/skillfactory/blog/536410/>
7. Analysis of existing approaches to face recognition <https://heor.com/ru/company/synthesis/blog/238125/>
8. Face Recognition with MTCNN - A guide to face extraction with a focus on speed <https://towardsdatascience.com/face-detection-using-mtcnn-a-guide-for-face-extraction-with-a-focus-on-speed-c6d59f62d49>
9. Introduction to FaceNet: Unified Embedding for Face Recognition and Clustering <https://medium.com/analytics-vidhya/introduction-to-facenet-a-unified-embedding-for-face-recognition-and-clustering-d0dc8e6f02>