

Эффективное распознавание лиц на основе последовательного анализа нейросетевых дескрипторов и детектирования миноритарных классов

А.Д. Соколова
Национальный исследовательский университет Высшая
школа экономики
Нижний Новгород, Россия
adsokolova@hse.ru

А.В. Савченко
Национальный исследовательский университет Высшая
школа экономики
Нижний Новгород, Россия
avsavchenko@hse.ru

Аннотация—Исследуются способы повышения точности распознавания лиц на основе обнаружения входных изображений, которые редко встречаются в наборах данных, используемых для обучения нейросетевых дескрипторов. В современных свободно распространяемых обучающих выборках обычно представлены изображения людей в основном среднего возраста и европеоидной расы, из-за этого большинство алгоритмов ошибаются на изображениях пожилых людей или детей, лицах более редких национальностей и т.п. В работе предложен алгоритм детектирования таких данных с последующей их отбраковкой, на первом этапе которого используется сверточная нейронная сеть, предобученная на специально созданном наборе редких данных. Второй этап – применение последовательного анализа дескрипторов для повышения вычислительной эффективности классификации. Экспериментальное исследование на наборе данных VGGFace2 с использованием нейросетевых дескрипторов, в том числе современных моделей InsightFace, продемонстрировало повышенную эффективность предложенного алгоритма по сравнению с известными аналогами.

Ключевые слова—распознавание лиц, аномалии, детектирование редких данных, последовательная классификация

1. ВВЕДЕНИЕ

Алгоритмы по распознаванию лиц достигли значительных улучшений за последнее время. Однако на ряде изображений до сих пор возникают ошибочные срабатывания классификатора. Это происходит за счет того, что обучающие выборки обычно содержат изображения знаменитостей достаточно хорошего качества [1]. Проблемы же возникают, когда появляется нестандартный масштаб, низкое разрешение, иное освещение и т.д. Более того человеку свойственно меняться в течение жизни, помимо возрастных изменений человек может изменить внешность: макияж, прическа, оттенок кожи и другое. Соответственно, в обучающей выборке сложно учесть такое разнообразие изображений для одного человека. В дальнейшем, подобные изображения будем называть аномальными.

В задаче распознавания лиц на открытом множестве (от англ. open-set) необходимо соотнести входное изображение с одним из существующих классов лиц или определить изображение в специальный дополнительный класс людей, отсутствующих в обучающей выборке. В данной работе предлагается

реализовать автоматический детектор аномальных изображений [2]. В таком случае, например, при классификации лиц по видео аномальным кадрам может присваиваться более низкий вес для принятия решения.

Чтобы обучить детектор аномальных изображений был создан способ автоматического сбора редких данных из известных наборов данных, а также применения особых техник по модификации изображений. Для ускорения алгоритма распознавания была предложена процедура последовательного анализа для классификации k-NN, которая повышает производительность с минимальными потерями точности. Алгоритм основан на использовании метода главных компонент [3] для определения наиболее значимых признаков с последующим разбиением дескриптора изображения на уровни.

Полученные результаты могут заинтересовать широкий круг специалистов, которые занимаются методами распознавания людей. Также предложенный подход может быть применен в таких сферах как охрана, сбор статистики по конкретным людям, автоматизация процессов, где требуются идентификация человека и т.п.

2. ПРЕДЛАГАЕМЫЙ ПОДХОД

На вход системе поступает изображение, где необходимо выполнить детектирование лица. В данной работе детектирование лиц происходит с помощью модели MobileNet SSD [4]. Далее из полученной области извлекаем вектор признаков, используя предварительно обученные сверточные нейронные сети. На следующем этапе проверяется, относится ли изображение к аномальным данным или нет. Для определения качества изображения используется простая полносвязная нейронная сеть «Рис. 1», обученная на специально собранном наборе аномальных данных, в качестве бинарного классификатора. Если изображение относится к аномальным данным, то происходит исключение изображения из анализа. Если нет, то применяется алгоритм классификации. В данной работе была приведена процедура последовательного анализа признаков для k-NN классификации.

Последовательный анализ признаков состоит в использовании метода главных компонент в целях сортировки элементов. Далее вектор признаков делится на уровни, число элементов на каждом уровне может быть как постоянным, так и выбранным с помощью

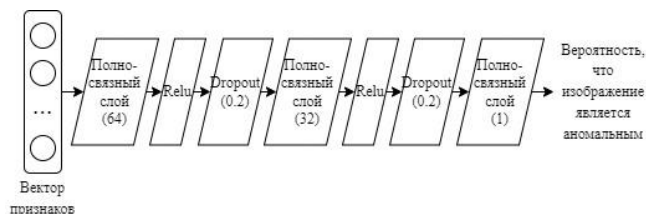


Рис. 1. Архитектура нейронной сети для детектирования аномальных изображений

объясненной дисперсии. После этого сравнение происходит последовательно между уровнями двух векторов. Если первый уровень не удовлетворяет условию принятия решения, то переходим к анализу второго уровня и т.д. В итоге решение, к какому объекту относится элемент, может быть принято при анализе начальных уровней или анализе всего вектора целиком.

Набор аномальных данных был создан следующим образом:

1. Использование наборов данных с изображениями людей разного возраста: Large age gap face verification [5] и All ages faces dataset [6];
2. Преобразование изображений (англ. data augmentation) с помощью библиотеки OpenCV и нейросетевых технологий [7, 8]: наложение двух видов шума с помощью генерации шума по нормальному распределению, размытие и переворот изображения, конвертация в черно-белое / нарисованное карандашом / нарисованное красками изображение, перенос макияжа, изменение возраста человека на изображении;
3. Применение алгоритма, где сначала происходит k-means кластеризация набора данных VGGFace2, число кластеров выбиралось равным числу различных людей в созданной обучающей выборке. Далее из каждого кластера бралось изображение и помещалось в обучающую выборку, а все остальные изображения в тестовую. Затем применялась классификация, в рамках которой собирался набор аномальных данных по ошибочным срабатываниям (точность распознавания для одного человека меньше 95%).

В рамках экспериментального исследования сопоставлялись различные сверточные нейронные сети для извлечения дескрипторов: VGGFace2 (SENet-50) [9], MobileNet [10], InsightFace (IRNet-50) [11], FaceNet (InceptionResNet) [12], а также различные классические алгоритмы классификации: k-NN, SVM, Random Forest, LDA. При этом сравнивалась классификация с исключением аномальных данных и без.

Наилучшие результаты точности были достигнуты с помощью InsightFace сети для извлечения признаков и k-NN классификатора с исключением аномальных данных: 97,9% на наборе данных VGGFace2 (в сравнении с 93,6% точностью без исключения аномальных данных). Последовательный анализ признаков позволил улучшить время в 6 раз с минимальными потерями в точности

(0,3%). При этом ошибки срабатывания детектора аномальных данных оказались следующими: TPR (True Positive Rate) 98,56%, TNR (True Negative Rate) 98,29%, FPR (False Positive Rate) 1,71%, FNR (False Negative Rate) 1,44%.

3. ЗАКЛЮЧЕНИЕ

В настоящей работе экспериментально показано, что использование предложенного подхода позволяет значительно повысить точность распознавания изображений лиц. Более того в работе используются легкие сверточные нейронные сети, что позволяет использовать алгоритм на различных устройствах в режиме реального времени. В будущем планируется разработать подход применения информации об аномальных данных без отбраковки таких изображений: присваивание весов, изменение дескрипторов аномальных данных и т.п.

БЛАГОДАРНОСТИ

Работа выполнена за счет гранта Российского научного фонда (проект № 20-71-10010).

ЛИТЕРАТУРА

- [1] Guo, Y. Ms-celeb-1M: A dataset and benchmark for largescale face recognition / Y. Guo, L. Zhang, Y. Hu, X. He, J. Gao // Proceedings of European Conference on Computer Vision (ECCV). – 2016. – P. 87-102.
- [2] Komkov, S. A. Distances between distributions for the belonging-to-the-distribution measurement of the image / pat. 2022045915A1 S. A. Komkov, A. A. Petiushko, I. L. Mazurenko, J. Li // Huawei Technologies CO., LTD. – PCT/RU2020/000450. – 25.08.2020.
- [3] Соколова, А. Д. Поиск редких данных в задаче распознавания лиц на изображениях / А. Д. Соколова, А. В. Савченко, С. И. Николенко // Компьютерная оптика. – 2022. – Т. 46. – С. 801-807.
- [4] Younis, A. Real-time object detection using pre-trained deep learning models MobileNet-SSD / A. Younis // Proceedings of 2020 the 6th international conference on computing and data engineering. – 2020. – С. 44-48.
- [5] Bianco, S. Large age-gap face verification by feature injection in deep apnetworks / S. Bianco // Pattern Recognition Letters. – 2017. – Vol. 90. – P. 36-42.
- [6] Cheng, J. Exploiting effective facial patches for robust gender recognition / J. Cheng, Y. Li, J. Wang, L. Yu, S. Wang // Tsinghua Science and Technology. – 2019. – Vol. 24(3). – P. 333-345.
- [7] Chen, Y. C. Facelet-bank for fast portrait manipulation / Y.C. Chen, H. Lin, M. Shu, R. Li, X. Tao, X. Shen, J. Jia // Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR). – 2018. – P. 3541-3549.
- [8] Zhang, H. Disentangled makeup transfer with generative adversarial network / H. Zhang, W. Chen, H. He, Y. Jin // arXiv preprint arXiv. – 2019. –P. 1907.01144.
- [9] Cao, Q. Vggface2 A dataset for recognizing faces across pose and age / Q. Cao, L. Shen, W. Xie, O. M. Parkhi, A. Zisserman // Proceedings of the 13th International Conference on Automatic Face and Gesture recognition (FG). – 2018. – P. 67-74.
- [10] Savchenko, A. V. Efficient facial representations for age, gender and identity recognition in organizing photo albums using multi-output ConvNet / A. V. Savchenko // PeerJ Computer Science. – 2019. – Vol. 5. – e197.
- [11] Deng, J. Sub-center arcfacenet: Boosting face recognition by large-scale noisy web faces / J. Deng, J. Guo, T. Liu, M. Gong, S. Zafeiriou // In European Conference on Computer Vision. – 2020. – P. 741-757.
- [12] Schroff, F. Facenet: A unified embedding for face recognition and clustering / F. Schroff, D. Kalenichenko, J. Philbin // Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – P. 815-823.