

УДК 004.8

DOI: 10.46548/21vek-2021-1056-0012

МЕТОДЫ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ В СИСТЕМАХ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ

© 2021

Маркин Евгений Игоревич, аспирант кафедры «Программирование»
Пензенский государственный технологический университет

(440039, Россия, Пенза, проезд Байдукова/ул. Гагарина, д. 1а/11, e-mails: evgeniymarkin1@gmail.com)

Аннотация. Обработка изображений, предваряющая их анализ с помощью алгоритма идентификации личности, является одним из ответственных этапов, определяющих точность распознавания человека. В настоящее время не существует единого метода решения этой задачи ввиду сильно различающихся внешних условий работы систем (разное освещение, качество камер, расстояние до объектов и т.д.). Поэтому одним из ключевых направлений разработки систем идентификации являются исследования в области предварительной обработки изображений. Классический подход подразумевает использование фильтров для нормализации контрастности и резкости изображений не всегда эффективен, поскольку при его применении могут теряться детали изображения лица. В статье приводится сравнительный анализ ряда современных методов повышения качества изображений, которые способны исключить потерю деталей.

Ключевые слова: нейронная сеть, сверточные нейронные сети, генеративно-состязательные сети, обработка фото-изображений, цифровая фотография.

IMAGE PREPROCESSING METHODS IN PERSONAL IDENTIFICATION SYSTEMS

© 2021

Markin Evgeniy Igorevich, postgraduate of sub-department «Programming»
Penza state technological university

(440039, Russia, Penza, Baydukov Proezd / Gagarin Street, 1a/11, e-mails: evgeniymarkin1@gmail.com)

Abstract. Image processing, which precedes their analysis by means of an identification algorithm, is one of the crucial steps determining the accuracy of human recognition. Currently, there is no single method for solving this problem due to widely differing external conditions of the systems (different lighting, camera quality, distance to objects, etc.). Therefore, one of the key areas in the development of identification systems is research in the field of image pre-processing. The classic approach involves the use of filters to normalize contrast and sharpness of images is not always effective, because it can lose details of the face image. This article provides a comparative analysis of a number of modern methods to improve image quality, which can eliminate the loss of detail.

Keywords: neural networks, convolutional neural networks, generative-adversarial networks, photo-image processing, digital photography.

Введение. Алгоритмы и методы распознавания лиц и идентификации личности за последние годы достигли высоких показателей благодаря развитию методов искусственного интеллекта, разработке процессоров для нейронных вычислений и росту заинтересованности компаний в данной области. Существующие технологии позволяют в режиме реального времени проводить одновременно довольно точное распознавание десятков лиц [1]. Но достижение такого уровня показателей возможно только при соблюдении ряда идеальных условий, таких как высокое разрешение фотографии, высокая четкость и контрастность снимка, равномерная освещенность, определенный угол наклона головы и т.д. [2]. Выполнение всех этих условий крайне затруднительно в реальном режиме работы оборудования, поэтому сейчас активно изучаются и испытываются различные методы для снижения воздействия внешних факторов на работу системы.

Целью настоящей работы является обзор и анализ методов предварительной обработки изображений для улучшения их качества и повышения точности идентификации личности.

Материалы и результаты исследования. Одной

из основных проблем распознавания лиц является качество изображений – недостаточное разрешение камеры, искажения оптических линз, перемещение объекта во время съемки ухудшают изображение, становится сложно извлечь полезную информацию. При этом системы распознавания и идентификации могут давать ложные сигналы или некорректные данные [3]. Традиционно для решения проблемы используется следующий комплекс методов [4]:

- для устранения размытия используется трекинг лица с поиском наилучшего кадра для распознавания;
- для извлечения максимальной информации из изображения проводится оценка его контрастности и ее усреднение с целью нормализации;
- для получения дополнительных деталей проводится оценка и увеличение резкости до заданных пороговых значений.

Данный подход характеризуется высоким быстродействием, но в процессе поиска наиболее «качественного кадра» могут теряться информативные кадры и увеличиваться время идентификации личности, что, в случаях, когда лицо попадает в кадр на несколько секунд, может стать критичным. Также из-за усреднения контрастности и повышения резкости одного

кадра могут проявляться артефакты, приводящие к ложным срабатываниям системы распознавания.

В качестве альтернативного метода для повышения качества изображений часто применяют метод «*Super Resolution*», данный подход основан на применении нейронных сетей.

На сегодняшний день в рамках указанного метода реализуются четыре основных подхода к улучшению изображения:

- *prediction models* (предсказательные модели) [5];
- *edge based methods* (краевые методы) [6];
- *image statistical methods* (статистические методы) [7];
- *patch based (or example-based) methods* (методы основанные на паттернах) [8].

Наилучшее качество дает модель *patch based (or example-based) methods* (методы основанные на паттернах).

Впервые использование сверточной нейронной сети, работающей по принципу *end to end*, для повышения резкости изображений было предложено Chao Dong в работе «*Image Super-Resolution Using Deep Convolutional Networks*» [9]. Предложенный подход состоит из трех основных функций (рис. 1):

- 1) извлечения и отображения патчей;
- 2) нелинейного отображения;
- 3) реконструкции.

Данный подход позволяет значительно улучшить качество изображения в сравнении с методами применения фильтров, при этом сохраняется высокая скорость работы (рис. 2).

В работе «*Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*» Christian Ledig [10] предлагает решение задачи улучшения фотографии путем применения генеративно-сопоставительной сети (GAN) [11-13] (рис. 3).

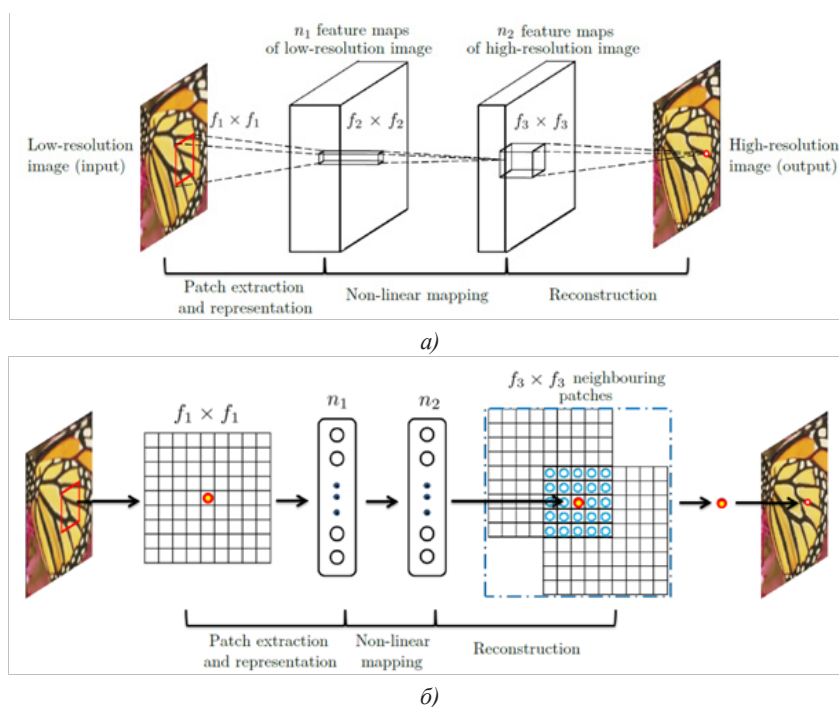
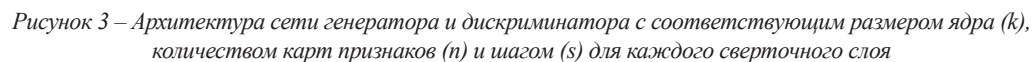


Рисунок 1 – Схема работы нейронной сети по методу разреженного кодирования (*Super-Resolution Convolutional Neural Network, SRCNN*)

а) двухслойная нейронная сеть, верхний уровень которой извлекает набор карт признаков, а второй слой нелинейно сопоставляет эти карты признаков с представлениями участков высокого разрешения; б) метод на основе разреженного кодирования в представлении сверточной нейронной сети



Рисунок 2 – Сравнение работы метода SRCNN с оригинальным изображением, бикубическим методом увеличения качества (*Bicubic*) и разреженным кодированием (*sparse-coding, SC*)



В большинстве подходов для подготовки изображения перед его распознаванием происходит поиск оптимального кадра, а соседние кадры полагаются

Однако Sam Hasinoff в своей работе «*Burst photography for high dynamic range and low-light imaging on mobile cameras*» [17] доказал, что использование соседних идентичных кадров позволяет значительно повысить качество изображений (рис 5.). В соответствии с данным подходом во избежание клиппинга (потери информации на светлых участках кадра) используются фотографии с достаточно низкой экспозицией [18], что дает дополнительные возможности для расширения динамического диапазона [19]. Помимо этого фотографии делаются с уменьшенной выдержкой, чтобы устранить размытие движущихся объектов [20]. Данные действия дают возможность получить дополнительную информацию для расширения динамического диапазона, однако увеличивают уровень шума, для компенсации которого необходимо объединение серии кадров (рис. 6).



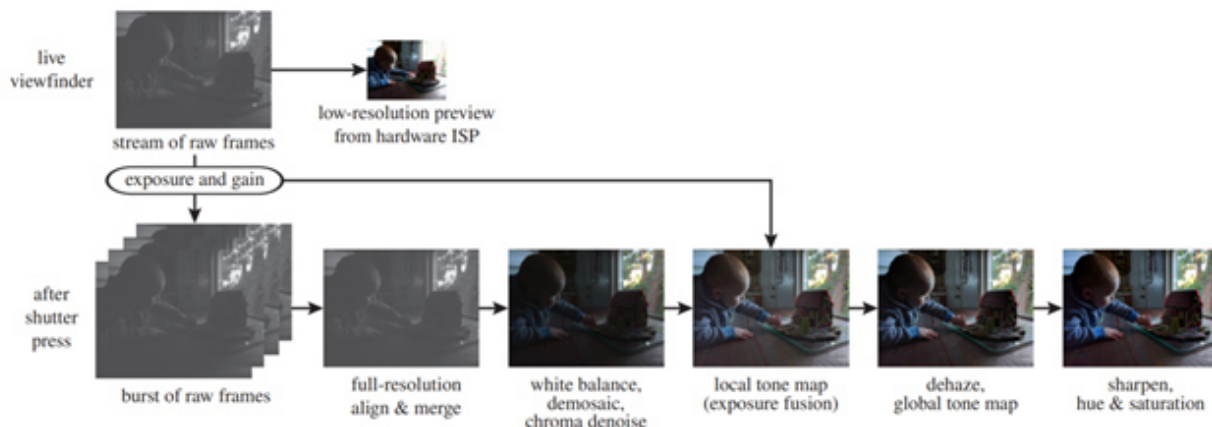


Рисунок 6 – Схема работы нейронной сети для объединения соседних кадров

При объединении нескольких соседних кадров получается промежуточное изображение с повышенной битовой глубиной, большим динамическим диапазоном и меньшим уровнем шума по сравнению с исходными кадрами. Данный подход позволяет получить высококачественное изображение с повышенными контрастностью и резкостью, а также устранить небольшие размытия. Все это дает дополнительную полезную информацию для идентификации личности, но требует более длительного времени порядка 1-2 секунд для объединения соседних кадров.

Заключение. Описаны основные методы для предварительной обработки изображений с целью повышения их качества, и как следствие, повышения достоверности идентификации личности. Показано, что наиболее эффективным по быстродействию и качеству получаемых изображений является метод «*Super Resolution*», но данный метод применим только к неподвижным объектам.

Подход Sam Hasinoff демонстрирует достаточно высокое качество изображений даже в случаях небольшого размытия перемещающегося объекта, но сопряжен со значительными затратами времени из-за сложности процесса объединения кадров, в связи с чем требует оптимизации для сценария идентификации личности. В то же время этот способ следует признать перспективным, если дополнить его процедурами предварительного извлечения изображений лиц из ряда соседних кадров с дальнейшим объединением информации о них, что позволит значительно сократить время объединения соседних кадров.

СПИСОК ЛИТЕРАТУРЫ:

1. Максименко, В. Н. Анализ системы распознавания лиц по алгоритму нейронной сети / В. Н. Максименко, Т. С. Волошина // Экономика и качество систем связи. – 2018. – № 4(10). – С. 31-37.
2. Юрко, И. В. Исследование основных проблем, связанных с распознаванием и идентификацией лиц по видеофиксации и совершенствование работы алгоритмов распознавания лиц по видеофиксации в режиме реального времени / И. В. Юрко, В. Н. Алдобаева // Проблемы современной науки и образования. – 2018. – № 6(126). – С. 36-42.
3. Трофименко, Я. М. Алгоритмы предобработки изображений в системе идентификации лиц в видеопотоке / Я. М. Трофименко, Л. Н. Виноградова, Е. В. Ершов // Вестник Череповецкого государственного университета. – 2019. – № 4(91).

– С. 21-29. – DOI 10.23859/1994-0637-2019-4-91-2.

4. Небаба, С. Г. Методы оценки и подготовки изображений в видеопотоке к распознаванию объектов / С. Г. Небаба // Труды Международной конференции по компьютерной графике и зрению "Графикон". – 2018. – № 28. – С. 450-453.
5. Onishi R., Sugiyama D., Matsuda K. Super-resolution simulation for real-time prediction of urban micrometeorology // SOLA. – 2019.
6. Zhou Q. et al. Edge-preserving single image super-resolution // Proceedings of the 19th ACM international conference on Multimedia. – 2011. – С. 1037-1040.
7. Goyal P., Gupta S. C. Statistical method for image registration & enhancement through super-resolution.
8. Rasti P. et al. A new low-complexity patch-based image super-resolution // IET Computer Vision. – 2017. – Т. 11. – №. 7. – С. 567-576.
9. Dong C. et al. Image super-resolution using deep convolutional networks // IEEE transactions on pattern analysis and machine intelligence. – 2015. – Т. 38. – №. 2. – С. 295-307.
10. Ledig C. et al. Photo-realistic single image super-resolution using a generative adversarial network // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 4681-4690.
11. Zhang W. et al. RankSRGAN: Super Resolution Generative Adversarial Networks with Learning to Rank // arXiv preprint arXiv:2107.09427. – 2021.
12. Guo Y. et al. Super-Resolution Image Reconstruction Based on Self-Calibrated Convolutional GAN // arXiv preprint arXiv:2106.05545. – 2021.
13. Goodfellow I. et al. Generative adversarial networks // Communications of the ACM. – 2020. – Т. 63. – №. 11. – С. 139-144.
14. O'Shea K., Nash R. An introduction to convolutional neural networks // arXiv preprint arXiv:1511.08458. – 2015.
15. Wei J. et al. PeerGAN: Generative Adversarial Networks with a Competing Peer Discriminator // arXiv preprint arXiv:2101.07524. – 2021.
16. Choi J., Han B. MCL-GAN: Generative Adversarial Networks with Multiple Specialized Discriminators // arXiv preprint arXiv:2107.07260. – 2021.
17. Hasinoff S. W. et al. Burst photography for high dynamic range and low-light imaging on mobile cameras // ACM Transactions on Graphics (ToG). – 2016. – Т. 35. – №. 6. – С. 1-12.
18. Gharbi M. et al. Deep bilateral learning for real-time image enhancement // ACM Transactions on Graphics (TOG). – 2017. – Т. 36. – №. 4. – С. 1-12.
19. Wei K. et al. A physics-based noise formation model for extreme low-light raw denoising // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020. – С. 2758-2767.
20. Dahary O., Jacoby M., Bronstein A. M. Digital Gimbal: End-to-end Deep Image Stabilization with Learnable Exposure Times // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2021. – С. 11936-11945.

Статья поступила в редакцию 10.11.2021

Статья принята к публикации 07.12.2021