

УДК 004.93

ИСПОЛЬЗОВАНИЕ АЛГОРИТМОВ ОСРДЛЯ РЕШЕНИЯ ЗАДАЧИ ВЕРИФИКАЦИИ ПЕРСОНАЛЬНЫХ ДОКУМЕНТОВ

Скурко И.А.

Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь

Научный руководитель: Тонкович И.Н. – канд.хим.наук, доцент, доцент кафедры ПИКС

Аннотация. Рассмотрен алгоритм выделения машиносчитываемых зон на фотографии персонального документа. Для оптического распознавания символов внутри машиносчитываемых зон предложена графовая глубокая нейронная сеть, показывающая наивысший показатель по гармоническому среднему.

Ключевые слова: оптическое распознавание символов, компьютерное зрение, нейронные сети

Введение. С расширением сферы применения электронной коммерции многие организации в той или иной мере сталкиваются с необходимостью автоматизации ручной обработки документов, представленных в электронном формате, для последующего ввода и проверки содержащейся в них информации. Автоматизация данного процесса требует использования технологий в области машинного обучения и компьютерного зрения. Так, эксперты крупнейшего независимого маркетингового и исследовательского агентства *MarketResearch.biz* прогнозируют рост мирового рынка оптического распознавания символов на 14,99% в течение десяти следующих лет.

Персональные документы, представленные в электронном формате, поступают в виде фотографий или сканов. Автоматизированная проверка содержимого документов, чтение данных и последующая верификация осуществляется на основе алгоритмов оптического распознавания символов (*OCR*). В данной работе предлагается использовать алгоритм выделения машиносчитываемых зон на фотографии документа с применением графовой глубокой нейронной сети для оптического распознавания символов.

Основная часть. Задача верификации персональных документов решается путем нахождения на фотографии удостоверения личности машиносчитываемых зон (*MRZ*) и последующего распознавания текста или кода внутри них. Страницы паспорта гражданина Республики Беларусь соответствуют единому международному стандарту транслитерации *ICAO Doc 9303* [1].

Это позволяет находить местонахождение сведений о личности владельца вычислительным путем. Для обнаружения машиносчитываемых зон на фотографии страницы паспорта используются базовые методы обработки изображений: методы уменьшения шума, методы выделения контрастных областей и морфологические преобразования. Данные методы обработки изображений могут быть реализованы с использованием мультипарадигменного языка программирования *Python* с применением библиотеки компьютерного зрения *OpenCV*. Алгоритм преобразования состоит в последовательности выполнения следующих шагов:

- 1 Применение размытия по Гауссу для устранения высокочастотного шума.
- 2 Использование морфологического оператора *black hat* для выделения контрастных зон.
- 3 Вычисление градиента яркости с помощью оператора Собеля.
- 4 Выполнение операции закрытия для заполнения промежутков между зонами отдельных символов по методу Оцу.
- 5 Вычисление координат прямоугольников (*bounding boxes*), заключенных вокруг машиносчитываемых зон, через определение контуров на двухуровневом изображении.

Полученные координаты определяют местонахождение личных паспортных данных.

Далее эти данные читаются с помощью одного из алгоритмов оптического распознавания символов.

Наибольшую производительность в оптическом распознавании символов на практике показывают глубокие нейронные сети [2]. Сравнение нейронных сетей по метрикам точности распознавания приведено в таблице 1.

Таблица 1 – Сравнение нейронных сетей для распознавания тестов

Нейронная сеть	Полнота (recall)	Точность (precision)	Гармоническое среднее (h-mean)
DBNet	0.731	0.871	0.795
DRRG	0.822	0.858	0.840
TextSnake	0.795	0.840	0.817
MaskRCNN	0.753	0.712	0.732
PANet	0.734	0.856	0.791

Каждая нейронная сеть была обучена на одних и тех же данных, полученных из базы данных изображений *ImageNet*. Для оптического распознавания символов внутри заранее определенных машиносчитываемых зон наиболее высоко результатной нейронной сетью является графовая глубокая нейронная сеть (*DRRG*), показывающая наивысший показатель по гармоническому среднему.

Нейронные сети уступают вероятностным моделям распознавания символов в скорости работы и в объеме требуемых ресурсов. Скрытая марковская модель (*НММ*) основана на составлении набора состояний, которые соединены между собой набором вероятностей переходов [3]. Данная модель возвращает значения вероятностей возникновения символов текста. Наиболее популярными методами по обучению модели являются метод максимального правдоподобия и алгоритм Баума-Велша, который, в свою очередь, использует алгоритм прямого-обратного хода. Скрытая марковская модель способна распознавать текст внутри машиносчитываемых зон с точностью до 80% [4].

Заключение. Применение рассмотренного алгоритма обнаружения машиносчитываемых зон на фотографии персонального документа с применением графовой глубокой нейронной сети для оптического распознавания символов обеспечит полную автоматизацию процесса верификации персональных документов.

Список литературы

1. Закон Республики Беларусь от 28 декабря 2009 года № 113-З «Об электронном документе и электронной цифровой подписи» в редакции от 08.11.2018.
2. Davies, E.R. *Computer Vision: Principles, Algorithms, Applications, Learning* / E.R. Davies. – London ; University of London : Academic Press, 2018. – 62 p.
3. Bonaccorso, G. *Mastering Machine Learning Algorithms* / Bonaccorso G. – Birmingham : Packt Publishing, 2018. – 127 p.
4. Hiral Modi, M. C. Parikh. *A Review on Optical Character Recognition Techniques* – London : International Journal of Computer Applications, 2017. – 47 p.

UDC 004.93

USING OCR ALGORITHMS FOR SOLVING THE PROBLEM OF PERSONAL DOCUMENTS VERIFICATION

Skurko I.A.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Tonkovich I.N. – PhD, assistant professor, associate professor of the department of ICSD

Annotation. An algorithm for selecting machine readable zones in a photograph of a personal document is considered. For optical character recognition within machine readable zones, a graph deep neural network is proposed, showing the highest indicator in terms of harmonic mean.

Keywords. optical character recognition, computer vision, neural networks