

ТЕХНИЧЕСКОЕ ЗРЕНИЕ И РАСПОЗНАВАНИЕ ОБРАЗОВ

УДК 004.932.75'1

НОВЫЙ КРИТЕРИЙ ОБУЧЕНИЯ НЕЙРОСЕТЕВОГО ЭНКОДЕРА В ЗАДАЧЕ СЕГМЕНТАЦИИ СТРОКИ НА СИМВОЛЫ

© 2019 г. А. В. Шешкус^{1,2,*}, Ю. С. Чернышова^{1,2}, А. Н. Чирвоная³, Д. П. Николаев⁴

¹ ООО “Смарт Энджинс Сервис”, 117312 Москва, проспект 60-летия Октября, 9, Россия

² Федеральное государственное учреждение “Федеральный исследовательский центр “Информатика и управление” Российской академии наук”, 119333 Москва, Вавилова, д. 44, кор. 2, Россия

³ Национальный исследовательский технологический университет “МИСиС”
119049 Москва, Ленинский проспект, д. 4, Россия

⁴ Институт проблем передачи информации им. А.А. Харкевича РАН
127051 Москва, Большой Каретный пер., 19, Россия

*E-mail: asheshkus@smartengines.biz

Поступила в редакцию 10.09.2018 г.

В данной работе рассматривается проблема сегментации текста на отдельные буквы. Зачастую для решения задачи сегментации применяются различные специализированные алгоритмы, использующие методы обработки изображений. Однако такие алгоритмы бывают сложны и слишком неробастны по отношению к алфавиту и шрифту сегментируемого текста и не решают задачу полностью. Существуют также нейросетевые методы, у которых такие недостатки отсутствуют, но классические функции потерь плохо подходят для поставленной задачи. Мы предлагаем способ обучения сверточного нейросетевого энкодера вместо классической схемы для построения оценок наличия разрезов между символами. В результате работы показана эффективность данного подхода на примере задачи сегментации строк полей паспорта гражданина РФ и показано преимущество предлагаемого метода обучения в сравнении с классическим.

Ключевые слова: сверточные нейронные сети, глубокое обучение, сегментация текста, кумулятивная сумма

DOI: 10.1134/S0235009219020094

ВВЕДЕНИЕ

Задача сегментации строки на символы является одной из основных при распознавании текста и остается актуальной на протяжении последних десятилетий (Casey, Lecolinet, 1996; Jacobsetal., 2005; Jiaetal., 2017). Сегментация может проводиться на изображениях в градациях серого (Grafmuller, Beyerer, 2013; Jacobsetal., 2005; Jiaetal., 2017; Maietal., 2016), на бинаризованных (Manwatkar, Yadav, 2015; Meietal., 2013; Radhiahetal., 2018; Sahare, Dhok, 2018; Vuckovic, Azinovic, 2017) или же на предварительно преобразованных (Chomphuwiset, 2017) изображениях.

Можно выделить два основных подхода к сегментации строки на символы: поиск компонент связности и анализ проекций. Анализ вертикальной проекции включает большое количество различных методов. Самой базовой версией можно считать поиск на бинаризованном изображении столбцов, не содержащих черных пикселей (Manwatkar, Yadav, 2015; Ryan, Hanafiah, 2015). Однако на изображениях, полученных в слабо контролируемых условиях, возникает множество искаже-

ний, которые после бинаризации приводят к зашумлению фона. В результате на проекциях возникает много ложных срабатываний, которые ведут к возникновению лишних разрезов и к потере настоящих, а также вызывают проблемы при сегментации курсива, прописных букв, а также лигатур и соприкасающихся символов, лишь частично решаемые с применением эвристик (Maietal., 2016; Mollahetal., 2010; Rehman, 2017; Ryan, Hanafiah, 2015; Shivakumaraetal., 2011). Ряд авторов предлагают анализировать проекции, учитывая результаты распознавания выделенных регионов. Так, в работе (Jacobsetal., 2005) распознаются все регионы строки с заданным шагом, а потом при помощи динамического программирования определяется лучшая кандидатура. В работе (Jiaetal., 2017) авторы строят вертикальную проекцию градиента, содержащую много потенциальных разрезов. Следующим шагом они строят дерево сегментации, содержащее все возможные пути из первого в последний разрез. В качестве ответа выбирают вариант с наибольшим весом, который рассчитывается исходя из оценок

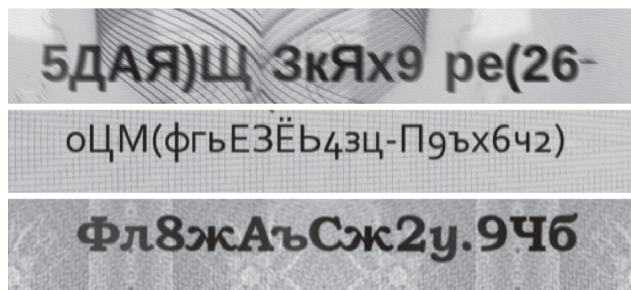


Рис. 1. Примеры изображений из обучающей выборки.

разрезов и уверенности распознавания символов. В работе (Чернов и др., 2016) анализ вертикальной проекции сравнивается с сегментацией при помощи машинного обучения. По результатам тестирования искусственные нейронные сети показывают результат лучший, чем классический “инженерный” подход.

Много работ рассматривают проблемы работы стандартных методов сегментации на символы для нетипичных западных пользователей языков, прежде всего арабского и китайского. Отличительной особенностью арабского языка является прежде всего связность букв (Firdausetal., 2017). Как правило, методы сегментации арабского языка основаны на том факте, что буквы соединены на уровне базовой линии (Radhiahetal., 2018), которую допустимо разделить. В случае китайского языка возможно не только соединение соседних иероглифов в одну связную компоненту, но и противоположная ситуация, когда один иероглиф состоит из нескольких несоприкасающихся радикалов (Meietal., 2013). Для решения этой проблемы используют в основном эвристические подходы, основанные на “квадратности” иероглифов, т.е. сначала разделяют символы на отдельные радикалы, а потом удаляют лишние разрезы так, чтобы отношение ширины и высоты каждого символа было примерно равно единице (Qietal., 2008).

При усложнении входных данных, например, если необходимо сегментировать текст, содержащий несколько алфавитов и знаки препинания, количество особенностей, которое нужно учитывать при построении оценки разреза, возрастает. При этом для получения стабильно качественного ответа на разных входных данных каждая составляющая суммарной оценки должна хорошо коррелировать с результирующим показателем. Кроме того, необходимо нормировать эти величины для совместного использования в общем оптимизируемом функционале. С увеличением количества компонент результирующей оценки растет и сложность их комбинирования, что, во-первых, может приводить к неожиданным или противоречащим здравому смыслу результатам, а во-вторых, требовать излишне много вычисли-

тельных мощностей. Таким образом, с одной стороны, хочется получить ответ, учитывающий все особенности входных данных, а с другой – опираться на разумно небольшое количество оценок.

Как уже отмечалось выше, алгоритм сегментации должен правильно работать и на касающихся, и на состоящих из нескольких частей (например, буква “Ы” в кириллице) символах. Мы предлагаем строить оценки разрезов, основываясь не на содержании каждого отдельного столбца, а на некоторой его окрестности. В такой постановке задача сегментации становится похожа на задачу распознавания образов, где уже много лет в качестве классификаторов успешно применяются сверточные нейронные сети (СНС). Мы считаем, что данный подход имеет большой потенциал в задаче сегментации текста на символы, так как в зависимости от обучающей выборки, СНС можно научить учитывать особенности письменностей, а также сделать ее нечувствительной к относительному сдвигу компонент символов и зашумлению изображения. В данной работе мы покажем, как обучить сегментирующую СНС и проверим ее применимость на строках, полученных с полей паспорта гражданина РФ.

ОБУЧАЮЩИЕ И ТЕСТОВЫЕ ДАННЫЕ

Для обучения нейронной сети мы использовали искусственные данные, созданные по методике (Chernyshovaetal., 2018). Такие данные хорошо подходят для обучения, так как не требуют ручной разметки и не содержат потенциальных ошибок, которые могли бы допустить эксперты. Примеры изображений, использованных для обучения, представлены на рис. 1. Всего было сгенерировано 50000 изображений строк по 100 случайных символов. Искусственные данные для обучения были выбраны по двум причинам: их можно сгенерировать в любое заданное количество и настоящие данные выгоднее использовать для тестирования.

Для тестирования обученной сети мы использовали набор из 11500 строк, собранных с полей “Кем выдано” и “Место рождения” паспорта гражданина РФ. Некоторые примеры из тестовой базы показаны на рис. 2.

Все данные перед подачей на вход нейронной сети были приведены к заранее заданной высоте (но по ширине они не изменялись). В данной работе высота составляла 33 пикселя.

Для каждой картинке из тестового набора данных был создан текстовый файл, содержащий значение распознаваемой строки. Для обучающей базы в дополнение к значению строки еще были известны положения разрезов между буквами.

АРХИТЕКТУРА СЕТИ

При создании архитектуры сети мы основывались на следующих соображениях:

- сеть должна быть полносверточной для удобного применения к строкам с заранее неизвестной длиной;
- на выходе сети должно быть изображение $w \times 1 \times 1$, где в каждой точке находится оценка разреза на изображении в соответствующей координате;
- количество обучаемых параметров сети должно быть не велико для уменьшения эффектов переобучения и для быстрой работы.

В итоге была создана сеть с архитектурой, представленной в табл. 1. Количество обучаемых параметров данной сети – 1741.

Важно добавить, что в качестве “идеального” ответа сети использовалось изображение, отражающее правильное расположение разрезов, т.е. изображение размера $w \times 1 \times 1$ (где w – ширина картинка), содержащее единичные значения в точках, координата x которых соответствует положению разреза на основном изображении, а во всех остальных – нули. Таким образом, сеть состояла из двух ветвей, одна из которых – обучаемая – содержала сверточные слои, а необучаемая – лишь передавала идеальное значение для вычисления функции потерь.

ПРЕДЛАГАЕМЫЙ КРИТЕРИЙ

Первые эксперименты обучения подобных сетей мы проводили, оптимизируя L2 норму между векторами, полученными от двух ветвей сети. Результат обучения сети таким способом не показал высокого качества, и впоследствии анализа процесса обучения было обнаружено несколько проблем. Первая из них заключается в том, что из-за сравнительно небольшого количества позиций в строке с разрезом относительно позиций в строке без разреза нейронная сеть учится говорить везде “нет разреза”, так как потенциальный рост оптимизируемого функционала из-за ложных сраба-

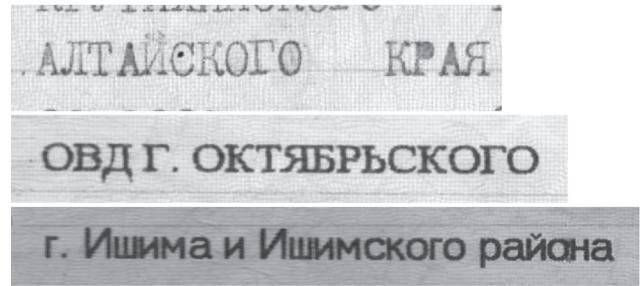


Рис. 2. Примеры изображений из тестовой выборки.

тываний слишком велик. Вторая проблема, связанная с подобным функционалом обучения, заключалась в том, что ошибка ответа в 1 пиксель вносит такой же штраф, как и ошибка в любое другое количество пикселей, что противоречит здравому смыслу в данной задаче: логично предположить, что допустима некоторая погрешность в ответах.

Для решения этих проблем был разработан слой кумулятивной суммы. После ее применения по отдельности к полученному и к ожидаемому ответу уже проводилась минимизация L2 нормы между ними. Таким образом, мы внесли в критерий обучения сети два фактора: количество разрезов в обоих векторах должно быть примерно одинаково, иначе оптимизируемый функционал будет очень велик, и небольшие отклонения от идеальных разрезов приводят к меньшему росту оптимизируемого функционала, чем большие отклонения. На рис. 3 показаны четыре вектора. При обучении классическим методом происходит минимизация расстояния между векторами из первой строки, тогда как при обучении предлагаемым методом минимизируется расстояние между векторами второй строки.

Стоит упомянуть еще одну деталь, имеющую отношение к реализации слоя. В нейронной сети такой слой был реализован в виде линейного оператора (право-верхне-треугольной матрицы):

Таблица 1. Архитектура сети

№ слоя	Тип	Описание
1	Сверточный	4 фильтра 5×5 , без отступов, шаг фильтра 2×1 , функция активации: relu
2	Сверточный	4 фильтра 5×3 , без отступов, шаг фильтра 1×2 , функция активации: relu
3	Сверточный	4 фильтра 5×3 , без отступов, шаг фильтра 1×1 , функция активации: relu
4	Сверточный	4 фильтра 3×3 , без отступов, шаг фильтра 1×1 , функция активации: relu
5	Сверточный	4 фильтра 5×3 , без отступов, шаг фильтра 1×1 , функция активации: relu
6	Сверточный	4 фильтра 5×3 , без отступов, шаг фильтра 1×1 , функция активации: relu
7	Сверточный	4 фильтра 5×3 , без отступов, шаг фильтра 1×1 , функция активации: relu
8	Сверточный	4 фильтра 5×3 , без отступов, шаг фильтра 1×1 , функция активации: relu
9	Сверточный	1 фильтр 3×2 , без отступов, шаг фильтра 1×1 , функция активации: rbf



Рис. 3. Визуализация применения кумулятивной суммы к идеальному вектору (слева) и к полученному (справа).

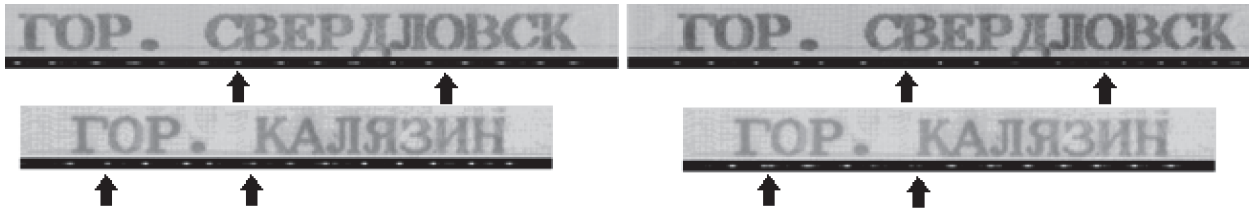


Рис. 4. Ответы сети, обученной с кумулятивной суммой (слева) и обычным способом (справа).

$$v_0 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} v_i \quad (1)$$

Если вспомнить, что при обратном проходе для распространения градиента линейные операторы транспонируются, то выгодно применить следующий трюк: считать не прямую кумулятивную сумму, а обратную (лево-верхне-треугольная матрица):

$$v_0 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 1 & 0 & \dots & 0 \end{pmatrix} v_i \quad (2)$$

Таким образом, матрица линейного оператора будет верхне-лево треугольной и при транспонировании перейдет сама в себя.

РЕЗУЛЬТАТЫ

Обе нейронные сети были проверены на тестовом наборе данных. Нейронная сеть, обученная по предлагаемой схеме, менее склонна пропускать разрезы и имеет более высокое качество. На рис. 4 представлены ответы обеих сетей, по которым видно, что сеть, обученная с кумулятивной

суммой, ставит разрезы точнее. Стрелками на рисунке отмечены основные отличия в ответах.

Для численного замера эффективности работы подобной сети мы построили ее в существующую систему распознавания (Vulatov et al., 2018) и замерили качество распознавания. Результаты экспериментов приведены в табл. 2.

Из результатов видно, что как посимвольное, так и построчное качество распознавания выше с сетью, обученной с использованием предложенного критерия. Анализ ошибок показал, что сеть, обученная предлагаемым способом, позволяет добиться меньшего количества ошибок распознавания.

ЗАКЛЮЧЕНИЕ

В данной работе была рассмотрена проблема сегментации текста на символы и ее решение с использованием нейросетевого энкодера для построения оценок наличия разрезов в каждой из позиций строки. Была предложена модификация оптимизируемого функционала, которая позволила улучшить точность сегментации. Эксперименты показали, что использование предложенной модификации позволяет сократить количество ошибок распознавания в 2 раза.

Дальнейшие работы планируются в основном по двум направлениям. Первое направление – исследование зависимости полученного улучшения качества от сложности сети. Второе – разделение ответов сети с “разрезы” на “начала” и “концы” букв, что должно привести к дальнейшему увеличению робастности метода.

Работа выполнена при финансовой поддержке РФФИ в рамках научных проектов № 17-29-07092 и № 18-07-01384.

Таблица 2. Результаты экспериментов

Нейронная сеть	Посимвольное качество, %	Построчное качество, %
Классический способ	98.8	86.3
Кумулятивная сумма	99.4	91.0

СПИСОК ЛИТЕРАТУРЫ

- Чернов Т.С., Ильин Д.А., Безматерных П.В., Фараджев И.А., Карпенко С.М. Исследование методов сегментации изображений текстовых блоков документов с помощью алгоритмов структурного анализа и машинного обучения. *Вестник РФФИ*. 2016. № 4 (92). С. 55–71. DOI: 10.22204/2410-4639-2016-092-04-55-71
- Bulatov K., Arlazarov V.V., Chernov T., Slavin O., Nikolaev D.P. Smart IDReader: Document Recognition in Video Stream. *The 14th IAPR International Conference on Document Analysis and Recognition*. 2018. P. 39–44. DOI: 10.1109/ICDAR.2017.347
- Casey R.G., Lecolinet E. A Survey of Methods and Strategies in Character Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1996. V. 18 (7). P. 690–706. DOI: 10.1109/34.506792
- Chernyshova Y., Gayer A., Sheshkus A. Generation method of synthetic training data for mobile OCR system. *Proc. SPIE 10696, Tenth International Conference on Machine Vision*. 2018. P. 1–7. DOI: 10.1117/12.2310119
- Chomphuwiset P. Printed Thai Character Segmentation and Recognition. *4th International Conference on Soft Computing & Machine Intelligence*. 2017. P. 123–127. DOI: 10.1109/ISCMCI.2017.8279611
- Firdaus F.I., Khumaini A., Utaminigrum F. Arabic Letter Segmentation using Modified Connected Component Labeling. *International Conference on Sustainable Information Engineering and Technology*. 2017. P. 392–397. DOI: 10.1109/SIET.2017.8304170
- Grafmuller M., Beyerer J. Performance improvement of character recognition in industrial applications using prior knowledge for more reliable segmentation. *Expert Systems with applications*. 2013. V. 40 (17). P. 6955–6963. DOI: 10.1016/j.eswa.2013.06.004
- Jacobs C.E., Rinker J.R., Simard P.Y., Viola P.A. *Low resolution OCR for camera acquired documents*. US Application. No. US20050259866A1. 2005.
- Jia F., Shi C., Wang Y., Wang C., Xiao B. Grayscale-projection based Optimal Character Segmentation for Camera-captured Faint Text Recognition. *14th IAPR International Conference on Document Analysis and Recognition*. 2017. P. 1301–1306. DOI: 10.1109/ICDAR.2017.214
- Manwatkar P.M., Yadav S.H. Text Recognition from Images. *IEEE International Conference on Innovations in Information, Embedded and Communication Systems*. 2015. P. 1–6.
- Mai B.Q.L., Huynh T.H., Doan A.D. An Independent Character Recognizer for Distantly Acquired Mobile Phone Text Images. *International Conference on Advanced Technologies for Communications (ATC)*. 2016. P. 85–90. DOI: 10.1109/ATC.2016.7764836
- Mei Y., Wang X., Wang J. An Efficient Character Segmentation Algorithm for Printed Chinese Documents. *Advanced Science and Technology Letters*. 2013. V. 22. P. 183–189.
- Mollah A.F., Basu S., Nasipuri M. Segmentation of Camera Captured Business Card Images for Mobile Devices. *International Journal of Computer Science and Applications*. 2010. V. 1 (1). P. 33–37.
- Qi W., Li X., Yang B. A Character Segmentation Method without Character Verification. *International Symposium on Intelligent Information Technology Application Workshops*. 2008. P. 581–584. DOI: 10.1109/IITA
- Radhiah A., Machbub C., Hidayat E.M.I., Prihatmanto A.S. Printed Arabic Letter Recognition Based on Images. *International Conference on Signals and Systems*. 2018. P. 86–91. DOI: 10.1109/ICSSIGSYS.2018.8373574
- Rehman A. Offline Touched Cursive Script Segmentation Based on Pixel Intensity Analysis. *Twelfth International Conference on Digital Information Management*. 2017. P. 324–327. DOI: 10.1109/ICDIM.2017.8244641
- Ryan M., Hanafiah N. An Examination of Character Recognition on ID card using Template Matching Approach. *Procedia Computer Science*. 2015. V. 59. P. 520–529. DOI: 10.1016/j.procs.2015.07.534
- Sahare P., Dhok S.B. Multilingual Character Segmentation and Recognition Schemes for Indian Document Images. *IEEE Access*. 2018. V. 6. P. 10603–10617. DOI: 10.1109/ACCESS.2018.2795104
- Shivakumara P., Bhowmick S., Su B., Tan C. L., Pal U. A New Gradient based Character Segmentation Method for Video Text Recognition. *International Conference on Document Analysis and Recognition*. 2011. P. 126–130. DOI: 10.1109/ICDAR.2011.34
- Vuckovic V., Azinovic B. General Character Segmentation Approach for Machine-Typed Documents. *Proceedings of 4th International Conference on Electrical, Electronics and Computing Engineering*. 2017. P. RT12.2.1-6.

New criteria for neural network encoder learning in the string segmentation problem

A. V. Sheshkus^{a,b,#}, Y. S. Chernyshova^{a,b}, A. N. Chirvonaya^c, and D. P. Nikolaev^d

^a Smart Engines Limited, 117312 Moscow, pr. 60-letiya Oktyabrya, 9, Russia

^b Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 119333 Moscow, Vavilova str., 44-2, Russia

^c National University of Science and Technology “MISIS”, 119049 Moscow, Leninsky pr., 4, Russia

^d Institute for Information Transmission Problems (Kharkevich Institute) RAS 127051 Moscow, Bolshoi Karetny per., 19, Russia

[#] E-mail: asheshkus@smartengines.biz

In this paper we consider the problem of character segmentation. Image processing methods are widely used in various specialized algorithms, which cannot solve the problem entirely. In some papers authors use differ-

ent customizations based on the structures on the image and improve the final segmentation on the basis of knowledge about the alphabet and the font. However, such algorithms can be complicated and too non-robust towards the alphabet and the font of the text. We suggest the neural network encoder learning approach instead of classic method for cuts estimation between characters. We demonstrate the effectiveness of this approach using strings from RF passport fields as an example and show the advantages of the suggested approach over the classic one.

Key words: convolutional neural networks, deep learning, text segmentation, cumulative sum

REFERENCES

- Chernov T.S., Il'in D.A., Bezmaternyh P.V., Faradzhev I.A., Karpenko S.M. Issledovanie metodov segmentatsii izobrazhenii tekstovuch blokov dokumentov pri pomochi algoritmov strukturnogo analiza I maschinogo obucheniy [Research of Segmentation Methods for Images of Document Textual Blocks Based on the Structural Analysis and Machine Learning] *Vestnik RFFI [RBRF Information Bulletin]*. 2016. № 4 (92). P. 55–71. DOI: 10.22204/2410-4639-2016-092-04-55-71 (In Russian).
- Bulatov K., Arlazarov V.V., Chernov T., Slavin O., Nikolaev D.P. Smart IDReader: Document Recognition in Video Stream. *The 14th IAPR International Conference on Document Analysis and Recognition*. 2018. P. 39–44. DOI: 10.1109/ICDAR.2017.347.
- Casey R.G., Lecolinet E. A Survey of Methods and Strategies in Character Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1996. V. 18 (7). P. 690–706. DOI: 10.1109/34.506792.
- Chernyshova Y., Gayer A., Sheshkus A. Generation method of synthetic training data for mobile OCR system. *Proc. SPIE 10696, Tenth International Conference on Machine Vision*. 2018. P. 1–7. DOI: 10.1117/12.2310119.
- Chomphuwiset P. Printed Thai Character Segmentation and Recognition. *IEEE 4th International Conference on Soft Computing & Machine Intelligence*. 2017. P. 123–127. DOI: 10.1109/ISCMI.2017.8279611.
- Firdaus F.I., Khumaini A., Utaminigrum F. Arabic Letter Segmentation using Modified Connected Component Labeling. *International Conference on Sustainable Information Engineering and Technology*. 2017. P. 392–397. DOI: 10.1109/SIET.2017.8304170.
- Grafmuller M., Beyerer J. Performance improvement of character recognition in industrial applications using prior knowledge for more reliable segmentation. *Expert Systems with applications*. 2013. V. 40(17). P. 6955–6963. DOI: 10.1016/j.eswa.2013.06.004.
- Jacobs C.E., Rinker J.R., Simard P.Y., Viola P.A. *Low resolution OCR for camera acquired documents*. US Application. No. US20050259866A1. 2005.
- Jia F., Shi C., Wang Y., Wang C., Xiao B. Grayscale-projection based Optimal Character Segmentation for Camera-captured Faint Text Recognition. *14th IAPR International Conference on Document Analysis and Recognition*. 2017. P. 1301–1306. DOI: 10.1109/ICDAR.2017.214
- Manwatkar P.M., Yadav S.H. Text Recognition from Images. *IEEE International Conference on Innovations in Information, Embedded and Communication Systems*. 2015. P. 1–6.
- Mai B.Q.L., Huynh T.H., Doan A.D. An Independent Character Recognizer for Distantly Acquired Mobile Phone Text Images. *International Conference on Advanced Technologies for Communications*. 2016. P. 85–90. DOI: 10.1109/ATC.2016.7764836.
- Mei Y., Wang X., Wang J. An Efficient Character Segmentation Algorithm for Printed Chinese Documents. *Advanced Science and Technology Letters*. 2013. V. 22. P. 183–189.
- Mollah A.F., Basu S., Nasipuri M. Segmentation of Camera Captured Business Card Images for Mobile Devices. *International Journal of Computer Science and Applications*. 2010. V. 1 (1). P. 33–37.
- Qi W., Li X., Yang B. A Character Segmentation Method without Character Verification. *International Symposium on Intelligent Information Technology Application Workshops*. 2008. P. 581–584. DOI: 10.1109/IITA.
- Radhiah A., Machbub C., Hidayat E.M.I., Prihatmanto A.S. Printed Arabic Letter Recognition Based on Images. *International Conference on Signals and Systems*. 2018. P. 86–91. DOI: 10.1109/ICSIGSYS.2018.8373574.
- Rehman A. Offline Touched Cursive Script Segmentation Based on Pixel Intensity Analysis. *Twelfth International Conference on Digital Information Management*. 2017. P. 324–327. DOI: 10.1109/ICDIM.2017.8244641.
- Ryan M., Hanafiah N. An Examination of Character Recognition on ID card using Template Matching Approach. *Procedia Computer Science*. 2015. V. 59. P. 520–529. DOI: 10.1016/j.procs.2015.07.534.
- Sahare P., Dhok S.B. Multilingual Character Segmentation and Recognition Schemes for Indian Document Images. *IEEE Access*. 2018. V. 6. P. 10603–10617. DOI: 10.1109/ACCESS.2018.2795104.
- Shivakumara P., Bhowmick S., Su B., Tan C. L., Pal U. A New Gradient based Character Segmentation Method for Video Text Recognition. *International Conference on Document Analysis and Recognition*. 2011. P. 126–130. DOI: 10.1109/ICDAR.2011.34.
- Vuckovic V., Azinovic B. General Character Segmentation Approach for Machine-Typed Documents. *Proceedings of 4th International Conference on Electrical, Electronics and Computing Engineering*. 2017. P. RT12.2.1–6.