

УДК 004.85

DOI: 10.15827/2311-6749.18.4.10

## ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕВОГО ПОДХОДА ДЛЯ СЕГМЕНТАЦИИ СЛОВ В РАМКАХ ЗАДАЧИ ОФФЛАЙН-РАСПОЗНАВАНИЯ РУКОПИСНОГО ТЕКСТА

А.С. Басанько, магистрант, antonbas1@rambler.ru;

Ю.С. Белов, к.ф.-м.н., доцент, fn1-kf@mail.ru

(Калужский филиал ФГБОУ ВО «Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет)», г. Калуга, 248000, Россия)

Задача оффлайн-распознавания рукописного текста состоит из множества этапов, связанных между собой, так как результаты выполнения каждого этапа являются входными данными для следующего. Один из важных этапов – этап сегментации слова на символы, напрямую влияющий на дальнейший результат классификации. Задача сегментации слова самая сложная в рамках оффлайн-распознавания рукописного текста и на данный момент полностью не решена. Существуют различные подходы для ее решения, однако они довольно сложны в реализации и не обеспечивают качественных результатов.

В данной статье рассматривается относительно простой подход с применением нейронных сетей, позволяющий достичь хороших показателей сегментации. Идея подхода в том, чтобы с помощью гистограммного метода выделить набор потенциальных точек сегментации, а с помощью обученной нейронной сети оставить только нужные точки сегментации. В статье описываются гистограммный метод для выделения точек сегментации, процесс выделения признаков для обучения нейронной сети, а также алгоритм ее обучения. Приводятся результаты применения данного подхода.

**Ключевые слова:** сегментация рукописного слова, гистограммный метод, нейронные сети, перцептрон, алгоритм Левенберга-Марквардта.

Одним из этапов в рамках задачи оффлайн-распознавания рукописного текста является этап сегментации. На этом этапе сначала с предварительно обработанного изображения по необходимости извлекаются строки, затем строки разбиваются на слова и, наконец, каждое слово разбивается на отдельные символы, которые в дальнейшем будут классифицироваться [1].

Самой сложной задачей на этом этапе является сегментация слова на символы [2]. В различных методах иногда даже избегают решения этой задачи и пытаются классифицировать не отдельные символы, а целые слова.

В рамках данной статьи предлагается нейросетевой подход для сегментации слов на символы, который позволяет достичь вполне неплохих результатов.

Суть подхода в том, что сначала находятся все возможные точки сегментации в слове, а после этого с помощью обученной нейронной сети исключаются некорректные точки сегментации.

Возможные точки сегментации находятся достаточно простым гистограммным подходом в вертикальной проекции, который подразумевает подсчет черных пикселей на изображении для каждого столбца пикселей в изображении слова. Далее, исходя из этой гистограммы, можно найти пики, которые будут ниже некоторого установленного порогового значения и являться возможными точками сегментации (рис. 1).



Рис. 1. Выделение возможных точек сегментации в слове

В качестве архитектуры для нейронной сети используется многослойный перцептрон с двумя скрытыми слоями, состоящими из 20 и 14 нейронов соответственно [3, 4]. В выходном слое будет всего 1 нейрон, который при возбуждении будет показывать, что точка сегментации является корректной, иначе она некорректна (рис. 2).

В качестве входных данных используется информация о плотности черных пикселей

в области точки сегментации. Для этого берутся прямоугольные области слева и справа от возможной точки сегментации размером в 15–20 пикселей и делятся на 28 небольших прямоугольных подобластей – окон. В каждом окне считается плотность черных пикселей, как показано в примере на рисунке 3. Полученные в результате этого значения плотностей являются вектором признаков для точки сегментации, которые будут использоваться для работы нейронной сети и для ее тренировки [5, 6].

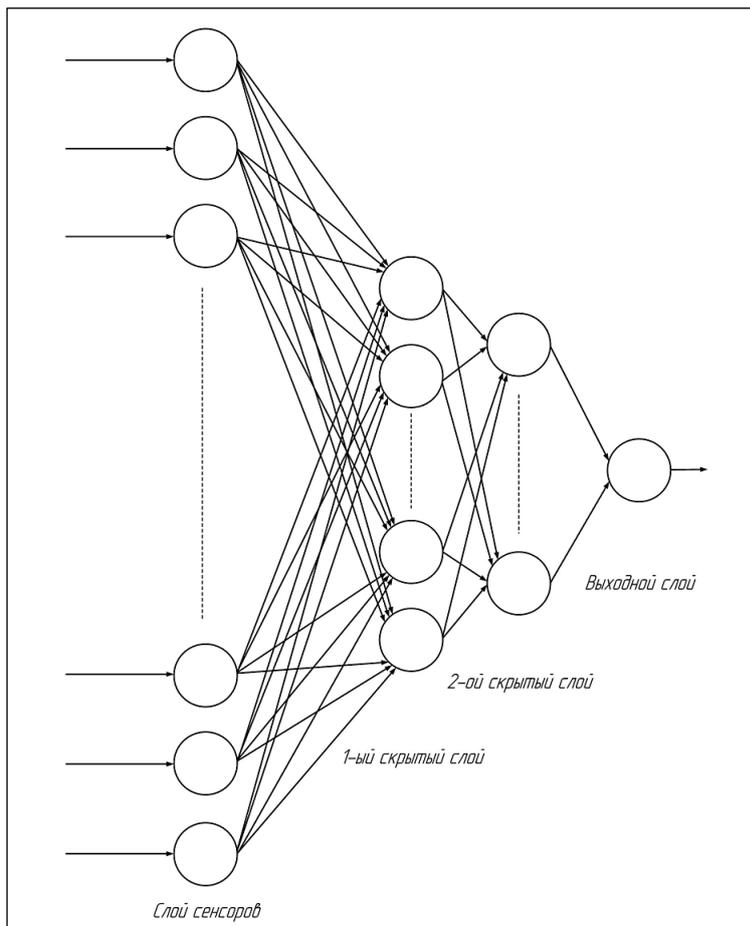


Рис. 2. Персептрон для определения корректности точки сегментации в слове

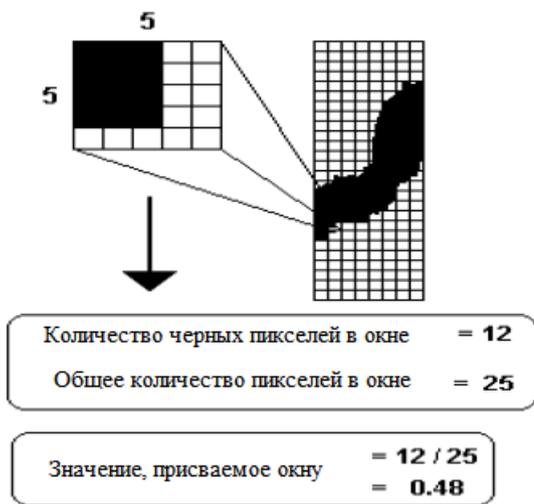


Рис. 3. Пример подсчета значения окна

Описанная архитектура сети и алгоритм обучения были подобраны путем оценки эффективности комбинаций различного количества слоев в нейронной сети и количества нейронов в каждом из скрытых слоев, а также использования разнообразных параметров алгоритмов обучения. В качестве возможных алгоритмов обучения рассматривались алгоритмы обратного распространения ошибки и Левенберга-Марквардта.

Алгоритм обратного распространения ошибки является наиболее популярным методом обучения, однако на практике он показал, что для обучения данной нейронной сети не совсем подходит. В целом алгоритму свойственны такие недостатки, как застревание в локальных минимумах и достаточно медленная сходимость. Для решения этих проблем обычно применяют градиентные алгоритмы второго порядка, к числу которых относится и алгоритм Левенберга-Марквардта.

Данный алгоритм является одним из приложений ньютоновской стратегии оптимизации. Основным выражением методов Ньютона является выражение

$$p_k = -[H(\omega_k)]^{-1} g(\omega_k), \quad (1)$$

где  $p_k$  – направление, гарантирующее достижение минимального для данного шага значения целевой функции;  $g(\omega_k)$  – значение градиента в точке последнего решения  $\omega_k$ ;  $H(\omega_k)$  – значение гессиана в точке последнего решения  $\omega_k$ .

При использовании алгоритма Левенберга-Марквардта точное значение гессиана  $h(\omega)$  в (1) заменяется аппроксимированным значением  $g(\omega)$ , которое рассчитывается на основе содержащейся в градиенте информации с учетом некоторого регуляризационного фактора [7].

Для обучения нейронной сети использовался набор из 50 различных слов, написанных несколькими людьми, которые предоставляют около 1 500 возможных точек сегментации. За счет использования алгоритма обучения Левенберга-Марквардта

процесс обучения занял небольшое количество времени по сравнению с алгоритмом обратного распространения ошибки и сошелся до небольшой величины ошибки [8].

Для тестирования обученной сети использовался набор из 10 слов, которые при этом не входили в обучающую выборку. В результате правильное разбиение слов на символы было получено в 85 % случаев. Наглядный пример разбиения слова на символы с помощью данной нейронной сети показан на рисунке 4.

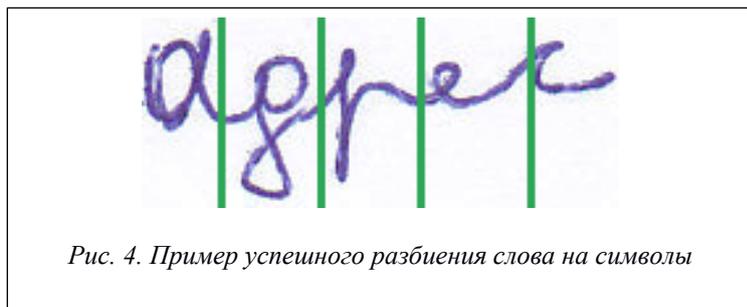


Рис. 4. Пример успешного разбиения слова на символы

Таким образом, в данной статье было рассмотрено использование нейросетевого подхода для сегментации символов в рамках задачи оффлайн-распознавания рукописного текста на примере русского языка. Предложенный подход показал довольно неплохие результаты и в целом может применяться в рассматриваемой задаче.

### Литература

1. Басанько А.С., Рыбкин С.В. Существующие проблемы распознавания рукописного текста // Научные исследования в области технических и технологических систем: сб. стат. Междунар. науч.-практич. конф. Казань. 2018. С. 25–28. URL: [https://elibrary.ru/download/elibrary\\_32246059\\_90455300.pdf](https://elibrary.ru/download/elibrary_32246059_90455300.pdf) (дата обращения: 15.10.2018).
2. Тетерин Д.М., Силаева А.Э. Проблемы распознавания рукописного текста с помощью нейронных сетей // Достижения естественных и технических наук в XXI веке: сб. науч. тр. по матер. Междунар. науч.-практич. конф. АПНИ. 2017. С. 212–215. URL: [https://elibrary.ru/download/elibrary\\_30407997\\_27709634.pdf](https://elibrary.ru/download/elibrary_30407997_27709634.pdf) (дата обращения: 09.11.2018).
3. Aiquan Yuan, Gang Bai, Po Yang, Yanni Guo, Xinting Zhao Handwritten English Word Recognition Based on Convolutional Neural Networks. Proc. Intern. Conf. on Frontiers in Handwriting Recognition, 2012, pp. 207–212. URL: <https://ieeexplore.ieee.org/document/6424393> (дата обращения: 20.10.2018).
4. Ибрафиллов Х.С. Применение нейронных сетей в распознавании рукописного текста // Молодой ученый. 2016. № 29. С. 24–27. URL: <https://moluch.ru/archive/133/37372/> (дата обращения: 02.12.2018).
5. Гришанов К.М., Белов Ю.С. Методы выделения признаков для распознавания символов // Наука, техника и образование. 2016. № 1. С. 110–119. URL: [https://elibrary.ru/download/elibrary\\_26747503\\_11900308.pdf](https://elibrary.ru/download/elibrary_26747503_11900308.pdf) (дата обращения: 11.11.2018).
6. Фазылов Ш.Х., Мирзаев Н.М., Каримов И.К., Даминов О.А. Выделение признаков в задаче распознавания рукописных текстов // Вестн. Камчатского гос. технич. ун-та. 2013. № 26. С. 38–42. URL: [https://elibrary.ru/download/elibrary\\_21157527\\_95947091.pdf](https://elibrary.ru/download/elibrary_21157527_95947091.pdf) (дата обращения: 28.10.2018).
7. Махотило К.В., Вороненко Д.И. Модификация алгоритма Левенберга-Марквардта для повышения точности прогностических моделей связанного потребления энергоресурсов в быту // Вестн. НТУ ХПИ. 2005. № 56. С. 83–90. URL: [https://elibrary.ru/download/elibrary\\_21345270\\_96427743.pdf](https://elibrary.ru/download/elibrary_21345270_96427743.pdf) (дата обращения: 26.11.2018).
8. Моченов С.В., Шаронов М.А., Ахметгалеев Р.Р. О возможности использования нейронных сетей при решении задач распознавания рукописных текстов // Информационные технологии в науке, промышленности и образовании: сб. тр. регионал. науч.-технич. конф. 2014. С. 185–190. URL: [https://elibrary.ru/download/elibrary\\_23192023\\_99074075.pdf](https://elibrary.ru/download/elibrary_23192023_99074075.pdf) (дата обращения: 01.12.2018).