

МЕТОДИКА РАСПОЗНАВАНИЯ РУКОПИСНОГО ТЕКСТА

Рассматриваются основные этапы распознавания рукописного текста. Предлагается использование однослойной сети Кохонена для автоматизации распознавания рукописного текста.

ВВЕДЕНИЕ

Распознавание рукописного ввода — это способность компьютера получать и интерпретировать интеллектуальный рукописный ввод. Распознавание текста может производиться «оффлайновым» методом из уже написанного на бумаге текста или «онлайновым» методом считыванием движений кончика ручки, к примеру, по поверхности специального компьютерного экрана. Оффлайновый вид распознавания успешно применяется в сферах деятельности, где необходимо обрабатывать большое количество рукописных документов, к примеру, в страховых компаниях. Качество распознавания можно повысить, используя структурированные документы (формы). Кроме того, можно улучшить качество, уменьшив диапазон возможных вводимых символов [1]. Оффлайновое распознавание считается более сложным по сравнению с онлайн-овым. Точное распознавание латинских символов в печатном тексте в настоящее время возможно только если доступны чёткие изображения, такие как сканированные печатные документы [2]. Абсолютная точность может быть достигнута только путем последующего редактирования человеком. Проблемы распознавания рукописного «печатного» и стандартного рукописного текста, а также печатных текстов других форматов (особенно с очень большим числом символов) в настоящее время являются предметом активных исследований.

I. ОСНОВНЫЕ ЭТАПЫ РАСПОЗНАВАНИЯ ТЕКСТА

Перечислим основные этапы распознавания текста:

Этап 1. Предобработка.

На этом этапе выполняются следующие задачи: повышение качества изображения за счет фильтрации, шумоподавления и других, имеющих своей целью повысить качество изображения. На этом этапе происходит очистка изображения от дефектов сканирования. В частности, в самом начале работы к изображению в целях шумоподавления часто применяется фильтр Гаусса. Важную роль играет пороговая бинаризация, то есть перевод изображения в чёрно-белый формат из цветного или оттенков серого [3]. Это позволяет резко разделить текст и фон, упрощает в дальнейшем применение многих алгоритмов, а

также избавляет от некоторых шумов на изображении. При этом используется гистограмма яркости изображения текста, на котором наблюдается два пика: высокий пик, соответствующий белому фону, то есть цвету бумаги, и пик в области тёмных пикселей, соответствующих яркости символов текста.

Этап 2. Выделение региона интереса.

На этом этапе на бинаризованном изображении выделяется непосредственно область, на которой находится распознаваемый текст, и отбрасываются элементы, текстом не являющиеся [3,4]. К ним относятся такие объекты, как кляксы, пятна на бумаге, не удалённые в процессе бинаризации, картинки и др. Для их удаления можно, например, выделять компоненты связности на изображении, вычислять геометрические признаки и на их основе классифицировать компоненту связности как часть текста или дефект, используя методы машинного обучения или эвристики.

Этап 3. Сегментация и нормализация текста.

На этом этапе текст разделяется, или сегментируется, на удобные для анализа составные части [5]. Наиболее естественными действиями на данном этапе является разделение текста на отдельные строки (сегментация строк) и разделение строк на слова (сегментация слов), а также, теоретически, разделение слов на элементарные составные части. Кроме того, на данном этапе проводится нормализация текста приведение выделенных составных частей к некоторому стандартному виду для снижения вариативности и упрощения распознавания.

Этап 4. Сегментация слов.

На этом этапе работы системы распознавания выделенные строки текста разделяются на отдельные слова. В отличие от машинописного текста, в котором расстояние между словами более-менее постоянно, а интервалы между символами внутри слова гораздо меньше, чем интервалы между словами, в рукописном тексте размер интервалов между словами может варьироваться в очень широких пределах. Компоненты связности текста, отнесённые к одной строке на предыдущем этапе работы системы распознавания, объединяются в слова на этом этапе.

Этап 5. Нормализация.

В силу высокой вариативности начертания слов их распознавание является очень сложным

процессом. Нормализация служит для приведения слова к некоторому стандартному виду без значительной потери информации, необходимой для распознавания. Одними из наиболее часто используемых методов нормализации является метод коррекции наклона слова от горизонтальной и вертикальной линии. Простейший метод коррекции по горизонтали состоит в выполнении максимизации его на некотором диапазоне. Существуют и другие методы нормализации, например, коррекция размера и выделение скелета текста, но они применяются реже.

II. РАСПОЗНАВАНИЕ РУКОПИСНОГО ТЕКСТА С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

Один из способов распознавания - использование нейронных сетей. Рассмотрим в данном случае однослойную сеть Кохонена.

Принцип работы нейронной сети таков, что, получив на входной слой нейронов новое изображение, сеть реагирует импульсом того или иного нейрона. Так как все нейроны поименованы значениями букв, следовательно, среагировавший нейрон и несет ответ распознавания. Углубляясь в терминологию сетей можно сказать, что нейрон помимо выхода имеет также множество входов. Данные входы описывают значение пикселя изображения. То есть, если имеется изображение 16×16 , входов у сети должно быть 256.

Каждый вход воспринимается с определенным коэффициентом и в результате, по окончании распознавания на каждом нейроне скапливается определенный заряд, чем заряд будет больше тот нейрон и испустит импульс.

Но чтобы коэффициенты входов были правильно настроены необходимо сначала обучить сеть. Этим занимается отдельный модуль обучения. Данный модуль берет очередное изобра-

жение из обучающей выборки и скармливает сети. Сеть анализирует все позиции черных пикселей и выравнивает коэффициенты минимизируя ошибку совпадения методом градиента, после чего определенному нейрону сопоставляется данное изображение.

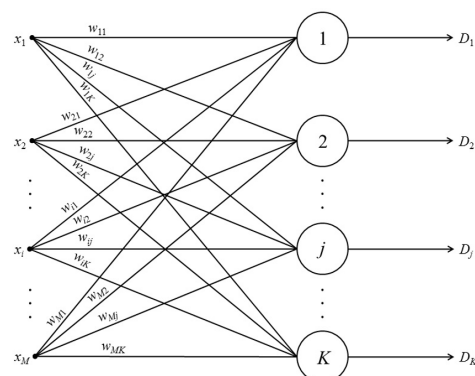


Рис. 1 – Общая структура нейронной сети Кохонена

III. Выводы

Предлагаемая нами методика может стать основой для разработки специальных приложений, предназначенных для автоматического распознавания рукописного текста.

1. А. Н. Абраменко Принципы распознавания // Компьютер-пресс. 1997. — № 12
2. Н. Г. Загоруйко Методы распознавания и их применение. — М.: Сов. радио, 1972. — 208 с.
3. У. Прэтт Цифровая обработка изображений. — М.: Мир, 1982. — Ч. 2 — 480 с.
4. Интернет. Сайт [openbooks.ifmo.ru](https://openbooks.ifmo.ru/ru/file/868/868.pdf)
<https://openbooks.ifmo.ru/ru/file/868/868.pdf>
5. И. И. Сальников Поэлементный анализ растровых изображений. — Пенза: Приволжский Дом знаний, 2015. — 180 с.

Бурак Анастасия Юрьевна, студент кафедры информационных технологий автоматизированных систем БГУИР, burak.67815@gmail.com.

Долголевец Алексей Васильевич, студент кафедры информационных технологий автоматизированных систем БГУИР, dolgolevets65@gmail.com.

Судаков Богдан Дмитриевич, студент кафедры информационных технологий автоматизированных систем БГУИР, sudakov.bogdan666@gmail.com.

Научный руководитель: Трофимович Алексей Фёдорович, старший преподаватель, заместитель декана ФИТУ trofimaf@bsuir.by.