

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК 004.[932.75'1+89]

Бурый  
Ярослав, Анатольевич

**Модель распознавания зашумленного и искаженного текста на основе  
модифицированной нейронной сети свертки**

**АВТОРЕФЕРАТ**

на соискание степени магистра технических наук  
по специальности 1-40 80 03 Вычислительные машины и системы

---

Научный руководитель

Новосёлова Наталья Анатольевна

кандидат технических наук, доцент

---

Минск 2016

## ВВЕДЕНИЕ

Точное распознавание латинских символов в печатном тексте в настоящее время возможно только если доступны чёткие изображения, такие как сканированные печатные документы. Точность при такой постановке задачи превышает 99%, 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.

Для решения более сложных проблем в сфере распознавания используются как правило интеллектуальные системы распознавания, такие как искусственные нейронные сети.

Задача распознавания текстовой информации при переводе печатного и рукописного текста в электронную форму является одной из важнейших составляющих любого проекта, имеющего целью автоматизацию документооборота или внедрение безбумажных технологий. Вместе с тем эта задача является одной из наиболее сложных и наукоемких задач полностью автоматического анализа изображений.

Моделирование работы нейронных сетей (НС) на сегодняшний день является одним из приоритетных направлений исследований в области искусственного интеллекта. Являясь, по сути, математической абстракцией, НС моделируют некоторые наиболее важные свойства высшей нервной деятельности, ключевыми из которых является способность к обучению и обобщению поступающей на вход информации.

Способность многослойных нейронных сетей, обученных методом градиентного спуска к построению сложных многомерных областей на основе большого числа обучающих примеров, позволяет применять их в качестве классификатора для распознавания символов. Однако если изображение текста было искажено и зашумлено, вариации написания символов содержат деформации. Свёрточная сеть – один из подвидов НС, который обладает описанными свойствами и гарантирует быстрое обучение и распознавание образов.

В качестве основного алгоритма распознавания выбрана модель нейронной сети свёртки ввиду её способности самостоятельного выделения и анализа информативных признаков путём обучения.

Выбранная модель устойчива к небольшим искажениям входного сигнала, а за счёт слоёв подвыборки и используемого алгоритма обучения количество связей в ней существенно меньше, по сравнению с другими типами нейронных сетей.

Указанные свойства позволяют модифицировать архитектуру выбранной нейронной сети с учётом возможных искажений входного сигнала, что позволит получить более качественное распознавание без дополнительных затрат процессорного времени при функционировании в режиме

распознавания.

В рассматриваемой магистерской диссертации предпринята попытка отойти от стандартного подхода к распознаванию с помощью обучающихся моделей, когда для устойчивости алгоритма к искажениям в обучающую выборку добавляются искажённые образы исходных сигналов. Вместо этого для достижения инвариантности распознающей системы к деформациям входного сигнала применяется модификация архитектуры нейронной сети.

Использование модели нейронной сети свёртки в качестве базового распознающего механизма позволяет объединить процессы выделения информативных признаков изображения текста и его распознавания в единой архитектуре.

Библиотека БГУИР

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

*Актуальность темы исследования* обусловлена тем, что задача распознавания текстов широко исследуется в области информационных технологий в связи с большим количеством практических приложений в области компьютерного зрения и автоматизации ввода информации в вычислительные устройства. Однако в настоящий момент не существует законченных решений по созданию эффективных систем распознавания текста, инвариантных к шуму и искажениям входного сигнала.

*Цель исследования:* разработка нейросетевой модели распознавания текстов, устойчивой к шуму и искажениям входного сигнала.

*Задачи исследования:* анализ существующих подходов к распознаванию текстов и внесению искажений в их изображения; разработка алгоритмов деформации изображения текстов; разработка модели распознавания текстов; тестирование устойчивости модели к искажениям входного сигнала.

*Объект исследования:* искажённое и зашумленное изображение текста.

*Предмет исследования:* метод модификации нейронной сети свёртки для обеспечения устойчивости к искажениям входного сигнала.

Основные положения работы и результаты диссертации изложены в опубликованной работе объёмом 4 страницы.

Структура диссертационной работы обусловлена целью, задачами и логикой исследования. Работа состоит из введения, трёх глав и заключения, списка использованной литературы и одного приложения. Общий объём диссертации 97 страница. Работа содержит 22 рисунков. Список использованной литературы включает 32 наименования.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении рассмотрено современное состояние проблемы распознавания искажённого и зашумленного текста, а также даётся обоснование актуальности темы диссертационной работы.

В общей характеристике работы сформулированы её цель и задачи, показана связь с научными исследованиями в данной области, даны сведения об объекте исследования и обоснован его выбор, представлены положения, выносимые на защиту, приведены сведения о личном вкладе соискателя, а также структура и объём диссертации.

В первой главе рассматривается научная литература по теме предобработки изображений в целях улучшения качества, обнаружения и сегментирования текста, распространённых алгоритмах распознавания текста и алгоритмах коррекции результатов распознавания. Также здесь рассматриваются типичные нейросетевые алгоритмы и описывается типичная сеть свёртки. Далее описываются модели когнитрона и неокогнитрона, а также алгоритмы искажения и зашумления изображений.

Во второй главе проводится анализ этапов проектирования и разработки модели распознавания искажённого и зашумленного текста, подробно описываются все модули системы, а также разработанные классы и используемые алгоритмы.

В третьей главе описывается проведение эксперимента с целью нахождения оптимальных параметров модификации нейронной сети, а также результаты тестирования подсистем.

В приложении приведён исходный текст разработанной модели распознавания искажённого и зашумленного текста. Также представлен графический материал в виде двух чертежей и одного плаката.

## ЗАКЛЮЧЕНИЕ

В соответствии с поставленными целями, в итоге проведенных исследований и разработок были получены следующие основные научные результаты:

Была создана действующая модель распознавания искажённого и зашумленного текста, которая в случае небольших искажений входных изображений сравнима по качеству распознавания с известными коммерческими системами оптического распознавания текста, а в случае существенных искажений – значительно превосходит их. Модули системы выполняют специализированные функции, такие как определение размера шрифта, сегментация строк, слов и, по возможности – отдельных символов изображения текста, а также распознавание символов.

Был разработан, реализован и протестирован алгоритм функционирования НС свёртки для распознавания символов. В качестве основного обучающего и тестового множества нейросетевого распознавателя символов в данной работе использовалась база обучающих эталонов MNIST, содержащая 60000 элементов для обучения модели и 10000 элементов для тестирования. Для обучения и тестирования модели распознавания текста использовались обучающие множества изображений символов и текста для различных шрифтов и начертаний, сформированные средствами платформы .Net.

Проектирование и конструирование производилось на языках Matlab и C#.

Было произведено упрощение алгоритма прохождения прямого сигнала подвыборочных слоёв НС для его ускорения.

В процессе обучения НС свёртки была достигнута точность распознавания 98.2% на обучающей выборке и 97.1% на тестовой выборке базы MNIST.

Также был разработан алгоритм центрирования изображений для улучшения точности распознавания.

Распознаватель системы способен распознавать изображения рукописных символов со скоростью около 85 образов в секунду.

Нейронная сеть настраивалась на распознавание рукописных арабских цифр при тестировании модели на обучающем множестве MNIST, а также на множество латинских символов, арабских цифр, знаков препинания и некоторых служебных символов для тестирования алгоритма распознавания текста.

Разработанная модель может использоваться в качестве подсистемы в более сложных системах распознавания, верификации и идентификации, таких, как идентификация отпечатков пальцев или человеческих лиц.

Описанная модификация архитектуры нейронной сети свёртки позволяет использовать мощностные парадигмы инвариантности распознающего алгоритма

к искажениям и зашумлению входного сигнала без использования дополнительных затрат на процессорное время во время обучения.

Несмотря на то, что время работы в режиме распознавания несколько увеличивается, в конкретно данной архитектуре это компенсируется, во-первых, слоями подвыборки, производящими локальное усреднение сигнала, в результате чего пространство признаков значительно уменьшается; а во-вторых, свойствами простых нейронов использовать один и тот же набор весовых коэффициентов, что открывает широкие возможности для многопоточной оптимизации при реализации алгоритма.

В процессе конструирования были реализованы алгоритмы зашумления и искажения текста, соответствующие коммерческим аналогам.

Также были разработаны и протестированы алгоритмы обнаружения текста и сегментации текстовых изображений. В процессе тестирования была установлена точность распознавания алгоритма, сопоставимая с точностью работы распознавателя – 94,3% на слабо искажённом тексте, и 82,4% на средне искажённом тексте, что значительно выше точности распознавания искажённого и зашумленного текста коммерческими аналогами.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Гальченко, Я. А. Роль антропного принципа в глобальном эволюционизме / Я. А. Гальченко, Д. В. Ермолович // Философия: материалы 51-й научная конференция аспирантов, магистрантов и студентов, Минск, 13-17 апреля 2015 г. / отв. за вып. Д. В. Ермолович – Минск: БГУИР, 2015.

2. Гальченко, Я. А. Модификация архитектуры нейронной сети свёртки с целью повышения устойчивости алгоритма распознавания к зашумлению и искажениям входного сигнала / Я. А. Гальченко, Н. А. Новосёлова // Сборник материалов VI Международной научно-практической молодёжной конференции «Научные устремления» / ООО «Лаборатория интеллекта и Центр молодёжных инновация. – Минск: «Энциклопедикс», 2015.