

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

УДК _____

Каханович
Александр Иванович

«Нейронные сети в системах распознавания текста»

Автореферат
на соискание степени магистра технических наук
по специальности: 1-40 80 02 «системный анализ, управление и
обработка информации»

Научный руководитель
Севернёв Александр Михайлович
доцент, кандидат тех. наук

Минск, 2018 г.

ВВЕДЕНИЕ

Общие сведения

Нейронные сети в последние годы набирают очень большую популярность, из-за их способности решать нетривиальные задачи, которые ранее были лишь фантастикой. В первой половине 2016 года мир услышал о множестве разработок в области нейронных сетей — свои алгоритмы демонстрировали *Google* (сеть-игрок в *AlphaGo*), *Microsoft* (ряд сервисов для идентификации изображений), стартапы *MSQRD*, *Prisma* и другие. Сама технология не нова. Первую формализованную нейронную сеть относят к 1943 году, когда два американских учёных (Уоррен Мак-Каллок и Уолтер Питтс) представили статью о логическом исчислении человеческих идей и нервной активности. Но только сейчас, ей нашли огромный спектр применений.

Нейронные сети — одно из направлений в разработке систем искусственного интеллекта. Идея заключается в том, чтобы максимально близко смоделировать работу человеческой нервной системы — а именно, её способности к обучению и исправлению ошибок. В этом состоит главная особенность любой нейронной сети — она способна самостоятельно обучаться и действовать на основании предыдущего опыта, с каждым разом делая всё меньше ошибок. Нейронные сети используются для решения сложных задач, которые требуют аналитических вычислений подобных тем, что делает человеческий мозг. К таким задачам относят ввод и обработку информации: распознавание рукописных текстов, отсканированных почтовых, платежных, финансовых и бухгалтерских документов; распознавание речевых команд, речевой ввод текста в компьютер.

Задача распознавания машинописного текста носит название оптического распознавания символов (*optical character recognition, OCR*). В настоящее время существуют высокоточные системы для распознавания машинописных и рукопечатных текстов (например, *ABBYY FineReader*). Распознавание же рукописных текстов является гораздо более сложной и на данный момент нерешённой задачей. Задача распознавания рукописного текста носит название *HWR (handwriting recognition)*.

ОБЩАЯ ХАРАКТЕРИСТИКА

Актуальность темы исследования

Так как задача распознавания рукописного текста полностью не решена, то актуальность данной темы исследования оправдана.

Системы распознавания рукописного текста могут применяться для решения многих практических задач: классификаций чеков в банках, почтовых отделениях, различных предприятиях, которые работают с рукописным текстом. Система, также может применяться для преобразования рукописного текста в машинный текст. Стоит отметить, это значительно увеличит скорость обработки данных (документов), тем самым позволит сэкономить ресурсы и денежные затраты. Даже сейчас можно заметить, какую финансовую прибыль приносят приложения с похожим функционалом. Для примера можно выделить систему, которая «читает» банковские чеки; она по эффективности в несколько раз превосходит оператора, а ошибка распознавания – в разы меньше.

Использование нейронной сети связана с тем, что она отлична подходит для решения данных задач, она обучается, это позволит распознавать текст с различным написанием и уменьшить ошибку распознавания до минимума.

Степень разработанности проблемы

Нейронная сеть – это последовательность нейронов, соединённых между собой синапсами. Структура нейронной сети пришла в мир программирования прямиком из биологии. Благодаря такой структуре, машина обретает способность анализировать и даже запоминать различную информацию. Нейронные сети также способны не только анализировать входящую информацию, но и воспроизводить её из своей памяти. Другими словами, нейросеть – это машинная интерпретация мозга человека, в котором находятся миллионы нейронов, передающих информацию в виде электрических импульсов.

Сложность задачи состоит в том, что рукописный текст содержит группы цифр, буквы, символы, разделительные знаки, соединённых в одно слово. При этом эти слова часто различаются, ведь у каждого человека присутствует свой стиль написания. Также некоторые буквы могут быть написаны неаккуратно и похожими при этом на совершенно другие буквы, что приводит к трудности

прочтения этого слова. Но самой главной проблемой является то, что почерк состоит из временной последовательности штрихов, то есть движения пера от касания к бумаге до его поднятия, что приводит к непониманию завершённости написания буквы и начала новой, поэтому нужно ограничить каждую букву и формировать символы последовательно, чтобы один символ завершался перед началом следующего. Также существует задача выделения области написанного текста из фотографий с нечётким изображением, с различной освещённостью, цифровым шумом и т. д.

Цель и задачи исследования

Целью исследования является создания алгоритма и приложения распознавания рукописных цифр с использованием нейронной сети. Для этого ставятся следующие задачи:

- а) Анализ и выбор подходящей нейронной сети;
- б) Спроектировать нейронную сеть для распознавания рукописных цифр, данная нейронная сеть должна быть обучаемая и для минимизации функции ошибки использовать градиентный метод;
- в) Программирование нейронной сети на языке *JavaScript*. Анализ и выбор библиотек нейронной сети. Для обучения использовать изображения открытой библиотеки *MNIST*;
- г) Анализ полученных результатов и перспективы улучшения работы алгоритма.

СТРУКТУРА РАБОТЫ

Введение содержит описание проблемы, сложность задачи, актуальность, мотивацию для её решения, и общие сведения о предметной области.

В главе №1 – АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ, описывается и анализируются существующие нейронные сети с указанием их достоинств и недостатков. Поясняется принцип работы различных нейронных сетей и её составляющих.

В главе №2 – ПОСТРОЕНИЕ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ ЦИФР, выбирается оптимальная нейронная сеть, а также база данных для распознавания рукописных цифр. Исходя из полученной информации рассчитывается свёрточная нейронная сеть, функция активации, функция ошибки. Анализируется обучение и коэффициенты ошибки нейронной сети, и составляется модель.

В главе №3 – ПРОГРАММИРОВАНИЕ НЕЙРОННОЙ СЕТИ НА ЯЗЫКЕ *JAVASCRIPT*, программируется нейронная сеть на языке *JavaScript* с использованием библиотеки *brain* и *MNIST*. Приводится часть программного кода и анализ полученных результатов.

В главе №4 – ПЕРСПЕКТИВЫ УЛУЧШЕНИЯ РАБОТЫ ПРИЛОЖЕНИЯ, приводятся перспективы улучшения работы приложения и несколько вариантов для дальнейшего развития алгоритма распознавания рукописного текста с использованием нейронной сети.

Заключение подводит итоги работы, включая достигнутые результаты и перспективы дальнейших исследований.

ЗАКЛЮЧЕНИЕ

Результатом выполнения данной работы является приложение и алгоритм по распознаванию рукописного текста, а именно цифр, на основе нейронной сети. В данной работе были описаны следующие этапы работы данного алгоритма, такие как: предварительная обработка входящих данных и преобразование их в двоичную форму, форму понятную для нейронной сети. Проанализированы различные виды использования нейронной сети и выбран лучший подход. По заданным требованиям была выбрана классификация нейронной сети.

Была построена свёрточная нейронная сеть, описан алгоритм обучения нейронной сети. Данная система распознавания была написана на языке *JavaScript* с использованием таких библиотек, как *brain*, *MNIST* и реализована в *web* приложении. Также произведена оценка результатов работы алгоритма. Результаты показали удовлетворительный уровень точности распознавания, который, тем не менее, может быть значительно улучшен. Были описаны также примеры улучшений, которые могут быть сделаны впоследствии.

Было установлено, что для увеличения количества правильного распознавания необходимо больше времени тренировать (обучать) нейронную сеть и увеличивать количество нейронов, находящихся в скрытом слое. Если есть возможность, то увеличить разрешение входных векторов. Самое важное, то, что в процессе обучения необходимо использовать большее количество наборов входных данных, по возможности с большим зашумлением полезной информации.

В дальнейшем данное приложение можно использовать как часть комплексной системы распознавания различных изображений содержащее рукописные цифры, для повышения качества и скорости выполнения данной работы, а также экономия денежных ресурсов.

СПИСОК СОБСТВЕННЫХ ПУБЛИКАЦИЙ

Каханович, А. И. Нейронные сети в системах распознавания текста / А. И. Каханович // Информационные технологии и управление: 53-я научная конференция аспирантов, магистрантов и студентов учреждения образования «Белорусский государственный университет информатики и радиоэлектроники» (Минск, 2 - 6 мая 2017 г.) / редкол. : Л. Ю. Шилин [и др.]. – Минск : БГУИР, 2017. – С. 94.