

РАЗРАБОТКА АЛГОРИТМА РАСПОЗНАВАНИЯ ОТВЕТОВ В БЛАНКАХ ЕГЭ С ПОМОЩЬЮ СВЁРТОЧНОЙ НЕЙРОННОЙ СЕТИ

Е.И. Имигеев

Научный руководитель: Ю.А. Болотова

Национальный исследовательский Томский политехнический университет
eii2@tpu.ru

Введение

Появление за последнее десятилетие достаточных мощностей для приемлемой скорости обучения [1] искусственных нейронных сетей (ИНС, далее - НС) позволило использовать цифровые технологии для решения тех задач, которые ранее могли выполняться только человеком, например, распознавание образов или речи, прогнозирование и т.п. В данной статье рассматриваются свёрточные НС (СНС), которые нацелены на эффективное распознавание изображений [2].

Существующие оптические способы распознавания текста (OCR-системы) в настоящее время достаточно эффективны, но не абсолютно – и точность достигает 99.998% [3], что все равно не исключает проверку результата человеком. Однако OCR-системы работают только с печатным не рукописным текстом. Для этого используются методы распознавания как при помощи метрик и шаблонов, так и с помощью НС [4]. Проблемы распознавания рукописного «печатного» и стандартного рукописного текста в настоящее время являются предметом активных исследований уже не один десяток лет [5].

Ответы в бланках ЕГЭ пишутся от руки «печатными» буквами, поэтому для определения ответов, как лучшим в настоящее время инструментом для распознавания объектов на изображениях, следует использовать СНС. Для этого необходимо реализовать алгоритм, которое будет обрабатывать отсканированные изображения бланков, распознавать на них рукописный текст с его последующей записью в необходимый текстовый формат для эффективной проверки работ.

Выбор платформы для реализации

Для реализации модели используется языки программирования C# и Python.

Фреймворков для создания программной НС существует предостаточно. Однако, особый интерес вызвал специализированный под быстрое обучение нейросетевой инструментариий CNTK, созданный компанией Microsoft [6].

Microsoft Cognitive Toolkit (CNTK) – бесплатный общедоступный набор инструментов с открытым исходным кодом для глубинного обучения. По популярности в веб-хостинг-сервисе GitHub специализированный пакет для глубинного обучения CNTK занимает третье место [7] после TensorFlow и Caffe, обогнав такие платформы, как MxNet, Theano, Torch и т.д.

В статье [7] описано сравнение двух фреймворков CNTK и TensorFlow, первый из которых преимущественен по большинству характеристик, таких как гибкость, скорость, масштабируемость и др.

Модель СНС

Используется архитектура, подобная СНС LeNet-5, с измененным количеством выходов для распознавания букв русского алфавита (см.рис.1).

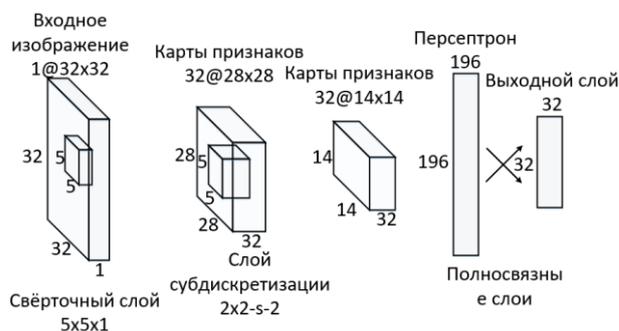


Рис.1. Структура исследуемой СНС

По результатам первых тестов распознавания цифр из базы данных MNIST с помощью данной СНС ошибка составила 2.25%, что является неплохим началом для распознавания объектов.

Сегментация бланка ответов

Бланки ответов ЕГЭ имеют строгую единую форму, на которой есть фотометки, по которой отсканированные изображения бланков необходимо выровнять, благодаря чему будет легче произвести сегментацию документа.

Выровненный бланк сегментируется по заданному шаблону, определяя на изображении места, где расположены номера вопросов и ответы на них. Ответы в бланках ЕГЭ записываются посимвольно в «клеточки», что также упрощает выделение символов с помощью шаблона. В конце сегментации получается массив изображений с символами, который отправляется в сеть для распознавания.

Проще говоря, этот процесс можно назвать парсингом (англ. *parsing* – синтаксический анализ) текста на фотографии (скане).

Распознавание ответов

Существует проблема распознавания цифр и букв одной НС – цифры распознаются хуже, чем буквы [8]. Это связано с тем, что букв больше, чем цифр, поэтому буквы имеют гораздо большее разнообразие написания (из-за почерка). Как следствие, обучающая выборка для распознавания

цифр имеет гораздо меньшее количество обучающих примеров, чем прецеденты для букв. Поэтому НС будет определять цифры, похожие на буквы, как буквы.

Для решения этой проблемы было решено создать две НС: одну для распознавания букв, вторую – для цифр. Однако появляется необходимость в использовании еще одного классификатора для определения типа ответа – является ли он числовым или буквенным.

После определения класса вопроса символ отправляется в свою НС для распознавания. На рис.2 показан процесс распознавания ответов на бланках ЕГЭ

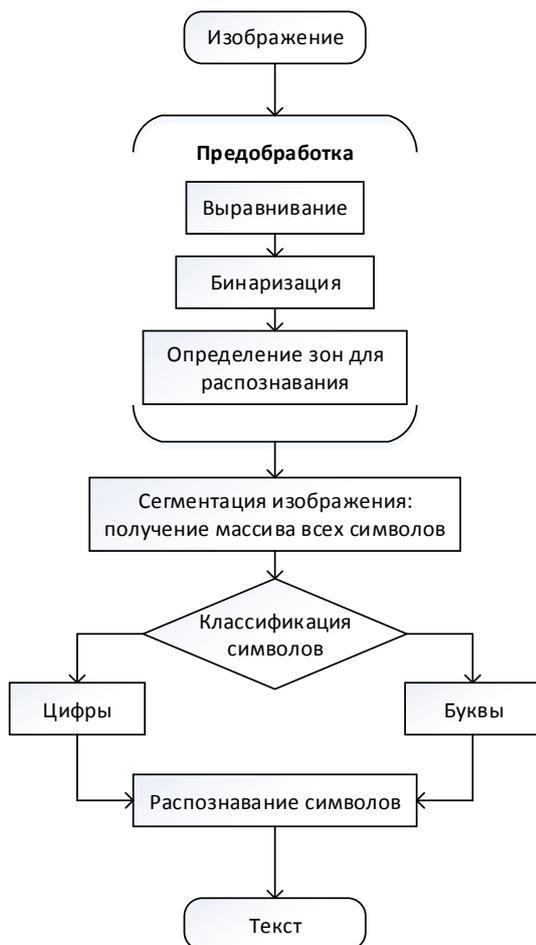


Рис.2. Схема процесса распознавания символов по изображению.

Синтаксический анализ

Из-за почерка человека чтение его рукописей может вызывать сложности даже у людей. СНС определенно будет допускать ошибки в ситуациях с непонятным почерком. Например, ученик написал в бланке ответов букву «И», однако СНС распознала ее как «Н». Корректно определить букву возможно только с контекстом (см.рис.3).



Рис.3. Пример сложного для распознавания символа

Решением этой проблемы может выступить синтаксический анализ, который позволяет проверять грамматику слов или фраз целиком. Ответы, которые содержат неясности и вызывают подозрение на наличие ошибок, отправляются оператору для уточнения.

Заключение

Большое количество машиночитаемых бланков нуждаются в обработке, а существующие системы не позволяют достаточно точно распознавать изображения, и полученные данные требуются в ручной проверке, что создает большую нагрузку на проверяющих.

В задачах распознавания применение нейронных сетей оправдано и в значительной степени экономит средства и ресурсы.

В дальнейшем планируется увеличить точность НС, ее скорость работы, а также добавить модуль грамматической проверки.

Список использованных источников

1. Бум нейросетей [Электронный ресурс] / Стартапы и бизнес. – URL: <https://vc.ru/16843-neural-networks> (доступ: 23.11.17)
2. S. Haykin. Neural networks. McMaster University, Ninth Indian Reprint, 2005
3. Report on the Meeting of Experts on Digital Preservation / U.S. Government Printing Office Washington, D.C., March 12, 2004
4. Методы распознавания текста [Электронный ресурс] // Хабрахабр. – URL: <https://habrahabr.ru/post/220077/> (доступ: 23.11.17)
5. Изменения в распознавании рукописного текста [Электронный ресурс] // Команда разработки Microsoft. – URL: <https://blogs.msdn.microsoft.com/e7ru/2009/02/18/303/>
6. CNTK – нейросетевой инструмент от Microsoft Research [Электронный ресурс] // Хабрахабр. – URL: <https://habrahabr.ru/company/microsoft/blog/275959>
7. Reasons to Switch from TensorFlow to CNTK [Электронный ресурс] // URL: <https://docs.microsoft.com/en-us/cognitive-toolkit/reasons-to-switch-from-tensorflow-to-cntk>
Разработка алгоритма и программного обеспечения сегментации и распознавания машиночитаемых бланков на основе сверточной нейронной сети [Электронный ресурс]. – URL: <http://earchive.tpu.ru/handle/11683/40514>