**МИНИСТЕРСТВО ОБРАЗОВАНИЯ, КУЛЬТУРЫ И ИССЛЕДОВАНИЙ**

**ГОСУДАСТВЕННЫЙ УНИВЕРСИТЕТ «АЛЕКУ РУССО» В БЕЛЬЦАХ**

**ФАКУЛЬТЕТ РЕАЛЬНЫХ НАУК, ЭКОНОМИКИ И ОКРУЖАЮЩЕЙ СРЕДЫ**

**КАФЕДРА МАТЕМАТИКИ И ИНФОРМАТИКИ**

**ОПТИЧЕСКОН РАСПОЗНАВАНИЕ СИМВОЛОВ ПРИ ПОМОЩИ НЕЙРОННЫХ СЕТЕЙ**

**Автор:**

Студентка группы IS11Z

**Юлия СТОЯН**

**\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Научный руководитель:**

**Олеся СКУТНИЦКИ**

магистр, асист. унив.

\_\_\_\_\_\_\_\_\_\_\_\_\_

**БЕЛЬЦЫ, 2021**

ОГЛАВЛЕНИЕ

[ВВЕДЕНИЕ 3](#_Toc71579216)

[1. Нейронные сети 5](#_Toc71579217)

[2. История 7](#_Toc71579218)

[3. Оптическое распознавание символов 11](#_Toc71579219)

[3.1. Алгоритм работы оптического распознавания текстов 14](#_Toc71579220)

[3.2. Возможности использования приложения ocr 15](#_Toc71579221)

[4. Преимущества работы с ocr- программами 17](#_Toc71579222)

[5. Заключение 19](#_Toc71579223)

[6. БИБЛИОГРАФИЯ 21](#_Toc71579224)

Введение

Одним из самых быстрых и удобных способов перевода информации из физического формата в электронный вид является сканирование документов. Результатом данного процесса будет электронный файл, представленный в виде графического изображения. Графическое изображение не позволяет производить необходимый набор действий, как при работе с текстом, что делает его менее функциональным. Основные отличия хранения текстовой информации, в отличии от графической: экономия затрат на хранении, более обширный список сценариев использование документа. OCR - это система оптического распознавания символов. В настоящее время данная система имеет большую популярность, она применяется в большом количестве программ, связанных с распознаванием текста.

Технология OCR существует уже давно и является ключом к облегчению работы многих людей, поскольку это способность преобразовывать символы, которые являются неотъемлемой частью изображения, в символы, которыми можно манипулировать, что позволяет избежать утомительной задачи расшифровка текста . Но что за всем этим стоит?

OCR сегодня используется круглосуточно и не только для оцифровки текстов, но и для таких вещей, как перевод в реальном времени текста, написанного на других языках, и мы даже можем преобразовать рукописный текст в печатный.

В 1960 году Лоуренс (Ларри) Робертс, исследователь Массачусетского технологического института, который, как ни парадоксально, позже стал одним из изобретателей того, что в конечном итоге стало Интернетом, создал систему распознавания символов и связанный с ней шрифт, предназначенный для оцифровки банковских чеков и т. Д. на. конфиденциальная информация, которую нужно было хранить на ранних компьютерах. Этот источник получил название OCR-A.

1. Нейронные сети

Нейронная сеть - математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей - сетей нервных клеток живого организма. Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы. Первой такой попыткой были нейронные сети У. Маккалока и У. Питтса. После разработки алгоритмов обучения получаемые модели стали использовать в практических целях: в задачах прогнозирования, для распознавания образов, в задачах управления и др.

ИНС представляет собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов). Такие процессоры обычно довольно просты (особенно в сравнении с процессорами, используемыми в персональных компьютерах). Каждый процессор подобной сети имеет дело только с сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим процессорам. И, тем не менее, будучи соединёнными в достаточно большую сеть с управляемым взаимодействием, такие по отдельности простые процессоры вместе способны выполнять довольно сложные задачи.

С точки зрения машинного обучения, нейронная сеть представляет собой частный случай методов распознавания образов, дискриминантного анализа, методов кластеризации и т. п.

Нейронные сети не программируются в привычном смысле этого слова, они обучаются. Возможность обучения - одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными данными и выходными, а также выполнять обобщение. Это значит, что в случае успешного обучения сеть сможет вернуть верный результат на основании данных, которые отсутствовали в обучающей выборке, а также неполных и/или «зашумленных», частично искажённых данных.

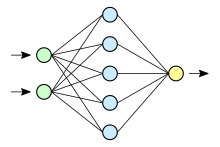


Рис. 1 Схема простой нейронной сети

1. История

Разработка OCR-систем основана на технологиях, связанных с телеграфией и созданием считывающих устройств для слепых. В 1914 году Эммануэль Гольдберг разработал устройство, считывающее символы и преобразовывающее их в стандартный телеграфный код. Одновременно Эдмунд Фурнье д'Альбе разработал «Оптофон», ручной сканер, который, при перемещении по напечатанной странице, вырабатывал тональные сигналы, соответствующие определенным буквам или символам.

В 1974 году Рэй Курцвейл создал компанию «Kurzweil Computer Products, Inc» и начал работать над развитием первой системы оптического распознавания символов, способной распознавать текст, напечатанный любым шрифтом. Курцвейл считал, что лучшее применение этой технологии - создание машины чтения для слепых, которая позволила бы слепым людям иметь компьютер, умеющий читать текст вслух. Данное устройство требовало изобретения сразу двух технологий -ПЗС (прибор с зарядовой связью планшетного сканера и синтезатора, преобразующего текст в речь.

Первой коммерчески успешной программой, распознающей кириллицу, стала программа «AutoR» российской компании «ОКРУС». Алгоритм «AutoR» был компактный, быстрый и шрифтонезависимый. Этот алгоритм разработали и испытали ещё в конце 60-х два молодых биофизика, выпускники МФТИ - Г. М. Зенкин и А. П. Петров. В настоящее время алгоритм Зенкина-Петрова применяется в нескольких прикладных системах, решающих задачу распознавания графических символов.

В 1993 году вышла технология распознавания текстов российской компании ABBYY. На её основе создан ряд корпоративных решений и программ для массовых пользователей. Технологии распознавания текстов ABBYY OCR лицензируют международные ИТ-компании, такие как Fujitsu, Panasonic, Xerox, Samsung, EMC и другие.

В 2000-х годах производительность и компактность OCR-системы позволила представить на рынок онлайн-сервисы по переводу текста с одного языка на другой. Со временем такие программы получили возможность обрабатывать изображения как печатного, так и рукописного текста.

С развитием технологий производства мобильных устройств и упрощения процесса разработки мобильных приложений, OCR-системы стали неотъемлемой частью разнообразных программ: от развлекательных до обучающих, от мобильных помощников до систем управления.

1. Оптическое распознавание символов
   1. Алгоритм оптического распознавания символов

Алгоритм работы распознания текста всегда строится одинаково.В систему загружается отсканированный файл, представленный в виде растрового изображения страницы документа. Качества изображения играет важную роль в распознании текста: чем выше качество, тем выше точность. Поэтому первым этапом будет являться обработка поступившего изображения: снижение шума, повышения контраста, повышение резкости, бинаризация изображения, выравнивание угла наклона [2]. Обработанный файл передается в модуль сегментации, задачей которого является выявление структурных единиц текста — страниц, строк, слов и символов. После сегментации полученные данные собираются в обратном порядке в готовый файл.



Рис. 2 Порядок сегментации

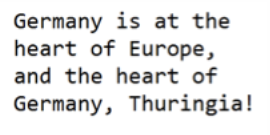
Для начала документ делится на страницы, далее определяются текстовые блоки. Для выявления слов из текстового блока производится определение угла наклона текста, для уменьшения будущих погрешностей, поиск вертикальных просветов в тексте, показывающих границы слова [3]. Для разбивки слова на символы проводится аналогичный процесс, только с меньшими просветами. Данные операции будут более точными, если текст будет черного цвета на белом фоне, если оригинальный текст иного цвета, то применяется бинаризация изображения.  
  


Рис. 3 Пример входного текстового блока

На выходе из модуля сегментации будут получены данные, в состав которых входят структуры и местоположение текстовых блоков на странице, строки в этих блоках и их сегментация на слова и символы. Данные могут содержать не только информацию об обычном текстом блоке, а также о колонках, таблицах и т. д.

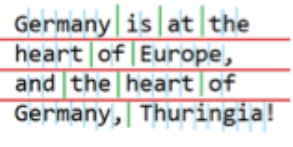


Рис. 4 Пример обработанного текстового блока

Определенные фрагменты слов и символов отправляются в модуль классификатора, результатом работы которого будет являться информация о принадлежности символа к определенной букве или символу. Нейронная сеть для каждого входящего символа, используя его пиксельное изображение, определяет признаки принадлежности буквы к нечеткому множеству. После определения признаков у символа начинается процесс составление из символов слов. Для этого нейронная сеть сравнивает возможность написания отдельных букв, частоту сочетаний букв в языке, производится проверка по модели слова и словарю [3]. Модель слова - модель, разделяющая слова на определенные типы, такие как сокращения, аббревиатуры, обычные слова, имена собственные, числа и т. д. С этого момента проверяется насколько хорошо подходит к данной модели полученное слово.

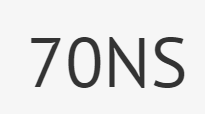


Рис. 5 Пример модели слова

Таблица 1 Пример списка моделей слов

|  |  |
| --- | --- |
| **Вариант распознавания слова** | **Модель** |
| TONS | Заглавные буквы |
| tons | Прописные буквы |
| Tons | Первая заглавная буква в слове |
| Tens | Первая заглавная буква в слове |
| Tans | Первая заглавная буква в слове |
| 70NS | Сокращение |
| 70ns | Число с подстрочным знаком |

* 1. Возможности использования приложения ocr

Наиболее совершенные системы распознавания символов, такие как ABBYY FineReader OCR, делают акцент на использовании механизмов, созданных природой. В основе этих механизмов лежат три фундаментальных принципа: целостность, целенаправленность и адаптивность (принципы IPA).

Изображение, согласно принципу целостности, будет интерпретировано как некий объект, только если на нем присутствуют все структурные части этого объекта и эти части находятся в соответствующих отношениях. Иначе говоря, ABBYY FineReader не пытается принимать решение, перебирая тысячи эталонов в поисках наиболее подходящего. Вместо этого выдвигается ряд гипотез относительно того, на что похоже обнаруженное изображение. Затем каждая гипотеза целенаправленно проверяется. И, допуская, что найденный объект может быть буквой А, FineReader будет искать именно те особенности, которые должны быть у изображения этой буквы. Как и следует поступать, исходя из принципа целенаправленности. Принцип адаптивности означает, что программа должна быть способна к самообучению, поэтому проверять, верна ли выдвинутая гипотеза, система будет, опираясь на накопленные ранее сведения о возможных начертаниях символа в данном конкретном документе.

1. Преимущества работы с ocr- программами

Преобразованием графического изображения в текст занимаются специальные программы распознавания текста (Optical Character Recognition - OCR). Современная OCR должна уметь многое: распознавать тексты, набранные не только определенными шрифтами, но и самыми экзотическими, вплоть до рукописных. Уметь корректно работать с текстами, содержащими слова на нескольких языках, корректно распознавать таблицы. И самое главное корректно распознавать не только четко набранные тексты, но и такие, качество которых, мягко говоря, далеко от идеала. Например, текст с пожелтевшей газетной вырезки или третьей машинописной копии. Само собой, распознать текст это еще полдела. Не менее важно обеспечить возможность сохранения результата в файле популярного текстового (или табличного) формата скажем, формата Microsoft Word.

1. Заключение
2. БИБЛИОГРАФИЯ

1. Бройдо, В. Л. Вычислительные системы, сети и телекоммуникации / В. Л. Бройдо. — СПб.: Питер, 2004. – 703 с.

2. Визильтер, Ю. В. Обработка и анализ цифровых изображений с примерами на LabVIEW IMAQ Vision / Ю. В. Визильтер, С. Ю. Желтов, В. А. Князь и др. – М.: ДМКПресс, 2009. – 465 с.

3. Волкова, М. А. Методы обработки и распознавания изображений / М. А. Волкова, В. Р. Луцив. – СПб: Университет ИТМО, 2016. – 40 с.

4. Липкина, А. Распознавание текста по структуре скелета букв / А. Липкина. – М.: МГУ им. Ломоносова, 2018. – 31 с.

5. Суясов, Д. И. Разработка алгоритмов распознавания текста на основе клеточных автоматов / Д. И. Суясов. – СПб.: ИТМО, 2007. – 88

6. Ломанов, Д. К. Алгоритм распознавания текстовой информации на изображении с помощью ЭВМ / Д. К. Ломанов.

7. статья «Что такое OCR» <https://pdf.abbyy.com/ru/learning-center/what-is-ocr/>

8. Беркинблит М. Б.  [Нейронные сети](http://www.katenke.net/static/berkinblit/neironnye_seti.html). — М.: МИРОС и ВЗМШ РАО, 1993. — 96 с. — [ISBN 5-7084-0026-9](https://ru.wikipedia.org/wiki/%D0%A1%D0%BB%D1%83%D0%B6%D0%B5%D0%B1%D0%BD%D0%B0%D1%8F:%D0%98%D1%81%D1%82%D0%BE%D1%87%D0%BD%D0%B8%D0%BA%D0%B8_%D0%BA%D0%BD%D0%B8%D0%B3/5708400269). [Архивная копия](http://web.archive.org/web/20110512120321/http:/www.katenke.net/static/berkinblit/neironnye_seti.html) от 12 мая 2011 на [Wayback Machine](https://ru.wikipedia.org/wiki/%D0%90%D1%80%D1%85%D0%B8%D0%B2_%D0%98%D0%BD%D1%82%D0%B5%D1%80%D0%BD%D0%B5%D1%82%D0%B0#%D0%9F%D1%80%D0%BE%D0%B5%D0%BA%D1%82%D1%8B)

9. Горбань А.Н. Обучение нейронных сетей. — М.: СССР-США СП «Параграф», 1990. — 160 с.