

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Московский институт электроники и математики им. А. Н. Тихонова

Чанке Иван Сергеевич, группа БИВ192

Разработка нейронной сети для распознавания изображений.

Краткий анализ способов машинной обработки изображений. Обзор
используемой литературы.

Направление 09.03.01 «Информатика и вычислительная техника»

Студент:

Чанке И.С

Руководитель:

Романова И.И.

Москва 2019

Оглавление

Анализ существующих способов машинной обработки изображений	3
Обзор литературы, используемой в ходе выполнения курсовой работы	6

Анализ существующих способов машинной обработки изображений.

Вступление.

Сегодня в основе практически каждой автоматизированной системы обработки изображений лежат нейронные сети. Это неудивительно, ведь обычным алгоритмам не по силам задачи такого класса, а способность нейросетей обучаться на аналогичных реальным данным примерах делает их отличным инструментом для решения проблем, которые не поддаются строгой формализации. Несмотря на то, что период бурного развития технологий машинного обучения и их применения в промышленности начался относительно недавно, учёные стремились автоматизировать обработку изображений ещё с тех пор, как появилась возможность преобразовывать свет в электрические сигналы – задолго до появления компьютеров в современном виде. И эта задача была успешно решена. И хотя идеи, лежавшие в основе обработки изображений до эпохи нейросетей, сегодня кажутся не стоящими внимания, а устаревшие аналоговые и ранние цифровые решения были вытеснены машинным обучением десятки лет назад, необходимо понимать, что компьютерное зрение – не единственный способ автоматической обработки изображений. Ниже будут кратко описаны альтернативные способы решить эту задачу, их достоинства и недостатки.

Аналоговые решения. Использование фотоэлементов.

Оптофон Эдмунда Фурдьё Д'Альба (1913).

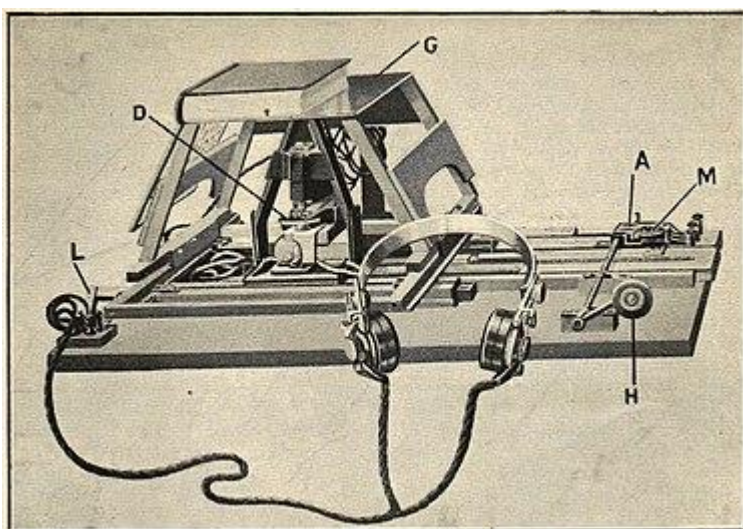


Рисунок 1. Устройство оптофона.

Изображённый на рисунке 1 оптофон, разработанный профессором Бирмингемского университета Эдмундом Фурдье Д'Альба, стал одним из первых механических приборов, способных обрабатывать представленную в графическом виде информацию. Прибор преобразовывал печатные строки в звуки, ставя в соответствие каждой букве определённую частоту. Устройством вывода служили наушники. Предполагалось, что оптофон даст слепым людям возможность читать. В основе работы устройства лежали селеновые фотоэлементы, преобразовывавшие свет в электрический сигнал. Главным недостатком прибора была скорость обработки информации - приблизительно одно слово в минуту. Устройство не получило массового распространения, однако именно на впервые применённых в нём методах преобразования света в электричество основывались дальнейшие разработки в области машинной обработки изображений.

Читающая машина Густава Таушека (1929).

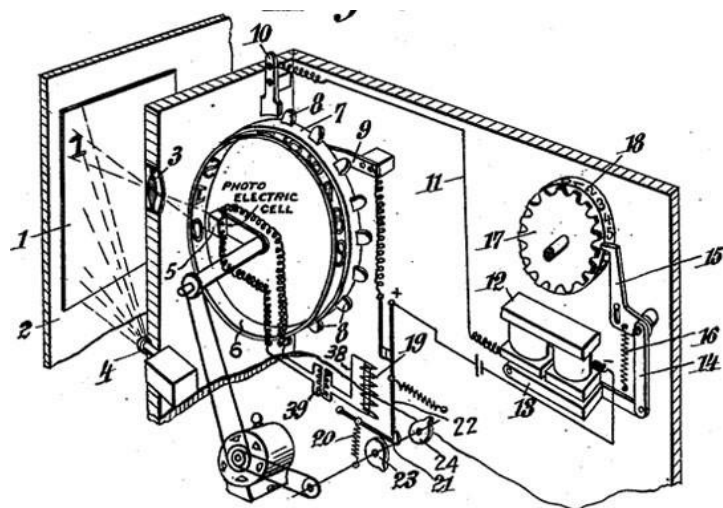


Рисунок 2.
Патентная схема читающей машины.

Читающая машина, изобретённая австрийским инженером-самоучкой Густавом Таушеком, стала первым устройством, непосредственно распознающим печатный текст. В основе её работы, как и в оптофоне, лежало использование фотоэлементов. Входные данные представляли собой текст, напечатанный на листе бумаги, результатом работы машины являлся этот же текст, перепечатанный на другой носитель. Принцип, по которому функционировало данное устройство, заключался в следующем: фотоэлемент скрывался за барабаном, имевшим отверстия в форме распознаваемых литер. Барабан вращался до тех пор, пока отверстие не совпадало с распознаваемой буквой по форме, в результате чего возбуждался фотоэлемент, посылавший сигнал устройству, ставящему печатающий барабан в соответствующее распознанной букве положение.

Очевидно, вышеперечисленные аналоговые решения обладали значительными недостатками: медленная работа, зависимость от шрифта, относительно низкая надёжность, невозможность гибкого использования результата работы (Так, печать распознанных литер была неотъемлемой частью механизма работы машины Таушека, использовать информацию о считанной букве иначе было нельзя). Тем не менее, такие решения позволяли автоматически решать задачу, прежде бывшую под силу только человеку – задачу классификации. Кроме того, описанные изобретения открывали новые возможности для дальнейших исследований в области машинной обработки изображений.

Ранние цифровые решения.

“Gismo” Дэвида Шепарда (1951).

Аппарат “Gismo” американского инженера Дэвида Шепарда стал, возможно, первым распознающим текст устройством, получившим широкое распространение и нашедшим применение в промышленности. Механизм распознавания символов наследовал некоторые идеи Таушека, однако был в целом значительно более сложным. Главным достоинством машины Шепарда стала возможность преобразовывать данные о распознанных литерках в сигналы, пригодные для обработки компьютерами того времени. Проблема гибкого использования выходных данных была решена. Компания Шепарда продавала системы распознавания текста крупным корпорациям, сыгравшим впоследствии огромную роль в развитии IT-индустрии в целом, например AT&T, через некоторое время вошла в IBM, что только способствовало развитию идей Шепарда и систем распознавания изображений. К середине 1970-х годов такие системы уже решали довольно сложные задачи, такие как сортировка писем. Примерно в то же время к задаче классификации впервые были применены технологии машинного обучения.

Современные решения.

Как уже было сказано ранее, нейросетевые технологии быстро вытеснили подобные “Gismo” системы из области машинной обработки изображений. Современные системы, основанные на нейронных сетях, нашли массу применений и сейчас используются практически в каждой отрасли для решения задач классификации. Например, инженеры компании «Яндекс» разработали систему для оценки уровня усталости водителей сервиса «Яндекс.Такси». Входные данные – изображение лица водителя – обрабатываются в реальном времени, ответ системы используется для мониторинга уровня внимания водителей и, следовательно, безопасности пассажиров. И это лишь один из множества примеров современного использования машинной обработки изображений.

Обзор литературы, используемой в ходе выполнения курсовой работы.

1. Rashid T. Make Your Own Neural Network. 2016 / Рашид. Т. Создаём нейронную сеть. изд-во «Диалектика». 2019

Книга Т. Рашида «Создаём нейронную сеть» является введением в теорию нейронных сетей, не требующим для понимания абсолютно никаких начальных знаний в предметной области. Я познакомился с этой книгой достаточно давно, когда во мне впервые возник интерес к машинному обучению. Именно работа Рашида помогла мне усвоить основные идеи обучения с учителем, а также получить начальные представления о структуре и механизме работы простых нейронных сетей – перцептронов.

Книга поделена на три большие главы: «Как работают нейронные сети», «Создаём нейронную сеть на Python», «Несколько интересных проектов». Также представлены два приложения: «Краткое введение в дифференциальное исчисление» и «Нейронная сеть на Raspberry Pi». В первой главе автор знакомит читателя с механизмом обучения с учителем: постепенным подбором необходимой для решения поставленной задачи функции при условии наличия набора связанных входных/выходных данных методом постепенных поправок к искомой функции. Автор приводит в качестве примера такой обучающейся системы линейный классификатор. Далее в книге кратко обобщается структура нейронной сети и механизмы её обучения. В главе «Несколько интересных проектов» автор предлагает поэкспериментировать с готовой нейросетью, распознающей изображения. Я не ознакомился с данным разделом, так как в нём идёт речь об использовании готовой нейронной сети.

Весьма полезным я нашёл приложение «Нейронная сеть на Raspberry Pi». В нём автор показывает, как запустить готовую нейронную сеть на одноплатном компьютере. Этот материал пригодится мне для настройки сети на Raspberry Pi.

Главные достоинства книги – краткость, упрощение сложных концептов с целью сделать их доступными и понятными широкой аудитории, использование скорее публицистического, чем научного стиля – перетекают в недостатки по мере углубления читателем своих знаний в предметной области. Автор умышленно избегает детального описания математического аппарата, на котором базируется теория нейронных сетей, ограничивается лишь поверхностными рассуждениями, но пытается поднимать такие сложные темы, как минимизация функции ошибки методом градиентного спуска.

В результате читатель, получивший базовое представление о нейросетях, чувствует необходимость обратиться к строгой научной литературе, затрагивающей ту же тему. Далее речь пойдёт об одной из таких книг.

2. Haykin S. Neural Networks: A Comprehensive Foundation. Prentice Hall, 2006 / Хайкин. С. Нейронные сети. Полный курс. И.Д. «Вильямс», 2006

Работа Саймона Хайкина «Нейронные сети. Полный курс» представляет собой объёмный учебник по теории нейронных сетей. Автор затрагивает практически все концепты, применяемые в машинном обучении, подробно останавливаясь на каждом, начиная от простых и постепенно переходя к более сложным. При этом большое внимание уделяется математической составляющей теории нейросетей – автор подробно описывает математическую модель искусственного нейрона, затем модели одно- и многослойных персептронов, метод обратного распространения ошибки и многое другое.

Книга состоит из 16 разделов, в свою очередь поделенных на множество глав. На данный момент я планирую ознакомиться с первыми четырьмя разделами: «Введение», «Процессы обучения», «Однослойный персептрон», «Многослойный персептрон». Изучения покрываемых в этих разделах тем будет достаточно для самостоятельной реализации модели персептрона. Далее я продолжу проработку материала книги с целью углубления своих знаний в этой области.