

Распознавание рукописных образов цифр на основе сверточных нейронных сетей

Сиражетдинов Руслан ФТ-401

16 ноября 2021 г.

Содержание

1	Введение	3
2	Архитектура проекта	4
3	Модель классификатора	5
4	Обучение модели	7
5	Создание графического интерфейса	8
6	Заключение	9

1 Введение

Распознавание изображений компьютером является задачей компьютерного зрения и является актуальной в наше время цифровизации и автоматизации различных процессов. Одной из таких задач является распознавание рукописных записей, для её решения можно применять различные алгоритмы, в том числе и нейронные сети, которые широко распространены, применяются и очень хорошо справляются с обработкой и классификацией изображений.

Для решения задач этой области применяются различные технологии, одним из удобных инструментов для реализации поставленных задач является **MatLab** - пакет прикладных программ для технических вычислений и создания графических интерфейсов.

MatLab для решения задачи классификации был выбран ввиду того, что по приложению существует большое количество документации и есть возможность установки дополнительных пакетов специфичных для этой области.

2 Архитектура проекта

Проект состоит из трёх частей

- Реестр данных для обучения - набор изображений с цифрами
- Сверточная нейронная сеть - модель классификатора
- Графический интерфейс - для ввода и вывода данных пользователя

3 Модель классификатора

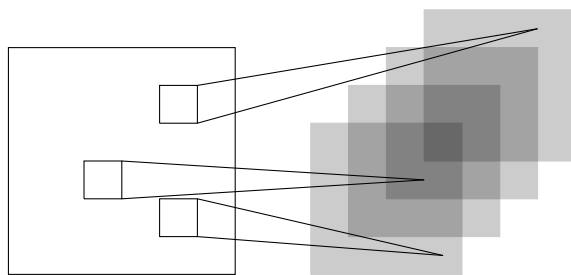
Классификатор представляет собой сверточную нейронную сеть. Работа свёрточной нейронной сети обычно интерпретируется как переход от конкретных особенностей изображения к более абстрактным деталям, и далее к ещё более абстрактным деталям вплоть до выделения понятий высокого уровня. Поэтому такой подход очень удобно использовать для распознавания образов.

Рассмотрим слои нейронной сети

- Входной слой - принимающий матрицу изображения 28 на 28 пикселей
- Сверточный слой - 20 сверток с фильтрующим размером 5x5

Входное изображение

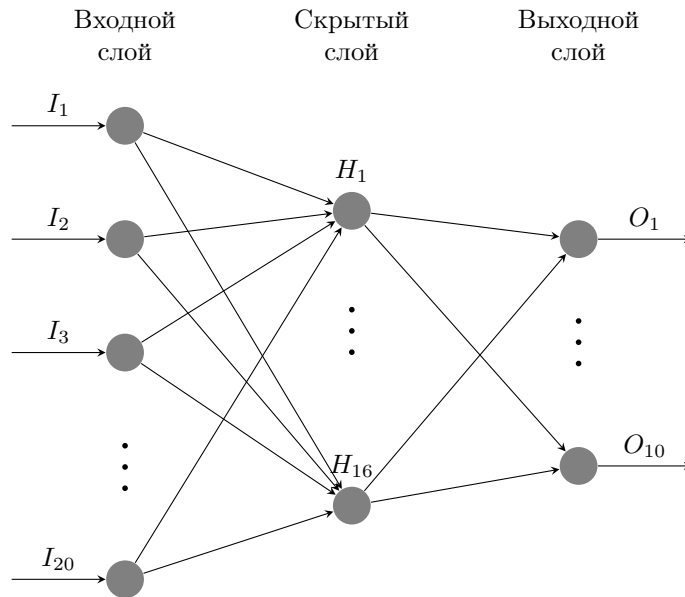
выходные свертки



- Слой пакетной нормализации - ускоряет вычисления с помощью обработки пачками.
- ReLU Слой - функция активации, выдает взвешенную сумму на нейроне.

$$Relu(x) = \max(0, x) \quad (1)$$

- Полносвязный слой - сеть из нейронов для учета всех признаков



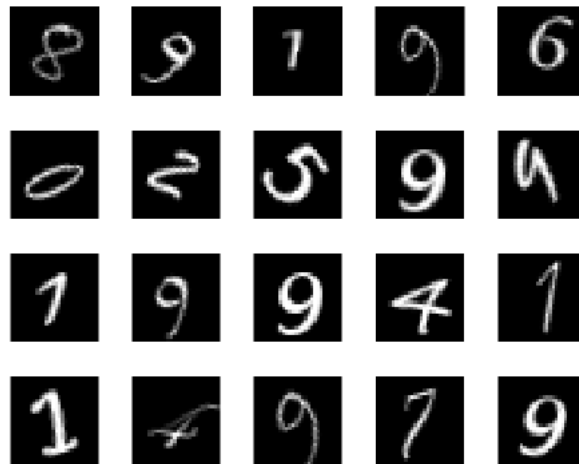
- Слой Softmax - на выходе вектор вероятностей принадлежности классу

$$\text{softmax}(x)_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (2)$$

- Классификационный слой - по вектору из вероятностей возвращает цифру

4 Обучение модели

Для обучения нейронной сети были использованы размеченные (пронумерованные) данные с изображениями цифр

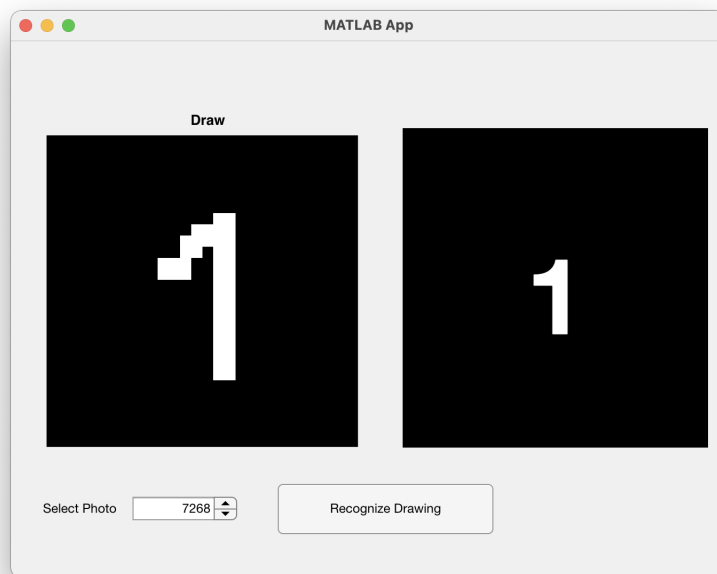


Обучение происходило со следующими гипер параметрами - 4 эпохи по 750 тренировочных подходов Для проверки качества модель была проверена на валидационной выборке

Метрика ассигасу показала 90%

5 Создание графического интерфейса

Графический интерфейс представляет собой панель для ввода произвольной графики слева и вывода классификации справа.



при запуске приложения в первый раз запустится обучение модели, что займет определенное количество времени, по завершению обучения, модель сохранится на диск и повторный запуск уже не потребует обучения.

6 Заключение

Мною было поэтапно реализовано приложение по распознаванию рукописных образов цифр. Для решения задачи я изучил различные сферы наук, такие как машинное обучение, анализ данных и разработка пользовательских интерфейсов.

Задача была декомпозирована на отдельные сущности: модель, реестр данных и интерфейс ввода-вывода. Такой подход позволит переиспользовать сущности в будущем в других проектах.

Список литературы

- [1] Ian Goodfellow, Yoshua Bengio, Aaron Courville Deep Learning // Adaptive Computation and Machine Learning series — P. 179–196.
- [2] Алан купер Интерфейс. Основы проектирования взаимодействия. 4-е изд. — Издательский дом «Питер» — 2016. — 449 p.