Министерство образования и науки Российской Федерации

Федеральное государственное автономное образовательное учреждение

высшего профессионального образования

«Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского»

Институт Информационных Технологий, Математики и Механики

Направление: Фундаментальная информатика и информационные технологии

Магистерская программа: Когнитивные системы

**ОТЧЕТ**

по лабораторной работе №3

Тема:

**«Разработка сверточной нейронной сети»**

Выполнили:

Краснов Александр (гр. 381806м4)

Разин Вячеслав (гр. 381806м4)

Ахмедов Саркар (гр. 381806м4)

Проверила:

доцент, к.т.н. Кустикова В.Д.

Нижний Новгород

2019

Оглавление

[Введение 3](#_Toc32832448)

[Постановка задачи 4](#_Toc32832449)

[Описание набора данных 5](#_Toc32832450)

[Метрика качества решения задачи 6](#_Toc32832451)

[Подготовка данных для нейронной сети 7](#_Toc32832452)

[Тестовые конфигурации сетей 8](#_Toc32832453)

[Результаты экспериментов 12](#_Toc32832454)

[Анализ результатов 12](#_Toc32832455)

# Введение

Пусть - множество изображений, – множество номеров (или наименований) классов. Существует неизвестная *целевая зависимость* – отображение , значения которой известны только на объектах конечной обучающей выборки . Требуется построить алгоритм , способный классифицировать произвольный объект .

В рамках данной лабораторной работы ставится [задача классификации немецких дорожных](https://www.kaggle.com/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign) знаков (The German Traffic Sign Benchmark).

# Постановка задачи

* **Цель работы**

Цель настоящей работы состоит в том, чтобы получить базовые навыки работы с одной из библиотек глубокого обучения (Caffe, Torch, TensorFlow, MXNet или какая-либо другая библиотека на выбор студента) на примере полностью связанных нейронных сетей.

* **Задачи**

Выполнение практической работы предполагает решение следующих задач:

1. Выбор библиотеки для выполнения практических работ курса.

2. Установка выбранной библиотеки на кластере (параметры аутентификации и инструкция по работе с кластером выложена в отдельной задаче в системе redmine).

3. Проверка корректности установки библиотеки. Разработка и запуск тестового примера сети, соответствующей логистической регрессии, для решения задачи классификации рукописных цифр набора данных MNIST (пример разобран в лекционных материалах).

4. Выбор практической задачи компьютерного зрения для выполнения практических работ.

5. Разработка программ/скриптов для подготовки тренировочных и тестовых данных в формате, который обрабатывается выбранной библиотекой.

6. Разработка нескольких архитектур полностью связанных нейронных сетей (варьируются количество слоев и виды функций активации на каждом слое) в формате, который принимается выбранной библиотекой.

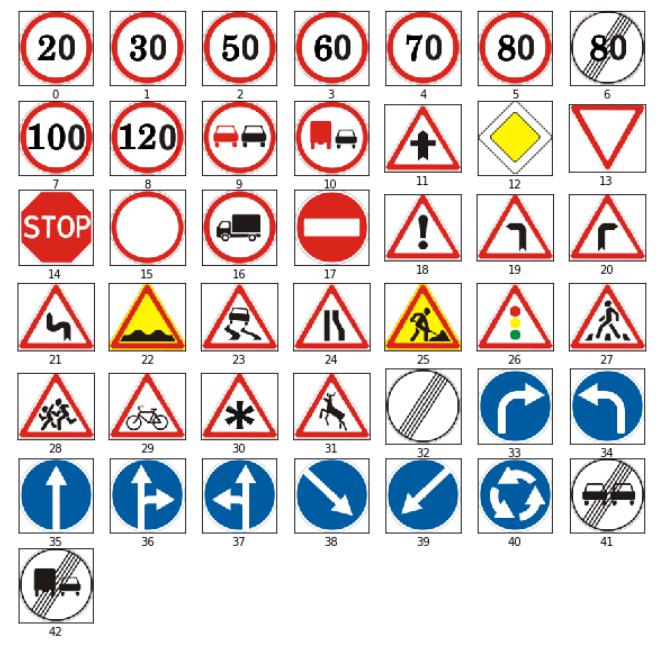
7. Обучение разработанных глубоких моделей.

8. Тестирование обученных глубоких моделей.

9. Публикация разработанных программ/скриптов в репозитории на GitHub.

10. Подготовка отчета, содержащего минимальный объем информации по каждому этапу выполнения работы.

# Описание набора данных

Для лабораторных работ был выбран один из представленных на Kaggle набор данных <https://www.kaggle.com/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign> (The German Traffic Sign Benchmark).

*Рис.1. Представители всех классов набора данных The German Traffic Sign Benchmark*

Данный набор содержит в общей сложности более 51840 изображений разных разрешений, 39209 из которых используются для тренировки нейронной сети, а 12631 - для тестирования. На каждом изображении содержится один дорожный знак. Набор данных содержит 43 класса изображений.

Для реализации различных архитектур нейронных сетей была выбрана открытая библиотека глубокого обучения Keras, написанная на языке программирования Python.

# Метрика качества решения задачи

Для оценки качества задачи классификации выбрана метрика "" (точность), показывающая отношение числа правильных предсказаний к числу предсказаний класса для всех изображений:  
где – количество изображений, – значения, выданные сетью при классификации изображений, – истинные значения классов элементов одежды с изображения.

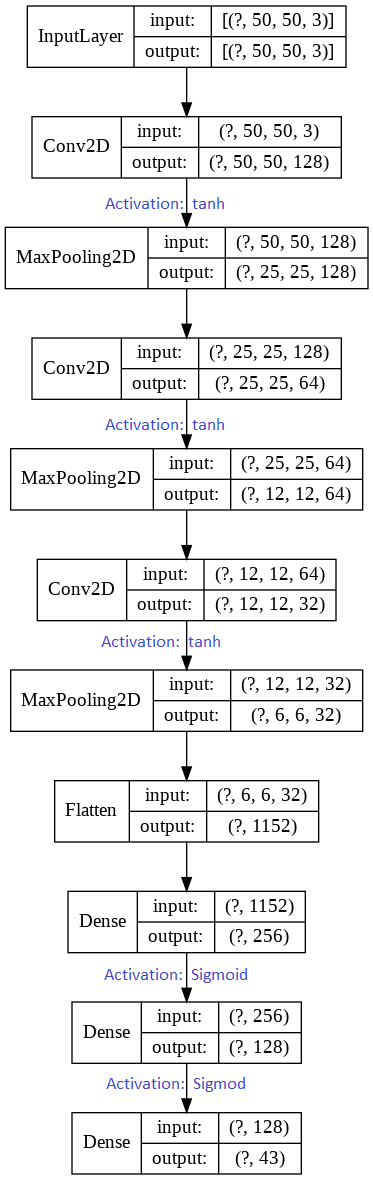
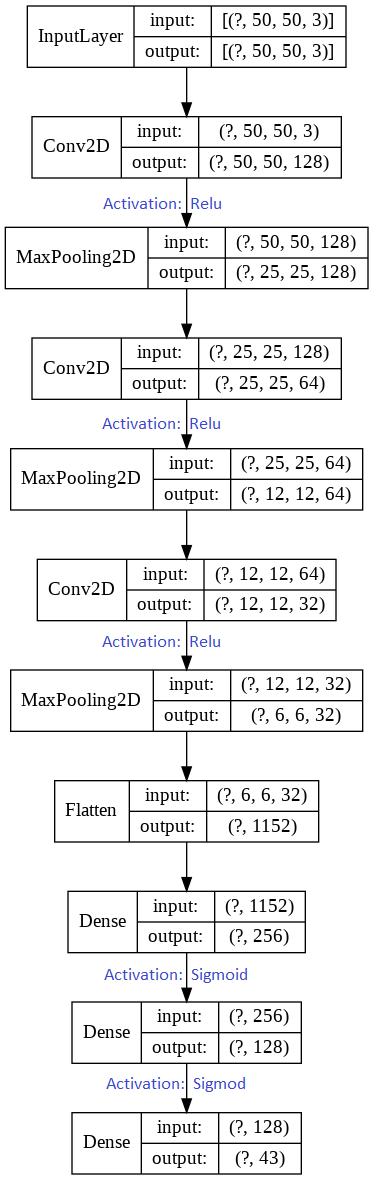
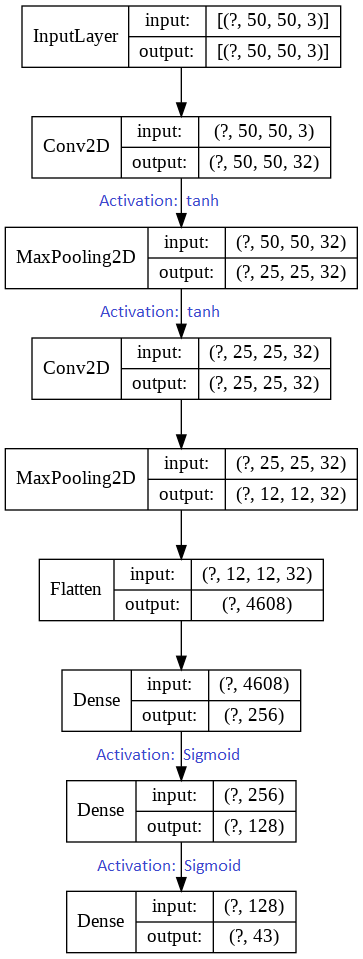
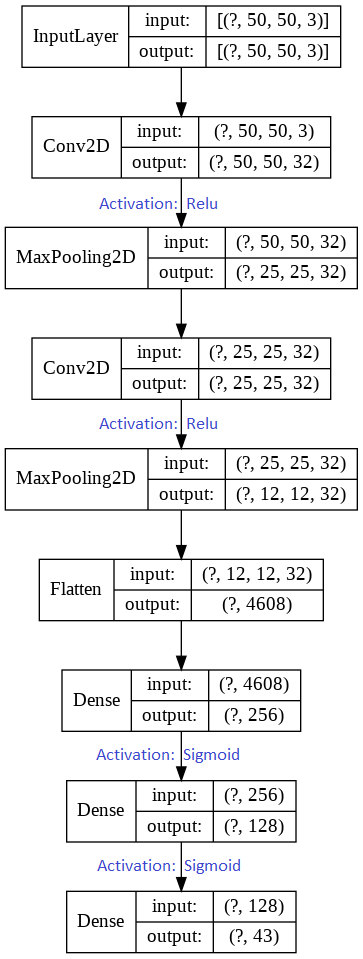
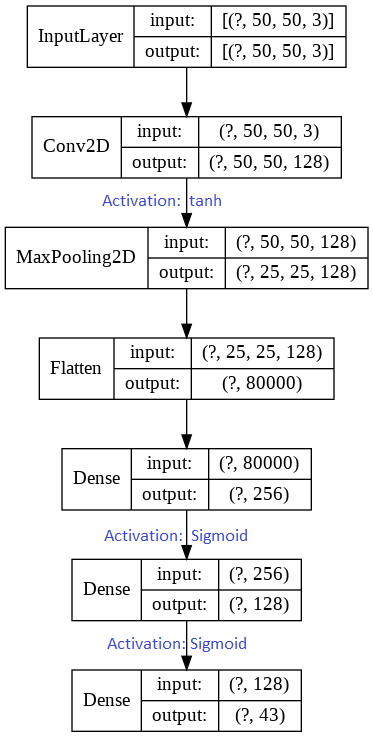
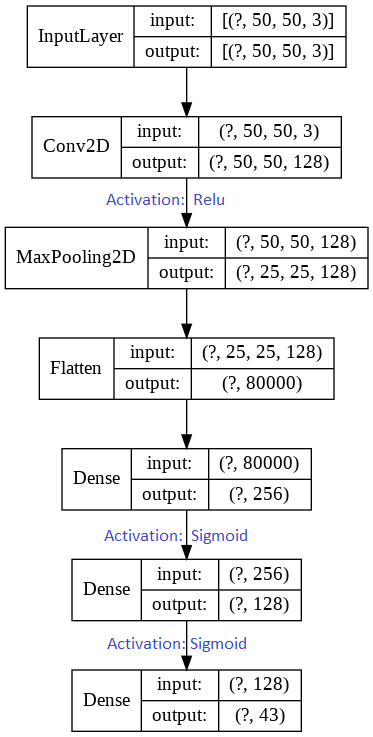
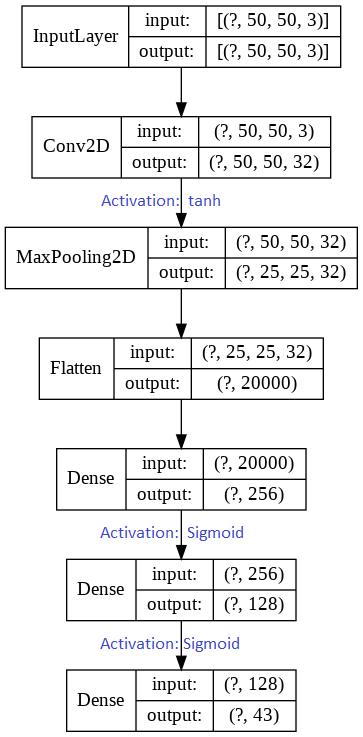
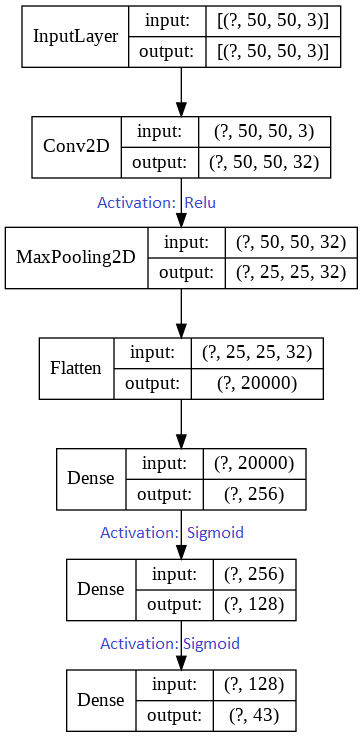
# Подготовка данных для нейронной сети

В силу большого количества входных изображений было решено сжать каждое из них до размера 50 x 50 пикселей. Сжатие происходит с сохранением соотношения сторон, уменьшенное изображение находится по центру, а недостающие пиксели имеют черный цвет. После этого преобразуем изображение (трёхмерный массив числе от 0 до 255 включительно) в один вектор и приводим каждый элемент вектора к значению с отрезка [0, 1].

Изначально каждое изображение обучающей выборки находится в пронумерованной папке, номер которой соответствует конкретному классу изображений от 0 до 42. Используя метод one-hot encoding, представляем каждый класс как вектор состоящий из 43 чисел. В итоге мы имеем два двумерных массива: массив изображений, готовых к классификации, и массив меток, отвечающий за то, какому классу данные изображения принадлежат.

# Тестовые конфигурации сетей

В ходе лабораторной работы были изучены восемь различных конфигураций сетей. Варьировались количество скрытых слоев и функции активации. Схемы конфигураций представлены ниже:



# Результаты экспериментов

Все вычислительные эксперименты выполнялись с помощью Google Colaboratory режиме GPU.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Модель сети | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Количество скрытых нейронов | 256  128  43 | 256  128  43 | 256  128  43 | 256  128  43 | 256  128  43 | 256  128  43 | 256  128  43 | 256  128  43 |
| Количество фильтров | 32 | 32 | 128 | 128 | 32  32 | 32  32 | 128  64  32 | 128  64  32 |
| Количество свёрточных слоёв | 1 | 1 | 1 | 1 | 2 | 2 | 3 | 3 |
| Функция активации на свёрточных слоях | relu | tanh | relu | tanh | relu | tanh | relu | tanh |
| Батч | 256 | 256 | 256 | 256 | 256 | 256 | 256 | 256 |
| Количество эпох | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| Скорость обучения | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| Оптимизатор | adam | adam | adam | adam | adam | adam | adam | adam |
| Время обучения, сек | 31 | 31 | 86 | 87 | 38 | 37 | 100 | 100 |
| Точность (Accuracy) на тренировочном наборе, % | 99.96 | 99.93 | 99.99 | 99.97 | 99.97 | 99.98 | 99.98 | 100 |
| Ошибка на тренировочном наборе | 0.0123 | 0.0146 | 0.0057 | 0.0083 | 0.0075 | 0.0063 | 0.0076 | 0.0065 |
| Точность (Accuracy) на тестовом наборе, % | 0.9126 | 90.58 | 92.17 | 90.83 | 92.43 | 93.6 | 93.02 | **93.81** |
| Ошибка на тестовом наборе | 0.3625 | 0.3973 | 0.325 | 0.3733 | 0.3034 | 0.26 | 0.28 | 0.26 |

# Анализ результатов

Свёрточные сети показали более, чем на 5 процентов лучший результат по сравнению с полносвязными сетями. Последняя модель с 3 сверточными слоями ожидаемо показала лучший результат. Эксперименты показали, что сетях с большим количеством свёрточных слоёв функция активация tanh эффективнее, чем relu.