

Нижегородский государственный университет им. Н. И. Лобачевского  
Институт информационных технологий, математики и механики

Направление подготовки Прикладная математика и информатика  
Магистерская программа Вычислительные методы и суперкомпьютерные  
технологии

Образовательный курс «Методы глубокого обучения для решения задач  
компьютерного зрения»

## **Отчёт**

по лабораторной работе № 3

**«Разработка свёрточной нейронной сети с использованием  
библиотеки MXNet»**

задача

«Классификация персонажей Симпсонов»

***Выполнили:***

студенты гр. 381603м4

Вершинина О.

Розанов А.

Рой В.

Нижний Новгород  
2017

## Оглавление

Постановка задачи .....	3
Формат данных для предоставления нейронной сети .....	4
Тестовые конфигурации нейронных сетей .....	5
Результаты экспериментов .....	7

## **Постановка задачи**

Необходимо построить архитектуру свёрточной нейронной сети для задачи классификации персонажей из мультфильма «Симпсоны».

В ходе работы необходимо решить следующие задачи:

1. Разработать несколько архитектур свёрточных нейронных сетей, варьируя количество слоев и виды функций активации на каждом слое.
2. Обучить построенные глубокие модели.
3. Протестировать обученные нейронные сети.

## Формат данных для предоставления нейронной сети

Исходные данные представляют собой набор jpg изображений, различного разрешения. Несколько примеров представлены на рис. 1.

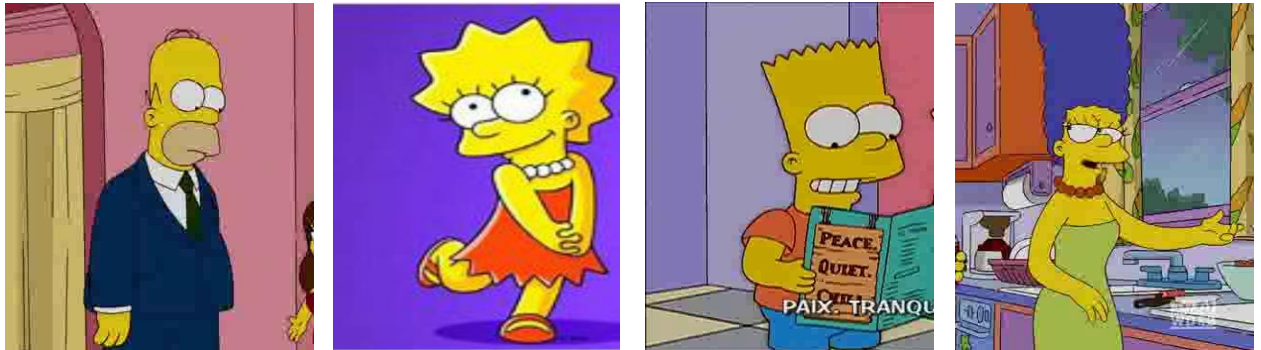


Рис.1. Персонажи из Симпсонов. Слева-направо: Гомер Симпсон, Лиза Симпсон, Барт Симпсон, Мардж Симпсон.

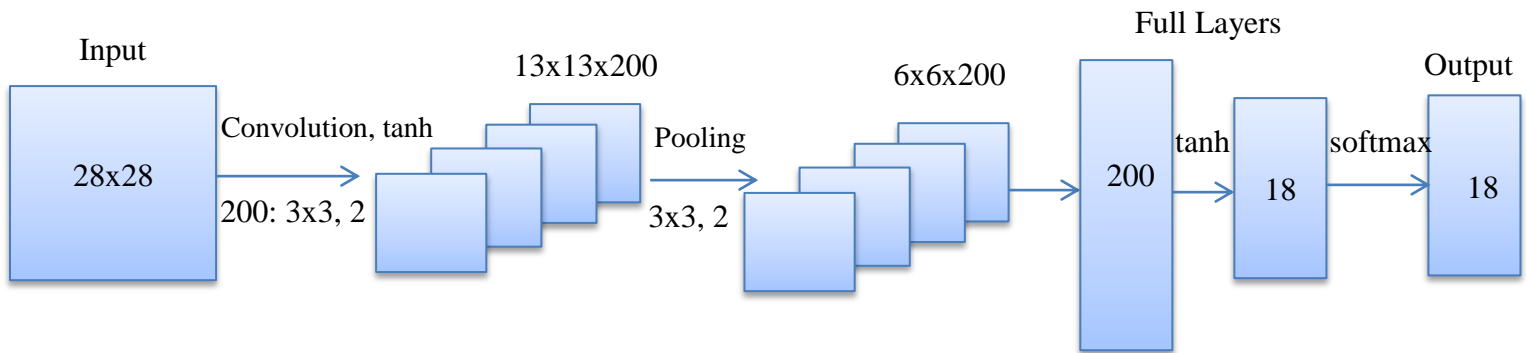
Для предварительной обработки данных использован скриптовый язык Python. Интерпретатор языка входит в дистрибутив Anaconda, который содержит ряд пакетов для анализа данных и машинного обучения.

Библиотека глубокого обучения MXNet может работать с различными типами входных данных, в том числе с однородными многомерными массивами ndarray из пакета NumPy. Для того, чтобы привести входные данные к такому формату, необходимо использовать библиотеку OpenCV для Python (opencv-python).

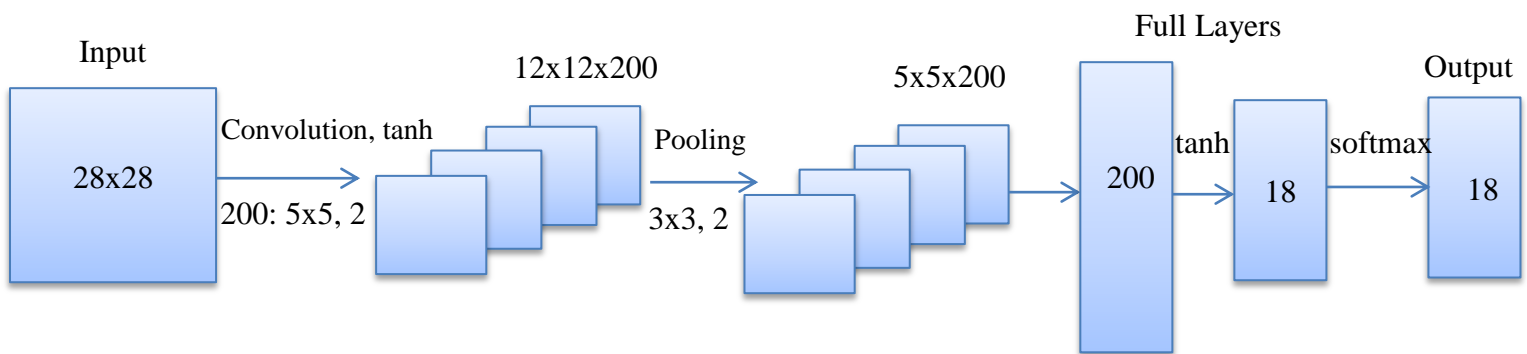
Был разработан скрипт, в котором с помощью функции `cv2.imread` считываются трёхканальные .jpg изображения в формате BGR (стандартное цветовое пространство OpenCV) и конвертируются в формат RGB; нормализуются (значение каждого пикселя делится на 255) и масштабируются до размера 28x28. Формируется массив меток от 0 до 17, соответствующих восемнадцати персонажам. Затем данные случайным образом делятся на обучающую и тестовую выборки в отношении 85% к 15%. Полученные ndarray-массивы `X_train`, `X_test`, `y_train`, `y_test` сохраняются в файлы для последующего использования нейронной сетью.

## Тестовые конфигурации нейронных сетей

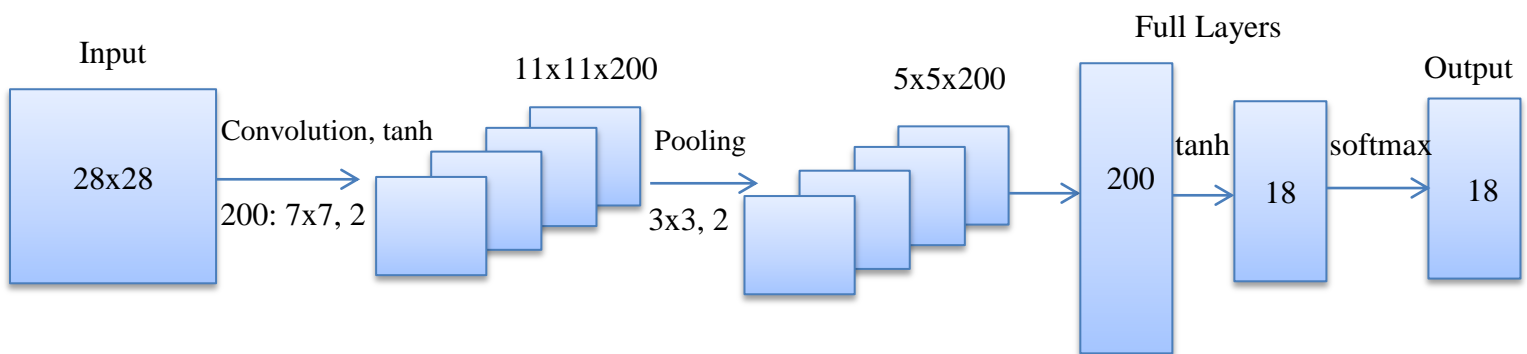
**Конфигурация №1.** Batch size = 10, optimizer = 'sgd', learning rate = 0.01.



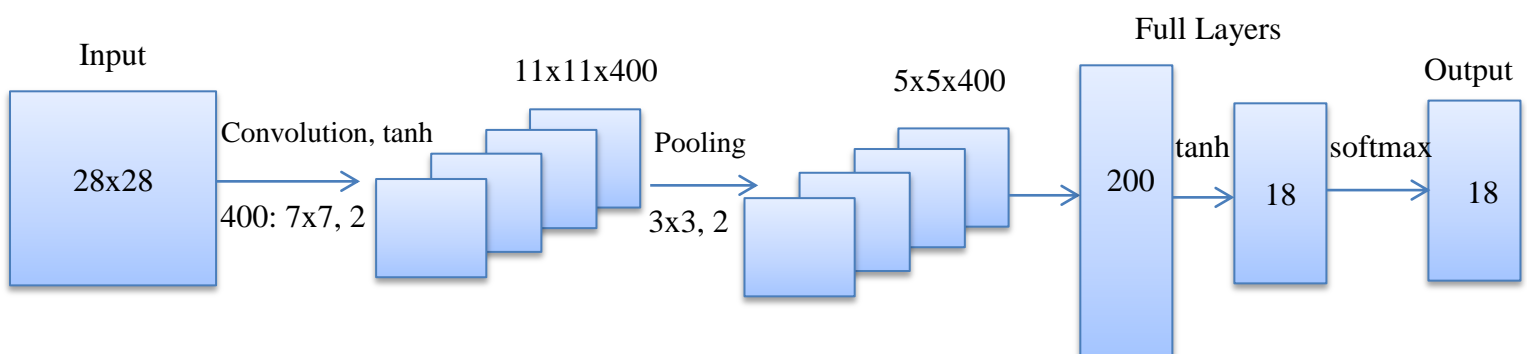
**Конфигурация №2.** Batch size = 10, optimizer = 'sgd', learning rate = 0.01.



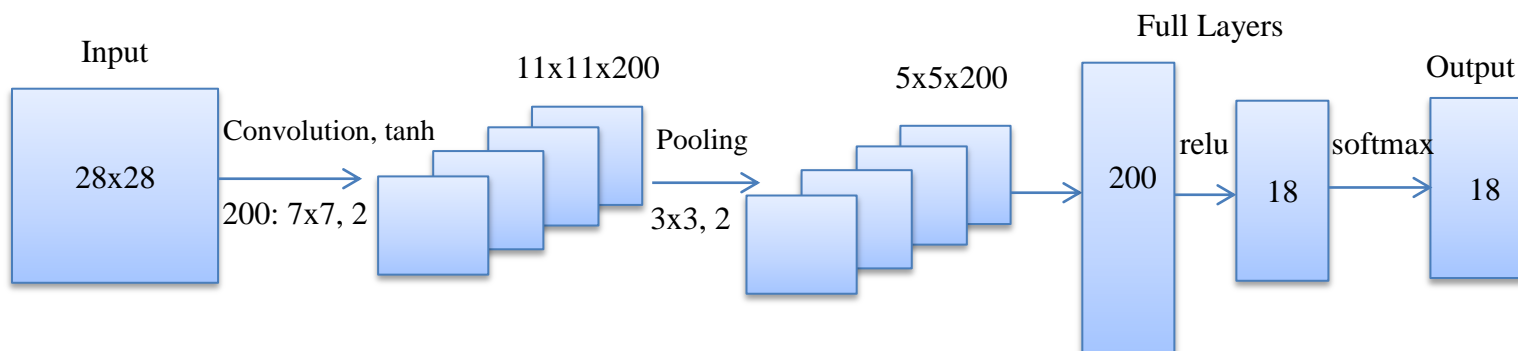
**Конфигурация №3.** Batch size = 10, optimizer = 'sgd', learning rate = 0.01.



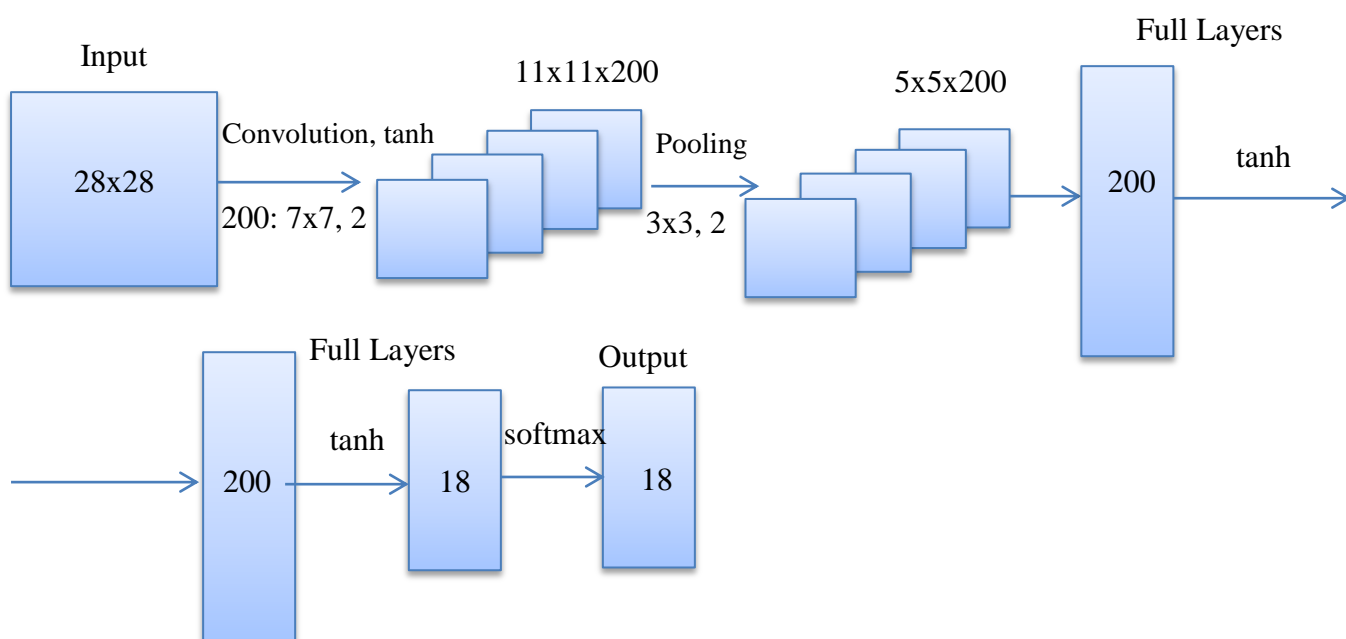
**Конфигурация №4.** Batch size = 10, optimizer = 'sgd', learning rate = 0.01.



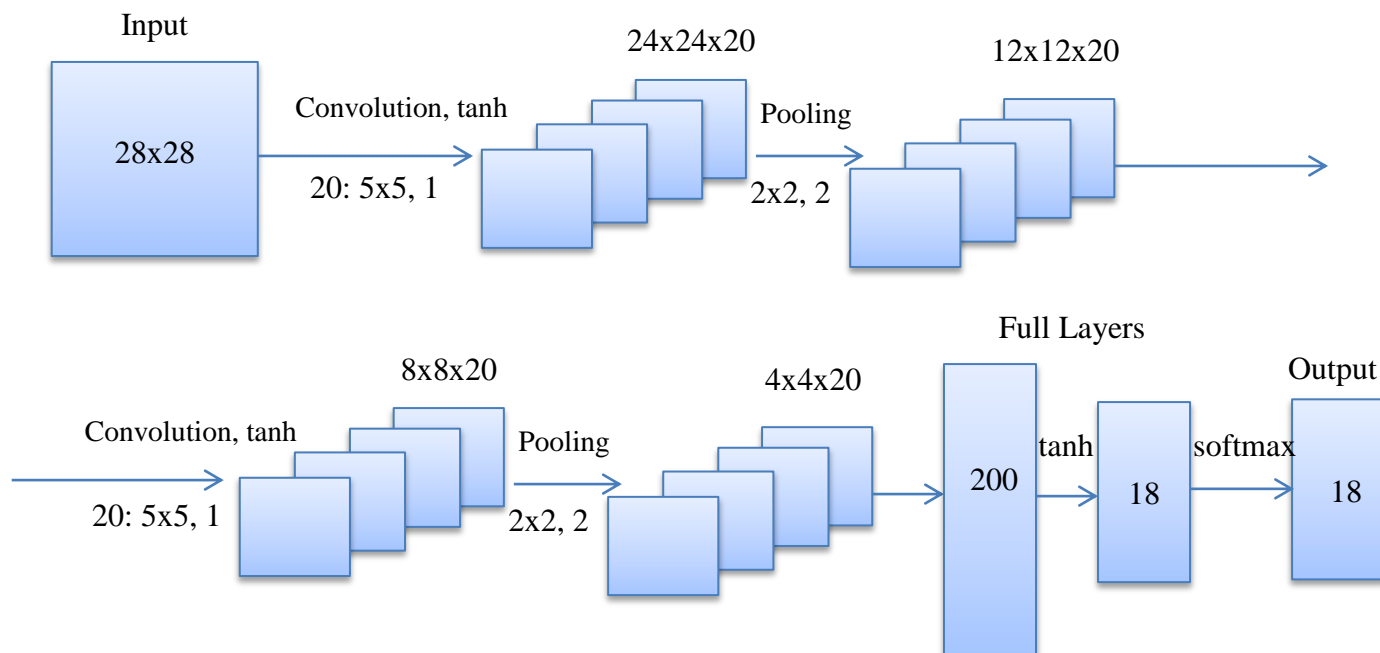
**Конфигурация №5.** Batch size = 10, optimizer = 'sgd', learning rate = 0.01.



**Конфигурация №6.** Batch size = 10, optimizer = 'sgd', learning rate = 0.01.



**Конфигурация №7.** Batch size = 10, optimizer = 'sgd', learning rate = 0.01.



## Результаты экспериментов

Эксперименты проводились на компьютере со следующей конфигурацией: процессор Intel® Core™ i3-5005U CPU @ 2.00GHz 2.00GHz, установленная память (ОЗУ) 4.00 ГБ, 64-разрядная операционная система Windows 8.1, интерпретатор Python 2.7.

Обучение проводилось до тех пор, пока точность на обучающей выборке не становилась равной 1.

Таблица 1. Результаты экспериментов

Конфигурация №	Время обучения модели, с	Точность классификации на тестовой выборке
1	2970	0.8760
2	2783	0.8802
<b>3</b>	<b>2990</b>	<b>0.8924</b>
4	4200	0.8837
5	3288	0.8866
6	3805	0.8502
7	3300	0.8597

Свёрточная нейронная сеть позволяет решать задачу классификации персонажей из Симпсонов с высокими показателями качества. Точность классификации лучше, чем в случае использования полностью связанной нейронной сети. Как можно заметить из таблицы 1, при увеличении глубины сети точность падает (конфигурации № 6 и №7).