МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования   
**«Национальный исследовательский   
Нижегородский государственный университет им. Н.И. Лобачевского»**

**(ННГУ)**

**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»

**ОТЧЕТ**

по лабораторной работе №4

**Начальная настройка весов полностью связанных**

**нейронных сетей**

**Выполнили:**

студенты группы 381603м4

Семеренко Александр

Кулдаев Александр

Горбунова Наталья

Третьякова Ольга

Морозова Юлия

Нижний Новгород

2018

Содержание

[Цели и задачи 3](#_Toc506444575)

[Выбор библиотеки 4](#_Toc506444576)

[Практическая задача компьютерного зрения 4](#_Toc506444577)

[Конфигурации нейронных сетей 5](#_Toc506444578)

[Результаты 7](#_Toc506444579)

Цели и задачи

В данной лабораторной работе необходимо использовать методы обучения без учителя для настройки начальных значений весов сетей, построенных при выполнении предшествующих практических работ для решения задачи бинарной классификации «еда – не еда».

Основными задачами данной лабораторной работы являются:

1. Выбор архитектур нейронных сетей, построенных при выполнении лабораторной работы №2.
2. Разработать модели автокодировщиков для выбранных архитектур нейронных сетей.
3. Обучить построенные глубокие модели.
4. Обучить выбранные архитектуры с начальной инициализацией весов сетей значениями, полученными в ходе обучения без учителя.
5. Протестировать обученные нейронные сети.

Выбор библиотеки

Для выполнения лабораторных работ была выбрана библиотека MXNet для языка программирования Python.

Практическая задача компьютерного зрения

Для выполнения лабораторной работы был выбран набор данных для решения задачи бинарной классификации: «еда» - «не еда». Были использованы картинки из набора данных https://www.kaggle.com/dansbecker/food-101/data в качестве «еды» и картинки из наборов данных https://www.kaggle.com/c/dogs-vs-cats/data и http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html в качестве «не еды». Итоговый набор данных состоит из 143125 изображений. С помощью скрипта im2rec.py, который входит в библиотеку MXNet, изображения были сконвертированы в формат .rec, который обрабатывается выбранной библиотекой, также картинки масштабировались до размера 128×128, и выборка разбивалась на тренировочную и тестовую в соотношении 60:40.

Вот несколько примеров изображений (рис.1, рис.2).



Рисунок 1. Примеры изображений из класса «еда»



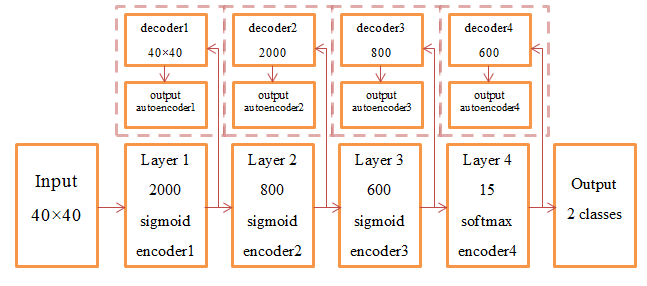
Рисунок 2. Примеры изображений из класса «не еда»

Конфигурации нейронных сетей

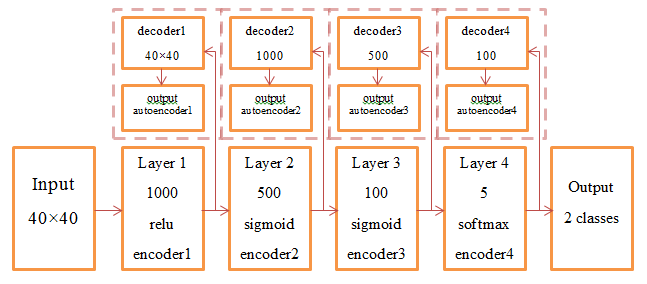
Для данной лабораторной работы данные были сжаты до размера 40×40, чтобы решить проблему с нехваткой оперативной памяти, также они были преобразованы в числовой формат .npy для работы с автокодировщиками.

Начальная настройка весов с помощью автокодировщиков была реализована для следующих конфигураций полностью связанных нейронных сетей, рассмотренных в лабораторной работе №2:

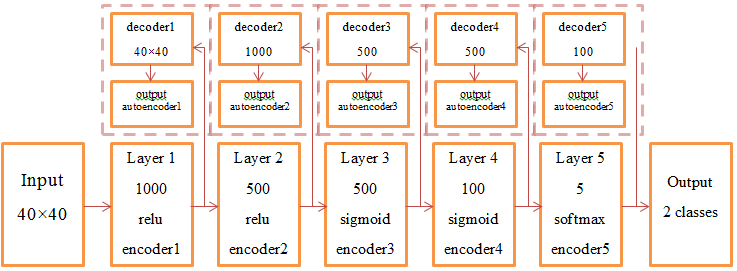
1. Конфигурация №1



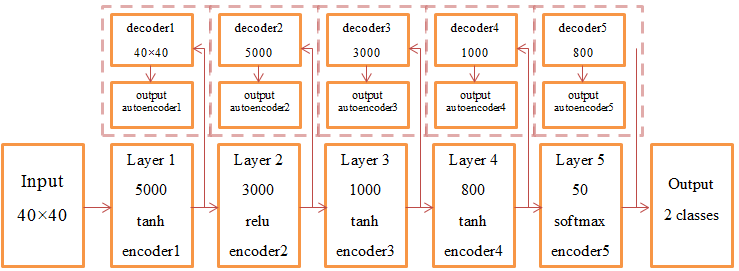
1. Конфигурация №2



1. Конфигурация №3



1. Конфигурация №4



Результаты

Во второй и третьей колонках в скобках указаны время обучения и точность классификации сети без начального приближения весов.

|  |  |  |
| --- | --- | --- |
| **Конфигурация** | **Время обучения модели, с** | **Точность классификации на тестовой выборке** |
| №1 | 276.53 (338.19) | 0.7055 (0.7055) |
| №2 | 131.8 (316.06) | 0.7055 (0.7616) |
| №3 | 140.38 (331.24) | 0.7055 (0.8166) |
| №4 | 299.28 (299.1) | 0.72669 (0.74347) |

На основе полученных результатов можно сделать вывод, что для данной задачи со значениями с начальной инициализацией весов, полученными в ходе обучения без учителя, точность классификации таких сетей уменьшилась или не изменилась по сравнению со случаем, когда веса были инициализированы случайными значениями. Возможно, это произошло из-за сжатия картинок с размера 128×128 до 40×40. Но при этом в среднем время обучения уменьшилось в 2 раза.