

**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**  
**Кафедра математического обеспечения и суперкомпьютерных технологий**  
Направление подготовки «Прикладная математика и информатика»  
Магистерская программа «Системное программирование»

**Отчет по лабораторной работе**  
**«Применение переноса обучения для решения задачи**  
**определения пола человека по фотографии лица»**

Выполнили:  
студенты группы 381603м4  
Гладилов, Волокитин, Левин,  
Новак

---

Нижний Новгород  
2017

## **СОДЕРЖАНИЕ**

<b>1</b>	<b>ПЕРЕНОС ОБУЧЕНИЯ</b>	<b>3</b>
<b>2</b>	<b>АРХИТЕКТУРЫ НЕЙРОННЫХ СЕТЕЙ</b>	<b>4</b>
<b>3</b>	<b>РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ</b>	<b>6</b>
<b>4</b>	<b>ИТОГИ</b>	<b>7</b>

# 1 Перенос обучения

Обучение глубоких нейронных сетей для решения всего многообразия задач, в которых они применяются, обычно не производится из случайной начальной инициализации весов. Наиболее распространенными причинами этого являются: невозможность найти необходимый набор данных достаточного размера для обучения сети нужной глубины; отсутствие вычислительных мощностей. В таких случаях на помощь приходит подход, основанный на переносе знаний уже полученных моделей.

Стратегии переноса обучения зависят от целого ряда факторов, но наиболее важными являются два: схожесть нового набора данных и исходного, а также размер нового набора. Основываясь на этих двух факторах можно различить 4 метода переноса знаний уже обученных моделей:

1. Новый набор данных меньше по размеру и аналогичен по содержанию исходному набору данных. Поскольку данные новой задачи схожи с изначальными, то можно предположить, что исходная нейронная сеть сможет решить новую задачу с приемлемой точностью.
2. Новый набор данных достаточно крупный и аналогичен по содержанию исходному набору. В таком случае можно попытаться использовать структуру глубокой модели, построенной для решения исходной задачи, однако обучения проводить на новом наборе.
3. Новый набор данных меньше по размеру и существенно отличается по содержанию от исходного набора. При таком сценарии можно попытаться использовать исходную модель, в качестве фиксированного метода извлечения признаков, заменив классификатор и до обучив его на новом наборе.
4. Новый набор данных относительно крупный и существенно отличается по содержанию от исходного набора данных. В таком случае можно обучать всю нейронную сеть с замененным классификатором начиная с весов заранее обученной модели – происходит тонкая настройка параметров модели, построенной для решения исходной задачи, для решения целевой задачи.

## 2 Архитектуры нейронных сетей

В качестве исходной модели была взята модель, предложенная на сайте с исходными данными <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/>, основой которой является архитектура VGG-16.ю веса которой были предобучены на наборе данных ImageNet. Архитектура исходной нейронной сети доступна изображена на (Рис. 1) (Увеличенная версия доступна в папке resources).



Рис. 1. Архитектура исходной нейронной сети

В качестве модели с заменённым классификатором использовалась модель, изображенная на (Рис. 2)

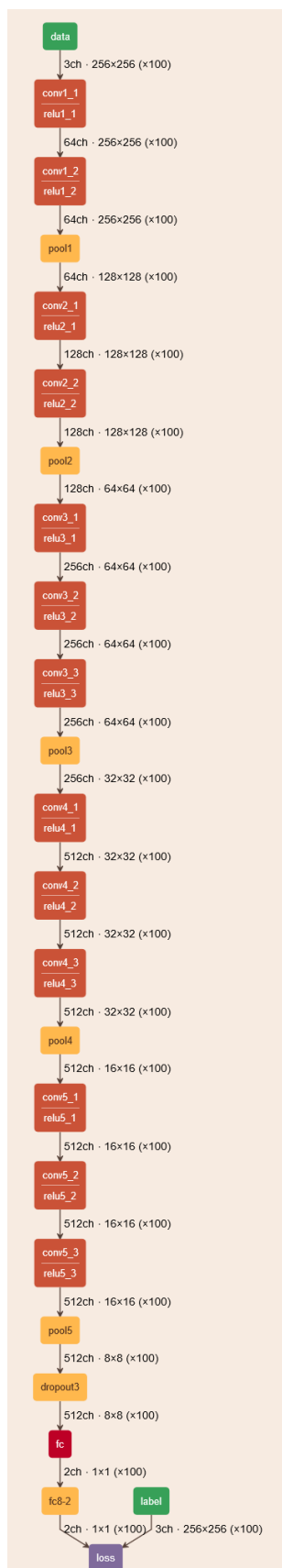


Рис. 2. Архитектура исходной нейронной сети с замененным классификатором

### 3 Результаты экспериментов

Тип эксперимента	Точность
<b>Эксперимент 1. Исходная нейронная сеть</b> (GeForce GTX 1080)	<b>0.8616</b>
<b>Эксперимент 2. Обучение исходной нейронной сети</b> (GeForce GTX 1080)	<b>0.768</b>
<b>Эксперимент 3. Обучение нового классификатора для исходной нейронной сети</b> (GeForce GTX 1080)	<b>0.9248</b>
<b>Эксперимент 4. Обучение исходной нейронной сети с новым классификатором</b> (GeForce GTX 1080)	<b>0.768</b>

В экспериментах 2 -4 критерий остановки – достижение 15000 эпох.

## 4 Итоги

В рамках данной лабораторной работы был изучен процесс переноса обучения глубоких нейронных сетей. Наилучший результат **92.48%** был показан в эксперименте 3, в рамках которого исходная нейронная сеть применялась для извлечения признаков. Что является наилучшим результатом среди всех рассмотренных нами нейронных сетей для решения задачи определения пола человека по фотографии лица.