

**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 распознавания**  
**жестов ASL»**

Выполнил:  
студент группы 381606-2м  
Пауль Э.А.

---

Нижний Новгород  
2018

## **СОДЕРЖАНИЕ**

<b>1</b>	<b>ПОСТАНОВКА ЗАДАЧИ</b>	<b>3</b>
<b>2</b>	<b>ПЕРЕНОС ОБУЧЕНИЯ</b>	<b>4</b>
<b>3</b>	<b>АРХИТЕКТУРЫ НЕЙРОННЫХ СЕТЕЙ</b>	<b>5</b>
<b>4</b>	<b>РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ</b>	<b>6</b>
<b>5</b>	<b>ВЫВОД</b>	<b>7</b>

# **1 Постановка задачи**

Целью данной лабораторной работы является изучение переноса обучения для решения задачи определения пола человека по фотографии лица.

В ходе лабораторной работы будут решены следующие задачи:

1. Выбор предобученной модели, решающей близкую задачу
2. Проведение четырех типов экспериментов по переносу обучения, описанных в секции 2

## 2 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.

### 3 Архитектуры нейронных сетей

В качестве исходной модели была взята модель с сайта <https://github.com/HolmesShuan/ResNet-18-Caffemodel-on-ImageNet>, основой которой является архитектура ResNet-18, веса которой были предобучены на наборе данных ImageNet. Архитектура исходной нейронной сети доступна по ссылке <http://ethereon.github.io/netscope/#/gist/bb14f8327c66e2419b21a81e0508e6d7>

В качестве модели с заменённым классификатором использовалась модель, полученная заменой последнего слоя fc1000 на такой же слой, но с количеством нейронов равным 31.

## 4 Результаты экспериментов

Тип эксперимента	Точность
Эксперимент 2. Обучение исходной нейронной сети	<b>0.999355</b>
Эксперимент 3. Обучение нового классификатора для исходной нейронной сети	<b>0.999193</b>
Эксперимент 4. Обучение исходной нейронной сети с новым классификатором	<b>0.999839</b>

## **5 Вывод**

В рамках данной лабораторной работы был изучен процесс переноса обучения глубоких нейронных сетей. Все полученные результаты высоки, однако можно выделить некоторые тенденции. Так как размер тестовой выборки нашей задачи довольно велик, поэтому наиболее неподходящий для нас эксперимент 3 с обучением только классификатора дал наименьшую точность. Также видно, что использование предобученных весов в эксперименте 4 дает преимущество по сравнению с экспериментом 2.