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: Прикладная математика и информатика

Магистерская программа: Компьютерные науки и приложения

**ОТЧЕТ**

по лабораторной работе №5

Тема:

**«Применение переноса обучения для задачи, поставленной во второй лабораторной работе »**

Выполнила:

Ягудина Анастасия (гр. 381803м4)

Проверил:

доцент, к.т.н. Кустикова В.Д.

Нижний Новгород

2020

Оглавление

[Постановка задачи 3](#_Toc33477186)

[Перенос обучения 3](#_Toc33477187)

[Пофайловое описание содержимого директории 5](#_Toc33477188)

[Вычислительные эксперименты 5](#_Toc33477189)

[Анализ результатов 6](#_Toc33477190)

# Постановка задачи

* **Цель работы:**

Цель настоящей работы состоит в том, чтобы исследовать возможности переноса обучения для решения целевой задачи, выбранной изначально для выполнения практических работ.

* **Задачи работы:**

1. Поиск исходной задачи (близкой по смыслу к целевой задаче ) и поиск натренированной для решения исходной задачи.
2. Выбор трех типов экспериментов по переносу знаний (типы экспериментов описаны в лекции).
3. Сбор результатов экспериментов.

# Перенос обучения

Пусть имеется некоторая *исходная задача А,* и для ее решения собраны тренировочные данные и выполнена их разметка. Аналогичная процедура выполнена и для некоторой другой *целевой задачи В.* Если данные задачи связаны некоторым образом, то имеет смысл использовать обученную модель исходной задачи для решения целевой задачи. В этом случае используется перенос обучения.

Цель переноса обучения состоит в том, чтобы использовать некоторые знания, накопленные в процессе обучения некоторой исходной *задачи А* , для решения близкой по смыслу к исходной задачи целевой *задачи В.*

*Понятие домена и задачи:*

*Домен-* определяется парой {*X, P(X)*}, где *X-* пространство признаков, *P(X)-* распределение случайной величины *х = {x1, x2, x3, …} X.* Два домена являются несовпадающими, если не совпадают пространства признаков или распределения вероятности.

*Задача-* определяется парой {*Y, f(.)*}, где *Y-* пространство меток, *f(.)-* функция предсказания, представляющая собой распределение условной вероятности того, что при заданном входном признаковом описании наблюдается определенная метка *f(x)= P(x|y).*Две задачи – несовпадающие, если различаются пространства меток или распределение вероятностей.

Тогда цель переноса обучения исходной задачи в домене и целевой задачи в домене t – повышение функции предсказания целевой задачи в домене t с использованием знаний, полученных при обучении задачи в домене , где домены и задачи не совпадают.

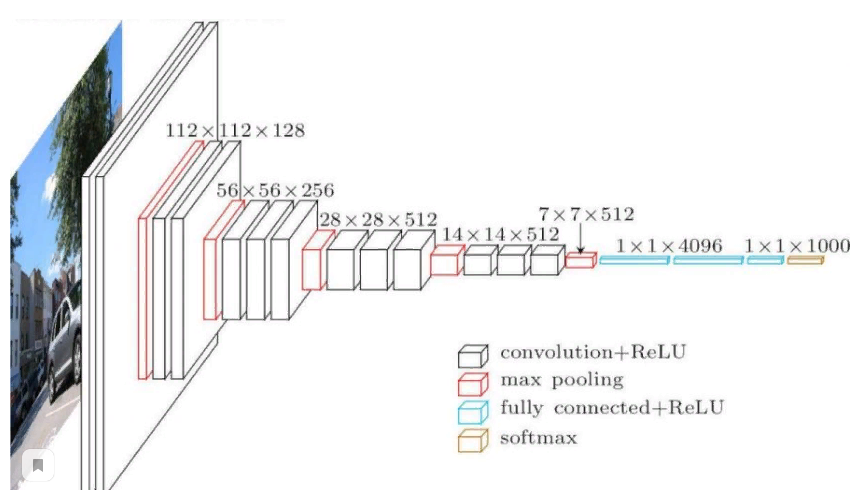
В данной работе рассматривается три эксперимента по переносу знаний, полученных при решении *исходной задачи А,*с целью обучить *целевую задачу В*.

* ***Использование структуры глубокой модели, построенной для решения задачи А, с целью обучения аналогичной модели для решения задачи В-*** предполагается, что модель, построенная для решения исходной задачи, обучается на данных, подготовленных для решения целевой задачи. При этом веса инициализируются случайным образом. Эксперимент реализует перенос знаний для родственных доменов.
* ***Использование модели, построенной для решения задачи А, в качестве фиксированного метода извлечения признаков при построении модели, решающей задачу В-*** идея: удалить из глубокой модели классификатор (последние полностью связанные слои) и рассматривать начальную часть сети как метод выделения признаков. Взамен старого классификатора помещается новый классификатор и обучается он на признаках, построенных с использованием начальной части сети. Эксперимент реализует перенос признакового описания.
* ***Тонкая настройка параметров модели, построенной для решения исходной задачи А, с целью решения задачи В-*** последние слои глубокой модели, решающие задачу А, заменяются новым классификатором и модель обучается как единая система. Эксперимент реализует перенос обучения на основе экземпляров.

Для проведения данных экспериментов была использована модель **VGG16**.

**VGG16** — модель сверточной нейронной сети, предложенная K. Simonyan и A. Zisserman из Оксфордского университета в статье “Very Deep Convolutional Networks for Large-Scale Image Recognition”. Модель достигает точности 92.7% — топ-5, при тестировании на ImageNet в задаче распознавания объектов на изображении. Этот датасет состоит из более чем 14 миллионов изображений, принадлежащих к 1000 классам.

VGG16 — одна из самых знаменитых моделей, отправленных на соревнование [ILSVRC-2014](https://neurohive.io/en/tag/segmentation/). Она является улучшенной версией [AlexNet](https://neurohive.io/ru/vidy-nejrosetej/alexnet-svjortochnaja-nejronnaja-set-dlja-raspoznavanija-izobrazhenij/" \t "_blank), в которой заменены большие фильтры (размера 11 и 5 в первом и втором сверточном слое, соответственно) на несколько фильтров размера 3х3, следующих один за другим. Сеть VGG16 обучалась на протяжении нескольких недель при использовании видеокарт NVIDIA TITAN BLACK.



*Рис1. Архитектура сверточной сети VGG16.*

[ImageNet](http://www.image-net.org/) — набор данных, состоящий из более чем 15 миллионов размеченных высококачественных изображений, разделенных на 22000 категорий. Изображения были взяты из интернета и размечены вручную людьми-разметчиками с помощью краудсорсинговой площадки Mechanical Turk от Amazon.

# Пофайловое описание содержимого директории

**src** – директория, содержащая разработанный для реализации архитектуры сверточной нейронной сети скрипт.

* lab\_5.ipynb –скрипт, содержащий реализацию трех экспериментов по переносу обучения, написанный на языке Python 3;
* lab\_5.docx – отчет.

# Вычислительные эксперименты

Рассмотрим результаты экспериментов по переносу обучения.

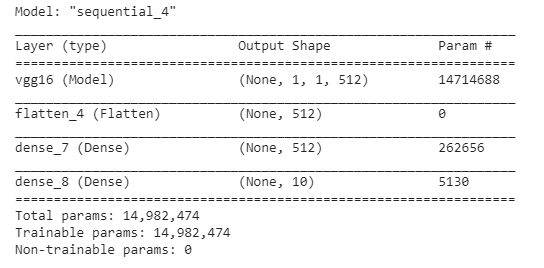
*Параметры:*

BATCH\_SIZE = 128

LEARNING\_RATE = 0.0001

EPOCHS = 20

OPTIMIZER=SGD

**

*Рис 2. Архитектура сети для экспериментов.*

Вычислительные результаты экспериментов представлены в таблице:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Experiments** | **Time train (sec)** | **Test accuracy** | **Train loss** | **Train accuracy** | **Epochs** |
| Experiment\_1 | 619 | 0.834 | 0.4693 | 0.8395 | 20 |
| Experiment\_2 | 495 | 0.922 | 0.0514 | 0.9824 | 15 |
| Experiment\_3 | 670 | 0.921 | 0.0296 | 0.9898 | 20 |

# Анализ результатов

В ходе выполнения работы можно сделать вывод, что эксперимент 3 решает поставленную задачу переноса обучения с высокими показателями и низкими потерями. Но использование переноса обучения малоэффективно по причине высоких затрат на обучение.