

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»  
(СПбГУ)

Направление: Прикладные математика и физика



Полуавтоматическое структурирование изображений  
в социальной сети с помощью методов машинного  
обучения

Санкт - Петербург  
2019

# Содержание

Введение	3
<b>1 Постановка задачи</b>	<b>4</b>
<b>2 Используемые инструменты</b>	<b>5</b>
<b>3 Сверточные нейронные сети</b>	<b>6</b>
3.1 Введение в теорию сверточных нейросетей . . . . .	6
3.2 Обзор используемых архитектур . . . . .	6
<b>4 Работа с данными</b>	<b>6</b>
4.1 Датасет SUN . . . . .	6
4.2 База WordNet . . . . .	6
4.3 Адаптация датасета SUN . . . . .	6
<b>5 Численные эксперименты</b>	<b>6</b>
5.1 Архитектура программы . . . . .	6
5.2 Особенности реализации . . . . .	6
5.3 Полученные результаты . . . . .	6
5.4 Обсуждение результатов . . . . .	6
<b>6 Валидация результатов</b>	<b>6</b>
Выводы	6
Список литературы	7

# Введение

С каждым днем пользователи социальных сетей создают и потребляют все бóльшие и бóльшие объемы информации, в том числе огромное количество фотографий. Построение системы для быстрой и точной навигации в миллионах изображений — не тривиальная задача. Одно из самых распространенных решений этой проблемы заключается в использовании тегов. Частный случай такого подхода — использование хештегов в социальных сетях.

*Хештег* — это любое слово или фраза без пробелов, перед которой стоит символ #, который называется *диез* или *решетка*, а в англоязычном варианте — *hash*, отсюда и название. Приведем несколько примеров: *#masterwork*, *#spbu*, *#htaglovesport*. Обычно в браузере или приложении хештег отображается как гипертекст, кликнув по которому можно получить список публикаций, снабженных таким же тегом.

Кроме простоты и удобства теги обладают еще одним полезным свойством — они позволяют не думать об иерархии структурируемой информации. Например, набор изображений можно разложить по папкам, создав иерархию по датам, геолокациям или авторству. Причем в отдельных случаях подобрать наиболее подходящую иерархию бывает затруднительно. Проблему можно решить так: достаточно поставить несколько тегов для всех изображений, а сами они могут храниться в плоской системе файлов. Благодаря этому свойству тегирование используется для рубрикации контента не только только онлайн, но и в оффлайн приложениях, например, просмотрщиках фотографий.

Можно выделить две разновидности популярных хештегов в социальных сетях. Первые, используемые недолго и посвященные каким-то социальным явлениям или событиям, например: *#elections2018*, *#metoo*. И вторые, широкораспространенные, но не связанные с новостной повесткой, например: *#sport*, *#cafe*; они и будут нас интересовать. Данная работа посвящена разработке интеллектуальной системы, подсказывающей пользователю релевантные хештеги к загружаемым фотографиям. Кроме того, с помощью такой системы можно решать и “обратную” задачу — определять, уместно ли поставлены те или иные теги к заданным изображениям. Способность системы давать ответ на такой вопрос можно использовать для выявления злоупотреблений со стороны пользователей. Например, зачастую в рекламных целях продвигаемые публикации снабжают множеством популярных тегов, не имеющих никакого отношения к публикуемой информации.

# 1 Постановка задачи

Целью настоящей работы является построение интеллектуальной системы для структурирования изображений в социальных сетях. А именно, предлагается автоматизировать процесс добавления пользовательских хештегов к загружаемым изображениям. С точки зрения машинного обучения решается задача классификации изображений.

В настоящее время наилучшие результаты в задаче классификации изображений удаётся получить, используя глубокие сверточные нейронные сети. В качестве примера можно привести один из самых известных и больших конкурсов по классификации изображений из датасета *ImageNet* [1], все призеры которого так или иначе используют сверточные нейронные сети.

## 2 Используемые инструменты

### Программные средства

В качестве языка программирования использовался *Python 3.6.7* (сборка *Anaconda*), в качестве среды разработки — *PyCharm Professional 2018.1*, операционная система *Ubuntu 16.04.4 LT*. Используются следующие сторонние библиотеки для *python*:

- *pytorch, torchvision* — построение и обучение нейронных сетей
- *tensorboardX* — визуальное логирование процесса обучения
- *PIL, opencv, scikit-image, scipy* — обработка изображений
- *matplotlib* — отрисовка графиков
- *numpy* — матричные вычисления
- *pandas* — работа с таблицами
- *nltk* — работа с текстом, в том числе с базой *WordNet*
- *scikit-learn* — библиотека машинного обучения общего плана
- *pip* — пакетный менеджер

### Вычислительные мощности

Обучение моделей производилось на удаленном сервере со следующей конфигурацией:

- видеокарта *GEFORCE GTX 1080 Ti*, (11 ГБ видеопамяти)
- процессор *AMD Ryzen Threadripper 1920X 12-Core*
- оперативная память объемом 62 ГБ

## 3 Сверточные нейронные сети

### 3.1 Введение в теорию сверточных нейросетей

### 3.2 Обзор используемых архитектур

## 4 Работа с данными

### 4.1 Датасет SUN

### 4.2 База WordNet

### 4.3 Адаптация датасета SUN

## 5 Численные эксперименты

### 5.1 Архитектура программы

### 5.2 Особенности реализации

### 5.3 Полученные результаты

### 5.4 Обсуждение результатов

## 6 Валидация результатов

## Выводы

## Список литературы

- [1] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. [2015] *ImageNet Large Scale Visual Recognition Challenge*. IJCV.