

Санкт-Петербургский государственный университет

Искусственный интеллект и наука о данных

Юдин Александр Владимирович

Выявление групп пользователей социальной сети  
на основе объектов на аватарах

Отчёт по учебной (ознакомительной) практике

Научный руководитель:

Старший преподаватель кафедры информатики, Ф.В. Бушмелев

Санкт-Петербург

2023

## Оглавление

Введение .....	3
Описание предметной области .....	5
Извлечение признаков.....	8
Проверка гипотезы .....	9
Заключение.....	10
Список литературы .....	12

## Введение

В современном обществе социальные сети играют ключевую роль в обмене информацией, коммуникации и формировании онлайн-сообществ, за последние годы наблюдается рост аудитории и ее активности в сети интернет, что вызывает интерес исследователей из разных областей [1], [2]. Пользователи, при оформлении своего профиля, используют различные аватары, на основе которых можно выявить дополнительные сведения о личности. В этом контексте интересным и перспективным направлением исследований является определение групп пользователей на социальных платформах путем анализа объектов, присутствующих на их аватарах.

Методы машинного обучения и компьютерного зрения стали мощными инструментами, которые используются для извлечения знаний, например, из цифровых следов пользователя. Одним из источников подобной информации могут выступать социальные сети, содержащие большое количество различного контента об активности пользователя, например, текстовые посты, графический контент, анкетная информация [2]. Стоит отметить, что в области изучения текстовых постов исследователями получены значимые результаты, в то время как, вопрос анализа цифровых изображений до сих пор остается актуальным.

Аватары служат визуальным представлением пользователей на платформах социальных сетей и могут содержать ценную информацию о предпочтениях, интересах и социокультурных связях. Существуют различные подходы к анализу аватаров пользователя, в данном исследовании рассмотрен метод, основывающийся на классификации графических объектов аватара.

Таким образом, в рамках данной работы ставится **цель** по проверке гипотезы о том, что психометрические характеристики пользователей социальной сети различаются в зависимости от числа живых и не живых объектов, изображенных на аватаре. Для достижения цели были поставлены следующие задачи:

1. провести анализ предметной области и выбрать предварительно обученную модель для детекции объектов для проведения эксперимента;
2. при помощи выбранной модели для каждого аватара определить количество живых и неживых объектов, изображенных на нем;
3. провести дисперсионный анализ для проверки поставленной гипотезы.

## **Описание предметной области**

Несмотря на то, что аватары предоставляют огромное количество визуальной информации, проблема заключается в эффективном извлечении значимых паттернов и идентификации групп пользователей на основе объектов, изображенных на аватарах. Для решения этой задачи требуются передовые методы компьютерного зрения и модели обнаружения объектов, способные различать и классифицировать объекты на изображениях.

Понимание групп пользователей в социальных сетях очень важно для различных приложений, включая целевую рекламу, рекомендации контента и управление сообществами. Традиционные методы анализа аватаров часто не справляются со сложностью и разнообразием визуальной информации. Данное исследование направлено на устранение этих недостатков путем использования современных моделей обнаружения объектов, включая Faster R-CNN, детекторы YOLO, SSD, Mask R-CNN и RetinaNet.

Теперь, рассмотрим выбранные для исследования модели для обнаружения объектов.

Архитектура Faster R-CNN (Region-based Convolutional Neural Network) включает в себя сеть предложений регионов (RPN) для генерации потенциальных местоположений объектов, механизм объединения регионов интереса (RoI) для извлечения признаков и опорную CNN для общего представления признаков. Процесс обучения включает в себя тренировку RPN для получения точных предложений регионов и тонкую настройку для конкретных задач обнаружения объектов [5].

Модель YOLO (You Only Look Once) делит входное изображение на сетку и одновременно предсказывает ограничительные рамки и вероятности классов. Детекторы YOLO работают эффективно, что делает их подходящими для приложений, где важна скорость.

Ключевые характеристики данной модели представляют собой, во-первых, одномоментное обнаружение, в частности детекторы YOLO используют стратегию "одного выстрела", предсказывая ограничивающие

рамки и вероятности классов за один проход, во-вторых, обработка в режиме реального времени, что означает способность обрабатывать изображения в режиме реального времени, что делает их подходящими для приложений с низкими требованиями к задержкам [8].

SSD (Single Shot MultiBox Detector) — это модель обнаружения объектов в компьютерном зрении, получившая признание благодаря своему подходу "одного выстрела" и способности предсказывать несколько ограничивающих боксов и вероятности классов за один проход. Модель делит входное изображение на сетку ограничительных рамок по умолчанию и предсказывает характеристики объекта в разных масштабах, позволяя обнаруживать объекты различных размеров.

Ключевые характеристики:

- 1) однопроходный подход: SSD предсказывает несколько ограничивающих рамок и вероятности классов за один проход, упрощая процесс обнаружения;
- 2) многомасштабное обнаружение: модель работает в нескольких масштабах, позволяя обнаруживать объекты разных размеров в рамках одной системы.

Mask R-CNN — это усовершенствованный вариант модели Faster R-CNN, в котором наряду с обнаружением объектов реализована возможность сегментации экземпляров. Она позволяет одновременно локализовать, классифицировать и сегментировать объекты на пиксельном уровне в изображении [6].

Расширение Faster R-CNN состоит в том, что Mask R-CNN расширяет архитектуру Faster R-CNN и включает в себя сегментацию объектов. Также, модель предоставляет маски сегментации на уровне пикселей для каждого обнаруженного объекта, что позволяет детально разграничить объекты [6].

RetinaNet — это модель обнаружения объектов, разработанная для решения проблемы дисбаланса классов, присущей многим задачам обнаружения объектов. Она использует фокальные потери для придания

большого веса трудно классифицируемым примерам, что улучшает способность модели обнаруживать объекты в различных масштабах.

RetinaNet использует фокальные потери для устранения дисбаланса между классами переднего и заднего плана во время обучения. Также, модель использует FPN для захвата признаков в нескольких масштабах, что повышает ее способность обнаруживать объекты различных размеров.

В зависимости от порога IoU (Intersection over Union), модели показывают разную эффективность, которая указана в таблице 1.

Название	AP25	AP50	AP75	mAP
Faster R-CNN	60.1%	38.3%	21.6%	39.67%
YOLO	52.3%	36.1%	20.7%	36.37%
SSD	50.6%	34.2%	18.3%	34.37%
Mask R-CNN	62.2%	39.5%	22.1%	41.27%
RetinaNet	40.3%	26.1%	14.5%	26.97%

Таблица 1 – сравнение эффективности разных моделей

В представленной выше таблице 1 эффективности различных моделей, обученных на датасете COCO, отображены значения метрик Average Precision (AP), которые измеряют точность обнаружения объектов. AP25, AP50 и AP75 представляют собой варианты Average Precision (AP), но с разными порогами IoU (Intersection over Union). IoU измеряет степень перекрытия между предсказанным объектом и истинным объектом на изображении. Также, в таблице отображены значения метрики Mean Average Precision (mAP) — это усредненное значение AP по всем классам. Обычно используется для общей оценки качества работы детектора объектов на всех классах. На основе таблицы 1 можно отметить, что модель Mask R-CNN является наиболее эффективной. Именно она была использована в данном исследовании.

## Извлечение признаков

В работе была использована готовая реализация Mask R-CNN из библиотеки PyTorch предварительно обученная на датасете MS COCOv2017. Данный датасет состоит из 91 класса, которые в результате работы были сведены к двум: живой и не живой.

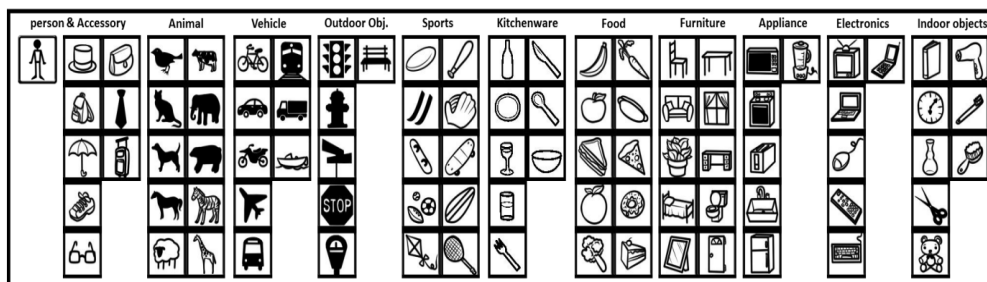


Рисунок 1 – классы в датасете COCO

На вход модели подается изображение аватара пользователя, на выходе же получаем список вероятностей принадлежности к конкретным классам.

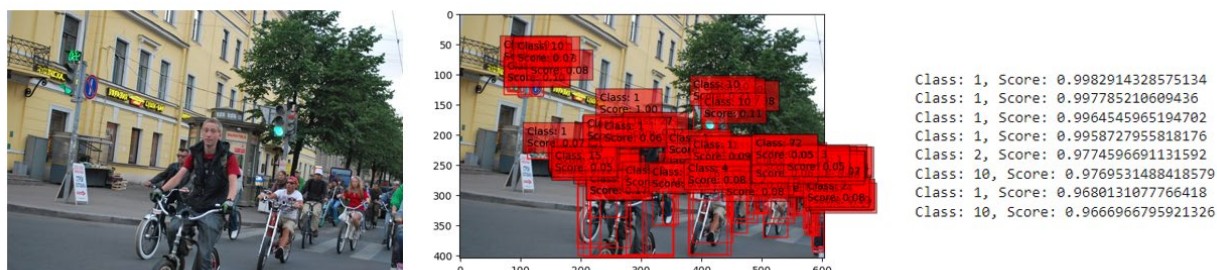


Рисунок 2 – пример работы модели

На рисунке 2 приведен пример: слева изображение, которое подается на вход, справа то, что получается на выходе. Модель определила объекты на изображении, ограничила области этих объектов, и выдала список с вероятностями.



## Проверка гипотезы

После того, как был получен список с классами для каждого изображения из датасета, был проведен дисперсионный анализ. Проверялась нулевая гипотеза о том, что не существует значимых различий между уровнями переменной `person_type`. Если посмотреть значение  $Pr > F$ , видим, что оно велико, следовательно, гипотеза верна, и статистических отличий между уровнями переменной `person_type` нет.

Multivariate linear model						
=====						
Intercept	Value	Num DF	Den DF	F Value	Pr > F	
-----						
Wilks' lambda	0.5991	2.0000	2061.0000	689.6498	0.0000	
Pillai's trace	0.4009	2.0000	2061.0000	689.6498	0.0000	
Hotelling-Lawley trace	0.6692	2.0000	2061.0000	689.6498	0.0000	
Roy's greatest root	0.6692	2.0000	2061.0000	689.6498	0.0000	
-----						
person_type	Value	Num DF	Den DF	F Value	Pr > F	
-----						
Wilks' lambda	1.0000	2.0000	2061.0000	0.0023	0.9977	
Pillai's trace	0.0000	2.0000	2061.0000	0.0023	0.9977	
Hotelling-Lawley trace	0.0000	2.0000	2061.0000	0.0023	0.9977	
Roy's greatest root	0.0000	2.0000	2061.0000	0.0023	0.9977	
=====						

Рисунок 3 – результаты дисперсионного анализа

## **Заключение**

В ходе исследования были рассмотрены следующие модели:

- 1) Faster R-CNN;
- 2) YOLO;
- 3) SSD;
- 4) Mask R-CNN;
- 5) RetinaNet.

Из представленных моделей была выбрана модель Mask R-CNN, так как она обладает наилучшими показателями эффективности, в зависимости от порога IoU (Intersection over Union).

При помощи выбранной модели для анализируемых аватаров было определено количество живых и неживых объектов, изображенных на них.

В результате проведенного дисперсионного анализа была проверена гипотеза о том, что психометрические характеристики пользователей социальной сети могут различаться в зависимости от числа живых и не живых объектов, изображенных на их аватарах. Анализ данных позволил выявить статистически значимые различия между группами пользователей, основанные на содержании их аватаров.

Результаты анализа указывают на важность визуального представления влияния на психометрические характеристики пользователей. Особенно интересным оказалось влияние изображения живых и не живых объектов на определенные аспекты поведения и взаимодействия в социальной сети.

Так, гипотеза о том, что психометрические характеристики пользователей социальной сети различаются в зависимости от числа живых и не живых объектов, изображенных на аватаре не подтвердилась.

Эти результаты могут иметь практическое значение для разработки и улучшения пользовательского опыта, а также для адаптации маркетинговых стратегий, учитывая визуальные предпочтения и представления пользователей и предоставляют базу для дальнейших исследований, направленных на

улучшение взаимодействия и вовлеченности пользователей в социальных сетях.

## **Список литературы**

1. Brandanalytics. Social media in Russia. - URL: <https://brandanalytics.ru/blog/social-media-russia-autumn-2023/> (дата обращения: 21.12.2023).
2. Datareportal. Russia 2023. - URL: [datareportal.com/reports/digital-2023-russian-federation](https://datareportal.com/reports/digital-2023-russian-federation) (дата обращения: 21.12.2023).
3. He K., Gkioxari G., Dollár P. and Girshick R., Mask R-CNN // IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980-2988.
4. Jiang S., Jiang H., Ma S., & Jiang Z. Detection of Parking Slots Based on Mask R-CNN // Applied Sciences, 2020, 10, 4295.
5. Majgaonkar S., Waghela K., Shah R., Pavate A. and Mishra P. Developing a Parking Monitoring System using Mask-RCNN // SAMRIDDI: A Journal of Physical Sciences, Engineering and Technology. 12, SUP 1 (Jun. 2020), pp. 386-389.
6. Naufal A.R., Fatichah C., & Suciati N. Preprocessed Mask RCNN for Parking Space Detection in Smart Parking Systems // International Journal of Intelligent Engineering and Systems, 2020, 13, pp. 255-265.
7. Nyambal J. and Klein R., Automated parking space detection using convolutional neural networks // Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech), 2017, pp. 1-6.
8. Patel R. and Meduri P., Car detection based algorithm for automatic parking space detection // 19th IEEE International Conference on Machine Learning and Applications (ICMLA), 2020, pp. 1418-1423.