

Санкт-Петербургский Государственный Университет  
Математико-механический факультет  
Математическое обеспечение и администрирование  
информационных систем

Фролова Марина Сергеевна, 342 группа

# Анализ зависимостей между жанром фильма и визуальными составляющими его обложки

Отчёт по учебной практике

Научный руководитель:  
к. ф-м. н., доцент Графеева Н.Г.

Санкт-Петербург  
2021

# Оглавление

<b>1. Введение .....</b>	<b>3</b>
<b>2. Цель и задачи .....</b>	<b>4</b>
<b>3. Обзор существующих методов выделения визуальных элементов у изображений .....</b>	<b>5</b>
3.1 Яркость, насыщенность, контраст .....	5
3.2 Основные цвета .....	6
3.3 Люди на изображениях .....	6
<b>4. Выделение признаков и анализ информации .....</b>	<b>7</b>
4.1 Используемые средства и технологии .....	7
4.2 Подбор данных .....	7
4.3 Яркость и насыщенность .....	8
4.4 Основные цвета .....	10
4.5 Люди на изображениях .....	11
<b>5. Заключение .....</b>	<b>13</b>
<b>6. Список литературы .....</b>	<b>14</b>

# 1. Введение

За 2020 год объем экономического сегмента развлекательного цифрового контента в России вырос на 44% и составил 123,4 млрд руб. [1]. Потребителей окружает огромное количество потенциально интересных аудио и видео записей в цифровом формате.

И пользователи, и создатели контента все более заинтересованы в цифровых методах анализа информации – потребители хотят сориентироваться в большом потоке медиа-данных и иметь возможность быстро находить интересный им контент, производителям же интересен анализ наиболее популярного контента с целью лучшего удовлетворения потребительского спроса.

В данной работе исследуется зависимость визуальной составляющей обложки фильма от его жанра. А именно, анализируются отдельные признаки изображения и их информативность для решения поставленной задачи. В наши дни это является актуальным, т. к. выявленные закономерности возможно применить не только для создания модели автоматического определения жанра произведения, но и при построении рекомендательных систем и отрисовке новых обложек, наиболее успешных с точки зрения привлечения внимания и побуждения к просмотру кинокартины.

## 2. Цель и задачи

Цель: проверить информативность отдельных признаков изображения для определения жанра.

Для достижения этой цели были поставлены следующие задачи:

1. Подбор данных в формате обложка фильма - жанр
2. Выделение визуальных элементов у изображений
3. Анализ закономерностей между визуальными признаками изображений и жанром картины.

## 4. Обзор существующих методов выделения визуальных элементов у изображений

### 4.1 Яркость, насыщенность, контраст

При формировании вектора признаков изображения зачастую выделяют яркость и насыщенность [10]. Их анализ естествен для человеческого восприятия изображений – при взгляде на картинку человек довольно быстро считывает ее освещенность и насыщенность цветов [8].

Традиционно, изображения находятся в цветовом пространстве RGB, кодирующем цвета с помощью трех каналов: Red (красный), Green (зеленый), Blue (синий). Для выделения и анализа яркости в работе [27] используется цветовое пространство Lab, кодирующие цвет с помощью канала яркости (L) и двух каналов цветности (от зеленого до пурпурного (a) и от синего до желтого (b)). Согласно [10], человеческий глаз воспринимает яркость отдельно от цветового тона, что учитывается в Lab. Кроме того, в Lab, в отличие от RGB, изменение цвета и яркости линейно с точки зрения человеческого восприятия [8].

В работе [21] для определения насыщенности изображения преобразовывают в цветовой формат HSV, кодирующий цвета с использованием каналов цветового тона (H), насыщенности (S), значения цвета (V). Это позволяет получать значение насыщенности напрямую.

Согласно [7], внимание человека при просмотре изображения в первую очередь привлекают неоднородные области – участки с большим контрастом. Контраст – это зрительное восприятие разницы между разными участками изображения. Человеческий глаз различает три вида контраста – по яркости, по насыщенности и по цветовому тону [8].

Существует много способов определения коэффициента контраста. Один из них – контрастность Майкельсона, расчет которого при анализе изображений рассмотрен в работе [22].

## 4.2 Основные цвета

Люди воспринимают цвет по-разному, но практически у каждого есть устойчивые ассоциации с определенными цветами и их оттенками/температурой [9]. Крупные производители часто пользуются этим – например, эко-товары часто упаковываются в «природные» цвета – зеленый, коричневый, цвет дерева. С помощью анализа цветов мы получаем большое количество информации о изображениях [8].

Именно поэтому выделение основных цветов на изображениях является распространенной задачей в сфере компьютерного зрения и обработки изображений. Для ее решения в работах [12, 23] используется метод k-means [11]. Для этого изображения преобразовываются в HSV формат и применяется алгоритм кластеризации, в котором каналы h (тон), s (насыщенность) и v (значение) рассматриваются как 3 независимые переменные.

## 4.3 Люди на изображении

Мы привыкли, что практически на каждой обложке фильма изображены его главные герои – это помогает зрителю быстро получить много информации о произведении (гендер, пол, возраст героев, их социальный статус и т. д.). Однако, это делает все обложки довольно похожими. В данной работе было интересно посмотреть, существует ли корреляция между наличием людей на изображении и жанром фильма. Чтобы решить данную задачу, необходимо автоматически определить количество портретов людей на изображениях.

В работе [14] рассматриваются классические и нейросетевые методы в задачах компьютерного зрения. Несмотря на то, что классические подходы обучаются и архитектурно проще, их точность зачастую уступает нейросетевым.

В задачах распознавания и детектирования в данный момент наиболее часто используются сверточные нейронные сети, т. к. их архитектура нацелена на эффективное распознавание образов [24,25]. Одной из популярных сверточных нейронных сетей для детектирования лиц людей является MTCNN [26]. Данная нейронная сеть предобучена (т. е. ее веса заданы), что позволяет сразу использовать модель.

## 5. Выделение признаков и анализ информации

### 5.1 Используемые средства и технологии:

- Python 3.8.1 — высокоуровневый язык программирования общего назначения [2].
- Jupyter Notebook — open-source среда разработки [15].
- OpenCV — библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом [3].
- MTCNN — open source пакет для решения задачи компьютерного зрения с помощью нейросетевых подходов [6] Использовался для обнаружения лиц на изображениях.
- Pandas - Python библиотека для анализа данных [16].
- Matplotlib - Python библиотека для построения графиков и визуализации данных [17].
- Seaborn — библиотека визуализации данных, построенная на основе matplotlib [18].
- Scikit-learn — Python-библиотека для выполнения высокоуровневых вычислений, анализа данных и машинного обучения [5].
- NetworkX — Python библиотека для построения и обработки графов [20].

### 5.2 Подбор данных

В данной работе используется набор данных, взятый с сайта kaggle.com. Данные содержат 15177 картинок - обложек фильмов в формате .jpg с разрешением 200\*300px и глубиной цвета 24 bit, а также подписи к ним, содержащих информацию об одном или нескольких жанрах, к которым принадлежит картина.

В исходном наборе данных содержались фильмы 16 различных жанров (один фильм мог принадлежать к нескольким жанрам одновременно). Я исключила из рассмотрения малочисленные (до 1000 фильмов). В итоге осталось 12: action (боевик), adventure (приключение), drama (драма), crime (криминал), thriller (триллер), comedy (комедия), horror (ужасы), romance (мелодрама), detective

(детектив), cartoon (мультфильм), sciencefiction (научная фантастика), fantasy (фэнтези).

### 5.3 Яркость, насыщенность, контраст

В данной работе было интересно проследить закономерности между усредненной яркостью и насыщенностью по обложке, контрастом по яркости и насыщенности и жанром, к которому обложка принадлежит. Для решения данной задачи были получены значения первого, пятого и последнего дециля яркости и насыщенности по каждому изображению. Яркость получена как канал L (lightness, яркость) в цветовом пространстве Lab, насыщенность – как S (saturation, насыщенность) в HSV.

Далее были построены распределения медианных яркостей и насыщенностей по жанрам, а также распределения яркостного контраста и контраста по насыщенности. Контраст рассчитывался как коэффициент Майкельсона - разность между девятым и первым децилем.

При анализе плотностей распределения медианных яркостей обложек выяснилось, что наибольшей медианной яркостью обладают обложки комедий, мелодрам и мультфильмов, наименьшей – ужасы, детективы и триллеры (рис. 1).

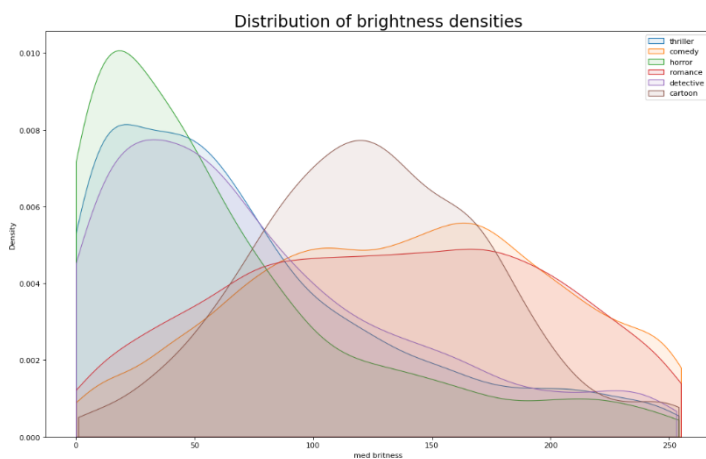


Рисунок 1

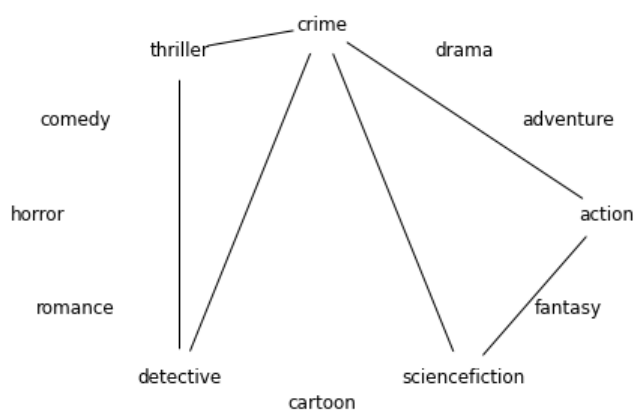


Рисунок 2

Распределения яркостей по жанрам не являются нормальными, поэтому для нахождения жанров со схожими распределениями (проверки гипотезы о принадлежности двух выборок одному закону распределения) я воспользовалась непараметрическим критерием



однородности Колмогорова-Смирнова[19]. При дальнейшем анализе я также пользовалась им.

В этом и дальнейших случаях для визуализации результатов применения критерия я строила граф, в котором ребро между двумя вершинами означает принятие гипотезы о равенстве законов распределений (т. е. значение статистики критерия меньше порогового). Таким образом, выделяются 2 группы со схожим распределением яркости: детектив - триллер – криминал и криминал – боевик – научная фантастика (рис. 2).

При визуальном анализе графика плотности распределения медианной насыщенности видно, что мультфильмы сильно смещены вправо относительно всех остальных, т. е. в среднем их обложки более насыщены (рис. 3). При применении критерия Колмогорова-Смирнова снова выделяется группа триллер – детектив – криминал, а также приключение – научная фантастика – фэнтези, боевик – ужас, боевик – научная фантастика (рис. 4).

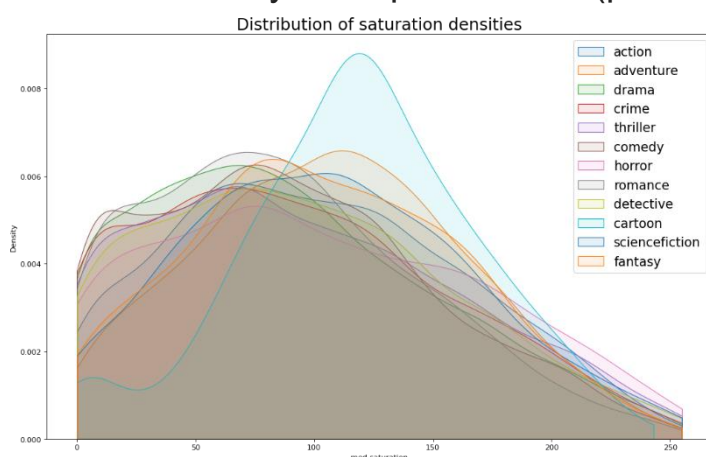


Рисунок 3

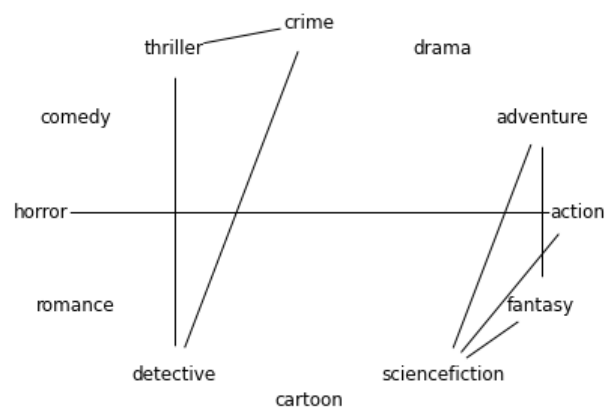


Рисунок 4

При анализе яркостного контраста получилось, что наименьшим медианным контрастом обладают мультфильмы (по-видимому, их обложки однородно яркие), наибольшим – ужасы, криминал и детектив. При применении критерия однородности выделяется группа криминал – детектив – триллер – боевик и триллер - ужас.

Анализ контраста по насыщенности показал, что наименьшим медианным контрастом обладают мультфильмы (получается, их обложки однородно насыщенные), наибольшим – комедии. При применении критерия однородности выделяются пары триллер – криминал, романтика – боевик, боевик – детектив, приключение – фэнтези.

## 5.4 Основные цвета

В данной работе было интересно проследить, есть ли корреляция между часто используемыми цветами на обложках фильмов и жанрами, к которым они принадлежат.

Для этого проводилась кластеризация каждого изображения по методу k-means. Изображения были переведены в пространство HSV, кластеризация осуществлялась по каналу H (hue, тон). Согласно результатам «метода-локтя», проведенному на небольшой подвыборке данных, изображения разбивались на 3 основных цвета. Для дальнейшего анализа используется 1 самый распространенный цвет на изображении.

Для анализа основных цветов по жанрам я использовала функцию интерполяции из модуля `scipy.interpolate`. Таким образом, для каждого жанра получена зависимость между цветовым тоном и количеством пикселей на картинке, имеющих этот тон. С помощью `scipy.optimize` я нашла точки максимума и минимума. Результаты представлены на рис. 5, 6 (точки первого и второго локальных максимумов), 7 (точка локального минимума). Точка глобального максимума совпадает с первым локальным максимумом. Таким образом, самый «частый» цвет на обложках – синий и его оттенки.



Рисунок 5

Рисунок 6

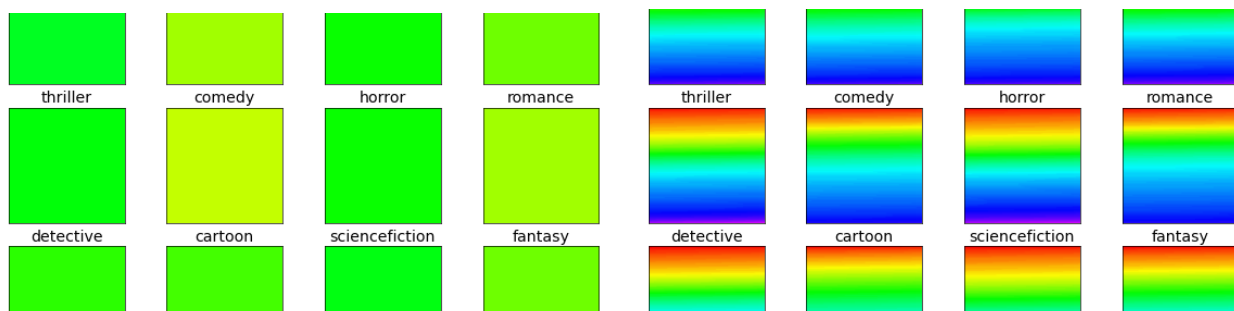


Рисунок 7

Рисунок 8

Точка глобального минимума у всех жанров – красный цвет (значение  $\text{hue} = 175$  в модели  $\text{hsv}$ ).

Также было интересно явно визуализировать распределение основного цветового тона на обложках по жанрам (рис. 8). Это позволяет сделать вывод, что наиболее равномерное распределение основного цвета у жанров «научная фантастика», «ужасы» и «триллер» (явно присутствует фиолетовый тон, полосы практически одинаковой ширины), т. е. их обложки наиболее разнообразны по цветам. Снова мультики наиболее выделяются среди всех жанров – у них наиболее выражен зеленый тон, а второй точкой локального максимума является желтый (в отличие от оттенков оранжевого у других жанров).

В результате применения критерия Колмогорова-Смирнова выделяются группы криминал – приключение – боевик и триллер – ужас – детектив со схожим цветовым распределением.

## 5.5 Люди на изображении

Для обнаружения портретов людей на обложках в данной работе используется предобученная сверточная нейронная сеть MTCNN. С ее помощью получено количество людей на каждом изображении.

В результате анализа данных о количестве портретов людей на обложках получилось, что среднее количество по всем жанрам – 1 человек (точное значение среднего: 1.12), а максимальное число – 18. На рис. 9 представлен box график по жанрам.

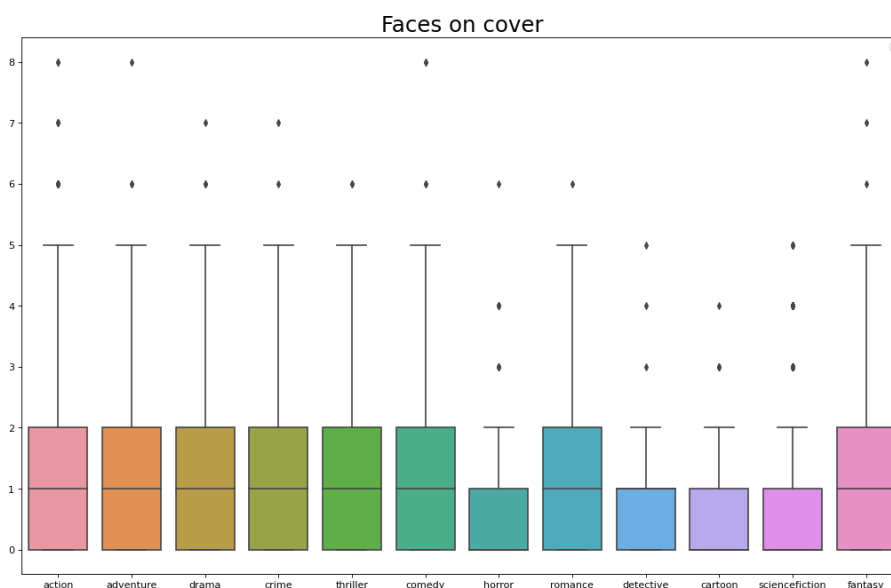


Рисунок 9

Из графика видно, что медианное значение людей на изображении у жанров “ужас”, “детектив”, “мультик”, “научная фантастика” равно 0, 75% значений равны 0 или 1. У остальных жанров 75% значений находятся в пределах 0 – 2, а медианное значение равно 1. В результате применения критерия Колмогорова-Смирнова снова выделяется группа триллер – боевик – детектив. Так же близкие распределения имеют жанры криминал – романтика, приключение – фэнтези, приключение – научная фантастика.

## 6. Заключение

После анализа исходного набора данных по вышеописанным критериям можно сделать вывод, что рассмотренных визуальных признаков недостаточно для однозначного определения жанра. Однако анализ показал, что уже с имеющимся вектором признаков возможно безошибочно отделить жанр мультфильм от всех остальных. Кроме того, можно с большой долей уверенности относить фильм к определенной подгруппе жанров, например, триллер – детектив – боевик и приключение – фэнтези – научная фантастика чаще всего выделялись в отдельные группы после применения критерия однородности Колмогорова-Смирнова.

В результате работы можно сделать вывод, что наиболее информативными оказались признаки «Медианная яркость», «Яркостный контраст», «Медианная насыщенность», «Основной цвет». Их распределения наиболее различны, а критерий Колмогорова-Смирнова позволяет явно определить отдельные подгруппы со схожими значениями признака.

В дальнейшей работе планируется увеличить количество рассматриваемых визуальных признаков изображений, а также построить модель классификации обложек фильмов по жанрам.

# Список литературы

1. Анализ экономики рунета за 2020г [Электронный ресурс] // [https://www.cnews.ru/news/top/2021-0520\\_ekonomika\\_runeta\\_v\\_2020\\_godu](https://www.cnews.ru/news/top/2021-0520_ekonomika_runeta_v_2020_godu)
2. Python [Электронный ресурс] // <https://www.python.org>
3. OpenCV [Электронный ресурс] // <https://opencv.org>
4. Openpyxl [Электронный ресурс] // <https://pypi.org/project/openpyxl/>
5. Scikit-learn [Электронный ресурс] // <https://scikit-learn.org/stable>
6. MTCNN [Электронный ресурс] // <https://pypi.org/project/mtcnn/>
7. Бабенко В.В., Явна Д.В. Конкуренция за внимание пространственных модуляций градиентов яркости // Российский психологический журнал. 2018. Т. 15, № 3. С. 160–189. DOI: 10.21702/rpj.2018.3.8
8. Железняков В.Н. Цвет и контраст: Технология и творч. выбор : Учеб. пособие для студентов вузов кинематографии // Всерос. гос. ин-т кинематографии (ВГИК), 2001. ISBN 5-87149-062-X
9. М.В. Миронова Экспериментальные исследования на предмет цветовосприятия // Амурский государственный университет, 2010.
10. В. В. Шаронов Свет и Цвет // Государственное издание физико-математической литературы, Москва. 1961.
11. Steinhaus H Sur la division des corps materiels en parties // Acad. Polon. Sci., C1. III vol IV: 801—804. 1956.
12. М. С. Шмаков, А. А. Товмасян Анализ цветовой палитры изображений для определения преобладающих цветовых тонов // Белорусский государственный технологический университет, 2015. ISSN 1683-0377. Труды БГТУ. 2015. № 9.
13. Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Yu Qiao Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks / 2016. DOI: 10.1109/LSP.2016.2603342
14. А. В. Грачев Классические и нейросетевые подходы к проблеме распознавания дорожных знаков / СПбГУ, Санкт-Петербург, 2020. DOI: 10.24411/2658-4964-2020-10025
15. Jupyter Notebook [Электронный ресурс] // <https://jupyter.org/>
16. Pandas [Электронный ресурс] // <https://pandas.pydata.org/>
17. Matplotlib [Электронный ресурс] // <https://matplotlib.org/>
18. Seaborn [Электронный ресурс] // <https://seaborn.pydata.org/>

19. Лемешко Б.Ю. Непараметрические критерии согласия. Руководство по применению. – М.: ИНФРА–М, 2014. – 163с.
20. NetworkX [Электронный ресурс] // <https://networkx.org/>
21. С.Н. Безрядин, П.А. Буров, Д.Ю. Ильиных, А.А. Львов Модели изменения яркости в современном программном обеспечении // Вестник Саратовского государственного технического университета, 2007. Т.1 № 4, С. 106-110.
22. Муравьева С. В., Пронин С. В., Шелепин Ю. Е. Контрастная чувствительность зрительной системы человека // Экспериментальная психология, 2010. Т. 3, № 3. С. 5–20. ISSN: 2072-7593 / 2311-7036
23. Н. О. Котелина, Б. Р. Матвийчук Кластеризация изображения методом k-средних // Вестник Сыктывкарского университета, 2019. № 3 (32). С. 101-112.
24. Ю. В. Алексеенко Разработка системы распознавания изображений на основе сверточных нейронных сетей // Евразийский союз ученых, 2017. № 7-1 (40). С. 8-11.
25. Д. В. Плотников, Е. А. Сопов Решение задач распознавания лиц и мимики с помощью сверточных нейронных сетей // Решетневские чтения, 2017. Т. 2. С. 234-236.
26. И. Л. Гринин, С. В. Кравченко Системы обнаружения распознавания лиц, основанные на различных архитектурах нейронных сетей: анализ и сравнение технологий // Инновации и инвестиции, 2021. № 6. С. 130-133.
27. Д. И. Зимин, В. А. Фурсов Технология определения восстанавливающего фильтра и обработки цветных изображений // Компьютерная оптика, 2005. № 27. С. 170-173.