| Министерство образования и науки Российской Федерации | | |
| --- | --- | --- |
| федеральное государственное бюджетное образовательное учреждение высшего образования | | |
| «Иркутский государственный университет» | | |
| (ФГБОУ ВО «ИГУ») | | |
| Факультет бизнес-коммуникаций и информатики | | |
|  | | |
|  | |
| **ОТЧЕТ ПО УЧЕБНОЙ ПРАКТИКЕ**  **Б2.В.02(У) Технологическая (проектно-технологическая) практика**  **по направлению 09.04.03 «Прикладная информатика»**  **(прикладная магистратура)**  **профиль «Прикладная информатика (Большие данные и интеллектуальные системы)»** | |
|  | |
|  |  |
|  | Студент 1 курса очной формы обучения, группа 14121-ДМ |
|  | \_\_\_\_\_\_\_\_\_\_ Д.С. Павлов |
|  |  |
|  | Руководитель: к.ф.–м.н., доцент  \_\_\_\_\_\_\_\_\_\_ С.М. Пещерова |
|  | |
|  |  |
|  |  |
|  |  |
|  |  |
|  | |

[**ВВЕДЕНИЕ 3**](#_qvgmz768xch0)

[**Теоретическая часть 4**](#_hk9a7waenvck)

[2.1 Искусственный интеллект и компьютерное зрение 4](#_kcxylqytv33t)

[2.2 ResNet: остаточные сети 5](#_g1w0hdg128se)

[2.3 BLIP: генерация описаний изображений 6](#_ij2s56l1ysp4)

[2.4 Извлечение визуальных признаков 7](#_ankd12dvvbjd)

[2.5 Визуализация внимания: Class Activation Map (CAM) 8](#_uwva39b7woq5)

[2.6 Семантический граф 9](#_yfzkbfz65ljh)

[**6. Заключение 9**](#_nv4vc7nk942u)

# **ВВЕДЕНИЕ**

На сегодняшний день информационные технологии переживают один из самых мощных витков развития — и в самом центре этого процесса стоит искусственный интеллект. Он уже не просто инструмент — это полноценный партнёр, способный решать задачи, ранее доступные только человеку. Особенно ярко это проявляется в области анализа визуальных данных.

В рамках своей учебной практики я сосредоточился на направлении, которое соединяет технологию и искусство — разработке интеллектуальной системы для анализа художественных изображений. Этот проект получил название **Art Analyzer**.

Что я поставил перед собой?  
Цель практики была ясна и амбициозна — получить реальный опыт в создании веб-приложения, предназначенного для анализа изображений методами ИИ

Для этого был сформирован комплекс задач:

* Изучение архитектур нейросетей **ResNet** и **BLIP** — основа анализа;
* Подготовка датасетов с картинами, разбитыми по жанрам и стилям;
* Обучение моделей для автоматической классификации этих жанров и стилей;
* Реализация алгоритмов извлечения признаков — от классических HOG и LBP до GLCM и ORB;
* Разработка удобного веб-интерфейса для анализа изображений;
* Создание визуализаций: карты внимания (CAM), цветовые палитры, семантические графы, гистограммы;
* И финальное тестирование и демонстрация работоспособности всей системы.

В условиях цифровизации культуры, когда визуальный контент растёт лавинообразно, становится жизненно необходимым иметь инструмент, который не просто «видит» картину, но и *понимает* ее. **Art Analyzer** отвечает на этот вызов.

Это не просто классификатор. Это полноценная платформа, которая определяет жанр, стиль, генерирует краткое описание, показывает, на что именно модель «смотрит», какие цвета преобладают, как соотносятся элементы изображения — и всё это доступно через интуитивный интерфейс.

Для кого это?

Для исследователей, галерей, художников, преподавателей и просто ценителей искусства — для всех, кто хочет получить глубинное понимание визуального материала с помощью современных технологий.

В результате мы получили работающий веб-сервис с мощным API, живой анимацией и встроенными примерами. Это — не просто учебный проект.

# **Теоретическая часть**

В основе проекта Art Analyzer лежат современные методы машинного обучения, ориентированные на обработку изображений. Для реализации задач классификации, описания и визуализации изображений были использованы нейросетевые архитектуры ResNet и BLIP, а также алгоритмы извлечения признаков, такие как HOG, LBP, ORB и GLCM. Ниже приведено краткое теоретическое описание используемых методов.

## **2.1 Искусственный интеллект и компьютерное зрение**

Давайте начнем с ключевого направления, которое лежит в основе всего проекта — *компьютерное зрение (CV)*.

Это не просто одна из ветвей искусственного интеллекта. Это технология, которая учит машины «видеть» — а главное, *понимать* то, что они видят. Компьютерное зрение занимается извлечением, интерпретацией и анализом информации из изображений и видео потоков.

Революционный скачок в этой области произошел с появлением *сверточных нейронных сетей* — CNN. Благодаря им уровень точности в распознавании объектов, лиц, сцен, а также стилистических и жанровых характеристик художественных изображений вышел на принципиально новый уровень.

Сегодня эти архитектуры не просто классифицируют — они *понимают контекст*, умеют выделять ключевые элементы и даже интерпретировать стиль и настроение визуального контента.

Именно на этих технологиях строится фундамент **Art Analyzer**. Я опирался на самые современные достижения в области CV, чтобы дать приложению способность видеть изображение *не глазами, а интеллектом.*

## **2.2 ResNet: остаточные сети**

В рамках проекта **Art Analyzer** я использовал одну из наиболее зарекомендовавших себя нейросетевых архитектур — **ResNet-50**. Это глубокая сверточная сеть, построенная на принципе остаточных соединений — так называемых *shortcut connections*.

Обычные глубокие сети сталкиваются с проблемой: по мере увеличения числа слоев обучение становится нестабильным — градиенты затухают. **ResNet** решает эту задачу элегантно: она позволяет сети «перепрыгивать» через слои, сохраняя информацию и улучшая обучаемость. Это значит, что можно строить действительно *глубокие модели* — вплоть до 152 слоёв — без потери качества.

Ключевые особенности ResNet-50, на которых я сделал акцент:

* Использование *остаточных блоков*, обеспечивающих стабильность и эффективность;
* Высокая устойчивость к переобучению;
* Проверенная точность на глобальных датасетах, включая **ImageNet**.

В проекте **Art Analyzer** ResNet-50 решает три ключевые задачи:

* *Классификация жанра* произведения;
* *Классификация художественного стиля*;
* *Генерация CAM-карт* — визуализаций, которые показывают, на что именно обращает внимание модель при анализе изображения.

Эта архитектура стала технологическим фундаментом моей системы — надёжным, масштабируемым и интеллектуально зрелым решением.

## **2.3 BLIP: генерация описаний изображений**

Для того чтобы сделать анализ картин по-настоящему многогранным, я добавил в систему компонент, который позволяет не только *видеть*, но и *говорить* — то есть *генерировать описания* на основе изображения. Эту задачу решает модель **BLIP —** Bootstrapping Language-Image Pre-training.

Это мощная трансформерная архитектура, построенная на симбиозе двух направлений: **Vision Transformer** анализирует изображение, а *GPT-подобный декодер* — формирует связный текст.

Что умеет BLIP?

* *Создаёт заголовки и краткие описания* на основе визуального контента;
* *Интегрирует визуальные и текстовые данные,* обеспечивая контекстно осмысленный вывод;
* Прекрасно масштабируется и обучается на разнородных мультимодальных датасетах.

В **Art Analyzer** модель BLIP интегрирована через экосистему **HuggingFace Transformers**, что позволило получить готовое API и гибко встраивать ее в веб-приложение.  
 Каждое загружаемое пользователем изображение автоматически сопровождается *текстовым описанием*. Это делает анализ не только визуальным, но и *семантически осмысленным*, добавляя системе уровень «интерпретации», а не просто распознавания.

## **2.4 Извлечение визуальных признаков**

В дополнение к нейросетевым моделям в проекте реализованы классические методы анализа изображений.

**HOG (Histogram of Oriented Gradients)**

Используется для выделения границ и контуров. Алгоритм разбивает изображение на ячейки, вычисляет градиенты и строит гистограммы направлений. Классика компьютерного зрения.

**LBP (Local Binary Patterns)**

Метод анализа текстуры. Для каждого пикселя формируется бинарный шаблон на основе его окружения. Это даёт возможность точно описывать текстуру — от мазков кисти до поверхности холста.

**ORB (Oriented FAST and Rotated BRIEF)**

Мощный и быстрый детектор и дескриптор ключевых точек изображения, устойчивый к масштабированию и поворотам. Используется в задачах сопоставления и сравнения картин.

**GLCM (Gray-Level Co-occurrence Matrix)**

Позволяет анализировать *текстурные характеристики*: контраст, однородность, энтропию и другие параметры, которые формируют "визуальный почерк" произведения. на основе матрицы совместной встречаемости серых уровней.

## **2.5 Визуализация внимания: Class Activation Map (CAM)**

CAM позволяет буквально *увидеть*, какие области изображения стали определяющими при принятии решения моделью. Это делается за счет *взвешенного суммирования признаков последних слоев нейросети*, после чего получается так называемая *карта внимания*. Она накладывается на исходное изображение и показывает — куда именно "смотрит" модель, когда делает вывод.

CAM используется в Art Analyzer для:

* Объяснения предсказаний — чтобы пользователь мог понять, почему система отнесла картину к тому или иному жанру или стилю;
* Повышения интерпретируемости модели — потому что в работе с искусственным интеллектом важно не только получить результат, но и понимать *механизм его появления*.

## **2.6 Семантический граф**

Чтобы сделать результат анализа ещё более наглядным и осмысленным, в проекте **Art Analyzer** я поэкспериментировал и реализовал инструмент визуализации *семантических связей* — динамически формируемый граф.

В этом графе каждый узел представляет собой жанр или стиль, который модель распознала в картине. Размер узла соответствует уровню уверенности — чем выше вероятность, тем больше и заметнее элемент на графе.

Связи между узлами могут отражать контекстную и стилистическую близость: например, как один стиль может быть связан с определенным жанром, или как модель трактует художественное содержание картины в более широком культурном контексте. В общем хороший инструмент для дальнейшего развития.

# **6. Заключение**

В ходе учебной практики мной был реализован полнофункциональный проект Art Analyzer, сочетающий в себе передовые технологии искусственного интеллекта и веб-разработки. Его основная задача — интеллектуальный анализ художественных изображений.

Проект наглядно продемонстрировал, как современные нейросетевые архитектуры и алгоритмы компьютерного зрения могут быть объединены в работающую систему с реальным пользовательским интерфейсом и интерактивной визуализацией.