

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра математической кибернетики и компьютерных наук

**РАЗРАБОТКА ВЕБ-ПРИЛОЖЕНИЯ РАСПОЗНАВАНИЯ ЖАНРОВ
ЖИВОПИСИ НА ОСНОВЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ С
ПОМОЩЬЮ КАМЕРЫ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студентки 5 курса 551 группы
направления 09.03.04 — Программная инженерия
факультета КНиИТ
Ковалевой Софьи Андреевны

Научный руководитель
доцент, к. т. н.

В. М. Соловьев

Заведующий кафедрой
к. ф.-м. н., доцент

С. В. Миронов

Саратов 2025

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Обзор принципов работы искусственных нейронных сетей и их применения для работы с изображением	5
1.1 Понятие искусственной нейронной сети.....	5
1.2 Обзор видов нейронных сетей.....	5
1.3 Описание особенностей сверточной нейронной сети	5
2 Применение трансферного обучения модели для классификации изображений	6
2.1 Определение трансферного обучения	6
2.2 Принцип работы трансферного обучения для классификации изображений	6
2.3 Сохранение и загрузка модели для дальнейшей работы после трансферного обучения	6
3 Практическая часть.....	7
3.1 Подготовка данных для обучения	7
3.2 Создание базовой модели MobileNetV3Large и сравнение её с другими моделями	7
3.3 Обучение модели	7
3.4 Тонкая настройка модели	8
3.5 Собственная оценка точности модели	8
4 Экспериментальная часть с созданием веб-приложения.....	9
4.1 Streamlit — инструмент для создания веб-приложений	9
4.2 Создание веб-приложения с помощью Streamlit	9
4.3 Размещение приложения в сети Интернет	10
ЗАКЛЮЧЕНИЕ	12

ВВЕДЕНИЕ

В настоящее время применение машинного обучения расширяется, охватывая всё больше сфер нашей повседневной жизни. Данные системы обучаются на больших наборах данных, что позволяет им качественно применять полученные знания при распознавании лиц, выявлении различных заболеваний на медицинских снимках и прогнозах каких-либо событий.

Одно из наиболее интересных и быстроразвивающихся направлений машинного обучения – компьютерное зрение. Задача этой области – извлечение информации из изображений или видео. Например, классификация изображений, поиск конкретного объекта на видео.

В этой работе будет рассмотрена одна из базовых задач компьютерного зрения – классификация изображений. Алгоритм распознавания изображений принимает изображение в качестве входных данных и выводит то, что содержит изображение. Другими словами, вывод – это метка класса.

Актуальность темы выпускной квалификационной работы объясняется тем, что разработка веб-приложения с распознаванием жанра живописи внесет свой вклад в цифровизацию культуры, так как поможет искусствоведам автоматизировать ручные задачи. Например, систематизировать художественное наследие, так как вручную произведения искусства может быть классифицировано с ошибкой, или же ускорить поиск жанра в цифровых музеях с помощью автоматической разметки картин для оцифровки коллекций.

Данный проект можно использовать в образовательных целях для онлайн-курсов по истории искусств с автоматической проверкой заданий или для геймификации викторин. Следовательно, веб-приложение позволит искусство сделать более доступным для образования и популяризации, а также помочь систематизировать художественное наследие.

Целью дипломной работы является изучение принципа работы сверточной нейронной сети, трансферного обучения, а также разработка классификатора изображений на языке программирования Python с помощью библиотеки TensorFlow. На основе разработанного классификатора необходимо создание веб-приложения для распознавания жанра живописи с помощью камеры.

Чтобы достичь этого должны быть решены следующие задачи:

1. описание видов нейронных сетей и их различие;
2. изучение принципа работы сверточных нейронных сетей;

3. описание принципа работы трансферного обучения;
4. описание принципа работы сохранения и загрузки модели после обучения;
5. изучение языка программирования Python и функций библиотеки TensorFlow;
6. сравнение различных моделей для обучения из Keras;
7. настройка и обучение выбранной модели;
8. разработка собственной оценки предсказаний;
9. создание веб-приложения с распознаванием жанра живописи с помощью камеры на основе Streamlit;
10. развертывание веб-приложения на облачной платформе для доступа из сети Интернет с применением Docker и Yandex Cloud.

Структура и объем работы. Бакалаврская работа состоит из введения, четырех разделов, заключения, списка использованных источников и трех приложений. Общий объем работы – 64 страницы, из них 50 страниц – основное содержание, включая 30 рисунков, цифровой носитель в качестве приложения, список использованных источников информации – 20 наименований.

1 Обзор принципов работы искусственных нейронных сетей и их применения для работы с изображением

1.1 Понятие искусственной нейронной сети

Искусственная нейронная сеть — упрощенная модель биологической нейронной сети, представляющая собой совокупность искусственных нейронов, взаимодействующих между собой.

Нейрон — это элемент, вычисляющий выходной сигнал по определенному правилу из совокупности входных сигналов.

Синапс — это входные сигналы, связь между двумя нейронами.

1.2 Обзор видов нейронных сетей

Сеть прямого распространения — искусственные нейронные сети, где сигнал строго идёт от входного слоя к выходному. Нейроны в этой сети не образуют цикла, они называются полносвязными.

Рекуррентная нейронная сеть (RNN) — нейросеть, которая имеет циклическую связь.

Сверточная нейронная сеть (CNN) — это нейронные сети, которые применяются для задач классификации, кластеризации и распознавания изображений.

1.3 Описание особенностей сверточной нейронной сети

CNN обрабатывает эти данные, применяя последовательность специализированных операций, таких как свертка, пулинг и активация.

— Сверточный слой.

Сверточный слой — это тип слоя нейронной сети, который применяет операцию свёртки к входным данным.

— Пулинговый слой.

Данный слой необходим для уменьшения размера сверточного слоя, понижая вычислительную мощность и количество весов, которая необходима для обработки данных.

— Полносвязный слой.

Нейроны в этом слое имеют полную связь со всеми нейронами предыдущего и последующего слоя. На этом этапе происходит изучение нелинейных комбинаций высокоуровневых признаков, устанавливая сложные нелинейные взаимосвязи между ними.

2 Применение трансферного обучения модели для классификации изображений

2.1 Определение трансферного обучения

Трансферное обучение позволяет использовать знания из ранее изученных задач и применять их к аналогичным с ними задачам. Этот факт означает быстрое развертывание сети, обеспечивая более быструю производительность, настройку и тестирование модели. Этот подход популярен в компьютерном зрении и в обработке естественного языка.

2.2 Принцип работы трансферного обучения для классификации изображений

- Выделение признаков.

Для выделения признаков используются представления, полученные предыдущей моделью, для извлечения признаков из новых образцов, которые позже пропускаются через новый классификатор. В данном методе добавляется классификатор, который будет обучаться с нуля, поверх предварительно обученной модели.

- Дообучение.

Суть дообучения заключается в размораживании последних слоёв нейронной сети и их обучении. Таким образом корректируются слои, которые имеют наиболее абстрактные представления. Когда производится обучение только нескольких слоёв, уменьшается риск переобучения.

2.3 Сохранение и загрузка модели для дальнейшей работы после трансферного обучения

Файл `.keras` хранит в себе полную информацию о нейронной сети, что позволяет восстановить модель в любом моменте времени без необходимости повторного обучения. Данный файл содержит в себе следующую информацию об архитектуре модели, ее состоянии, команд для компиляции и исторические данные обучения.

3 Практическая часть

3.1 Подготовка данных для обучения

Для обучения модели используется датасет галереи, где содержится 13 различных жанров живописи. Этот набор делится на два каталога `train` (данные для обучения) и `validation` (данные для проверки). Для обучения используется 50% данных от исходного датасета, а для проверки 50%.

В качестве базовой модели используем сеть `MobileNetV3Large`. Данная модель ожидает, что входные значения пикселей будут в диапазоне $[-1, 1]$, но пока значения пикселей в изображениях находятся в диапазоне $[0, 255]$. Это изменяется при загрузке изображения в скрипте.

3.2 Создание базовой модели `MobileNetV3Large` и сравнение её с другими моделями

`MobileNetV3Large` — это предварительно обученная модель на наборе данных `ImageNet`. Данная база знаний помогает нам классифицировать жанры живописи.

Выбор модели для обучения происходил следующим образом: обучение различных моделей проводилось на 10% данных от полного датасета. Далее сравнивался процент точности и скорость обучения.

Модели для сравнения:

- `InceptionV3` — точность обучения 48%;
- `MobileNetV2` — точность обучения 51%;
- `MobileNetV3Large` — точность обучения 56%;
- `ResNet50V2` — точность обучения 69%.

Несмотря на то, что `ResNet50V2` дает большую точность, использовать будем `MobileNetV3Large`, так как данная модель обучается быстрее.

3.3 Обучение модели

Модель будет обучаться на графическом процессоре GPU NVIDIA GeForce RTX 3070 8 GB.

Для обучения модели передаем в нее `train_dataset`, инициализируя количество эпох, равное 1000, и передавая функцию `get_callbacks()`. Данная функция заканчивает обучение сети, если на протяжении 10 эпох точность проверочных данных не возрастает. Это помогает избежать переобучения, из-

за которого сеть может распознавать особенности обучающей выборки, а не данных.

Далее вычислим точность обучения на тестовом наборе. Она составляет 68%.

3.4 Тонкая настройка модели

Одним из способов для повышения производительности и точности обучения является тонкая настройка модели, которая позволяет весам в процессе обучения настраиваться на объекты, связанные с конкретным набором данных.

Для этого надо выполнить разморозку верхних слоев модели, перекомпилировать ее, снизив скорость обучения из-за добавления слоев, и возобновить обучение.

С помощью данной настройки точность обучения на тестовом наборе повысилась и составляет 71%.

3.5 Собственная оценка точности модели

Для повышения точности была написана функция, обрабатывающая предсказанные метки с вероятностями, которые возвращает модель.

Обработка заключается в сортировке вероятностей и взятии двух метки с наибольшими вероятностями. Если же из набора, состоящего из двух меток, какая-либо предсказанная метка совпадает с реальной, то идет в счетчик для расчета оценки точности, иначе – нет.

Оценка точности происходит путем деления количества изображений с правильно предсказанными метками на количество всех изображений в наборе.

Таким образом, получилось повысить оценку точности с 71% до 87%.

4 Экспериментальная часть с созданием веб-приложения

4.1 Streamlit — инструмент для создания веб-приложений

Streamlit представляет собой современный фреймворк для разработки интерактивных веб-приложений на языке Python. Его ключевое преимущество заключается в простоте использования – разработчик может создавать полнофункциональные веб-интерфейсы без необходимости работы с традиционными веб-технологиями, такими как HTML, CSS или JavaScript.

4.2 Создание веб-приложения с помощью Streamlit

Был создан интерфейс загрузки файла и захвата изображения с камеры. На рисунке 1 продемонстрирована работа классификатора изображений с помощью захвата фото с камеры устройства.



Ваше изображение

Основной стиль: marina (94.87%)

Другие возможные стили:

- landscape: 4.55%
- mythological-painting: 0.23%

Рисунок 1 – Распознавание изображения с камеры

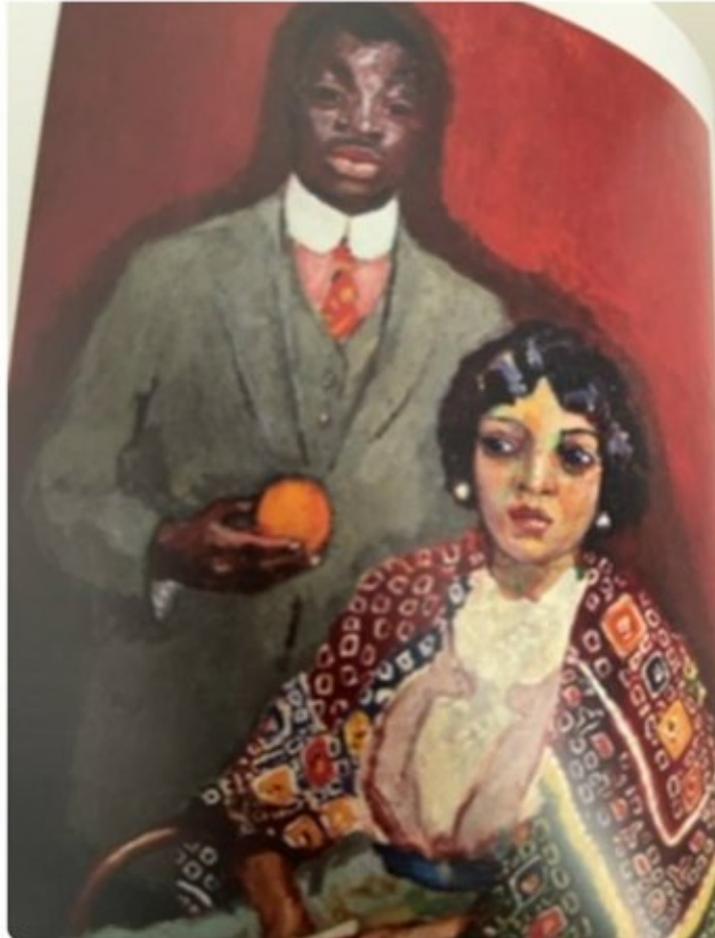
4.3 Размещение приложения в сети Интернет

Для размещения приложения в сети Интернет будет использоваться Docker, как технология контейнеризации приложений.

Был создан Docker-образ, а затем был выбран отечественный хостинг Yandex Cloud для размещения веб-приложения.

В Yandex Cloud создаем виртуальную машину на ОС Ubuntu Linux. Далее требуется регистрация домена. Для этого был выбран отечественный сервис рег.ру (<https://www.reg.ru/>) и создан домен ковалевадипломсгу.ру, который смотрит на IP-адрес виртуальной машины.

На рисунке 2 представлена работа приложения в сети Интернет с мобильного устройства.



Ваше изображение

Основной стиль: portrait (90.74%)

Другие возможные стили:

Navigation bar with address "ковалевадипломсгу.рф" and icons for back, forward, share, book, and tabs.

Рисунок 2 – Распознавание изображения с камеры мобильного устройства

ЗАКЛЮЧЕНИЕ

В данной выпускной квалификационной работе было проведено исследование возможностей применения сверточных нейронных сетей и технологий трансферного обучения для решения актуальной задачи классификации жанров живописи. Разработанное веб-приложение на основе библиотеки TensorFlow и фреймворка Streamlit демонстрирует практическую реализацию современных методов компьютерного зрения в сфере цифровизации изобразительного искусства.

Практическая значимость работы заключается в создании инструмента, который может быть использован.

- В музеях и галереях для автоматической сортировки коллекций
- В образовательных учреждениях для изучения истории искусств
- В цифровых архивах для обработки оцифрованных произведений искусства
- В целях геймификации для курсов по изобразительному искусству

Разработанное решение демонстрирует эффективность сочетания технологий Docker и облачной платформы Yandex Cloud для задач анализа изображений. Также стоит отметить, что веб-приложение доступно в сети Интернет и любой пользователь может попробовать распознавание жанра живописи.