

Пояснительная записка

к дипломной работе на тему:

**"Классификация драгоценных камней"**

Автор: Кормазова Вероника

Группа: DLL-24

# ОГЛАВЛЕНИЕ

[ОГЛАВЛЕНИЕ 2](#_Toc142091317)

[ПОСТАНОВКА ЗАДАЧИ 3](#_Toc142091318)

[АНАЛИЗ ДАННЫХ 5](#_Toc142091319)

[МЕТОДИКА РЕАЛИЗАЦИИ 6](#_Toc142091320)

[ИТОГИ ОБУЧЕНИЯ МОДЕЛИ 9](#_Toc142091321)

[ВЫВОДЫ 12](#_Toc142091322)

# ПОСТАНОВКА ЗАДАЧИ

Оценка ювелирных изделий, в том числе и драгоценных камней, всегда считалась сложной задачей, решаемой на основе предпочтений эксперта. Использование традиционных технологий искусственного интеллекта, таких как экспертные системы, не нашло в геммологии широкого распространения. Поэтому более перспективным представляется подход, основанный на использовании методов распознавания образов, классификации и кластеризации.

В данной работе поставлена и решается задача классификации природных драгоценных камней. Работа направлена на поиск оптимальной модели (моделей) для успешной реализации данной задачи классификации.

В работе применены архитектуры следующих сверточных нейронных сетей: ResNet18, VGG16.

В качестве метрик оценки обучения используются следующие метрики: Оценка F-1 (F-1 Score), Категориальная функция потерь (Categorical Cross-Entropy Loss), Доля правильных ответов (Accuracy).

Оценка F-1 является одним из распространенных показателей оценкиуспешности классификатора. Это среднее гармоническое двух других показателей, а именно: точности и отзыва. В задаче классификации формула равна:



precision

recall

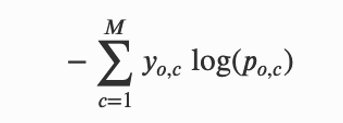
где TP - истинно положительные, FP- ложно положительные, FN - ложно отрицательные исходы.

В качестве другой метрики обучения используем Значение функции потерь (Loss), которое в данной задаче вычисляется с помощью функции Категориальная кросс-энтропия (Categorical Cross-Entropy Loss).

Категориальная кросс-энтропия - это функция потерь, которая используется в задачах многоклассовой классификации. Это задачи, в которых пример может принадлежать только к одной из многих возможных категорий, и модель должна решить, к какой именно.

Формально он предназначен для количественной оценки разницы между двумя распределениями вероятностей.

Если 𝑀**>2** (т. е. многоклассовая классификация), мы вычисляем отдельные потери для каждой метки класса за одно наблюдение и суммируем результат.



где M— количество классов, y — двоичный показатель (0 или 1), если метка класса **c** является правильной классификацией для наблюдения **o,** p— прогнозируемое вероятностное наблюдение **o** относится к классу **𝑐.**

Еще одна метрика обучения, рассмотренная в данной задаче, это Доля правильных ответов (Accuracy). Определяется по формуле:



где TP - истинно положительные, TN - истинно отрицательные, FP- ложно положительные, FN - ложно отрицательные исходы.

# АНАЛИЗ ДАННЫХ

Источник данных

В работе использованы данные с платформы Kaggle[[1]](#footnote-1).

Тип и описание данных

Набор данных состоят из 3200+ изображений различных драгоценных камней. Изображения сгруппированы в 87 классов, которые разбиты на тренировочную и тестовую выборки. Все изображения имеют различные размеры и представлены в формате .jpg. Каждый класс драгоценных камней представлен в различных формах - круглой, овальной, квадратной, прямоугольной и даже в форме сердца.

Пример изображения:

Рис. 1 - Alexandrite Рис. 2 - Coral

Работу с данными начнем с их загрузки и просмотра. Разделим загруженный тренировочный набор данных на собственно тренировочный и валидационный в отношении 0,7 к 0,3 соответственно. Преобразуем изображения в тензор, нормализуем и добавляем аугментацию. Набор данных готов к обучению. 2

# МЕТОДИКА РЕАЛИЗАЦИИ

Проведенные эксперименты

Для решения поставленной задачи было проведено 10 экспериментов (табл. 1), в которых использовались архитектуры нейронных сетей ResNet18[[2]](#footnote-2)3 - 5 экспериментов (рис. 4) и VGG164 - 5 экспериментов (рис. 5). Модели нейросетей ResNet185 и VGG166 рассматривались начиная с предварительно не обученной модели (модель с нуля) до предобученной с добавлением классификационных слоев и планировщика скорости обучения.

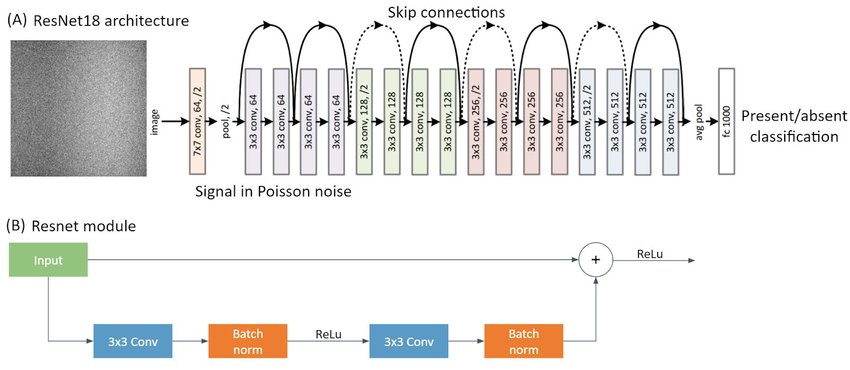


Рис. 4 – Архитектура ResNet18

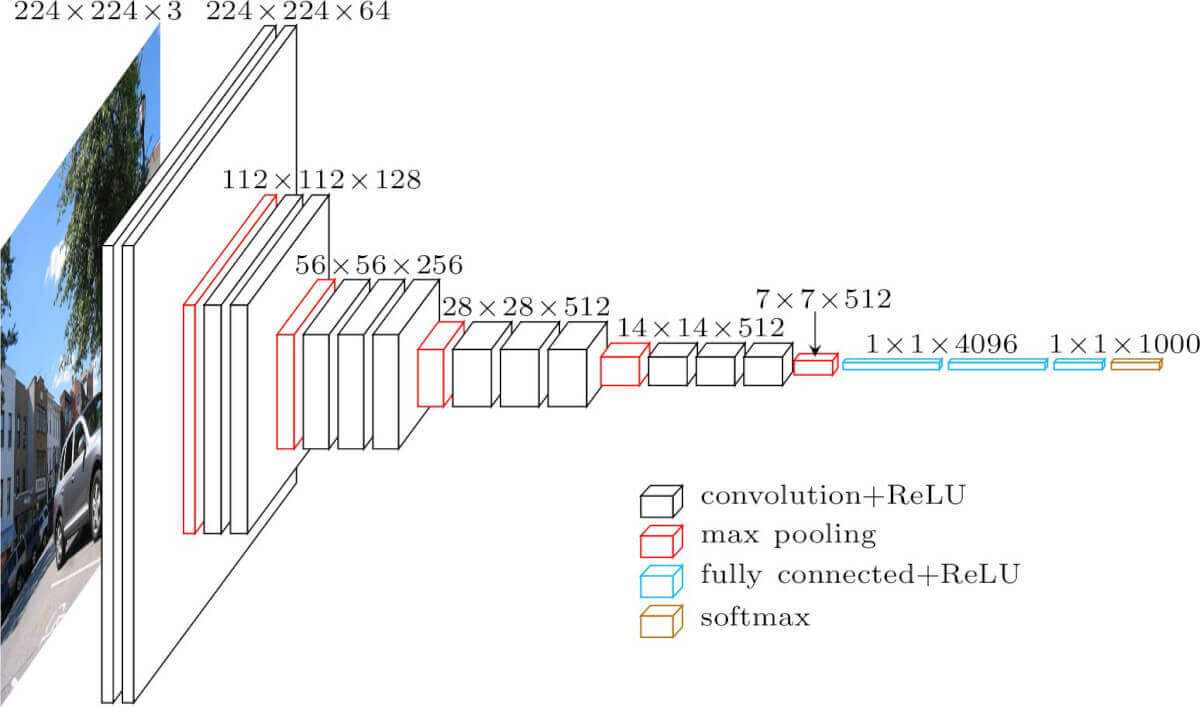


Рис. 5 - Архитектура VGG16

Параметры обучения моделей

ResNet187:

* Оптимайзер – SGD
* LOSS-функция – CrossEntropyLoss
* Функция активации – ReLU
* Аугментация – RandomHorizontalFlip, RandomVerticalFlip, RandomResizedCrop
* Планировщик - StepLR , ExponentialLR

VGG168:

* Оптимайзер – Adam
* LOSS-функция – CrossEntropyLoss
* Функция активации – ReLU
* Аугментация - RandomHorizontalFlip, RandomVerticalFlip, RandomResizedCrop
* Планировщик - StepLR , ExponentialLR

Таблица 1 - Результаты экспериментов

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **№ эксперимента** | **Архитектура** | **Модель** | **Значение функции потерь (на валидационной выборке)**  **Validation loss** | **Доля правильных ответов  (на валидационной выборке)**  **Validation accuracy** |
| 1 | ResNet187 | Предварительно не обучена | 1.46 | 58.79 |
| 2 | Предобученная | 1.22 | 69.94 |
| 3 | Предобученная с добавлением классификационных слоев | 0.93 | 75.48 |
| 4 | Предобученная с добавление планировщика StepLR | 0.98 | 71.97 |
| 5 | Предобученная с добавлением классификационных слоев и планировщика ExponentialLR | 0.72 | 80.59 |
| 6 | VGG168 | Предварительно не обучена | 2.38 | 57.75 |
| 7 | Предобученная | 1.2 | 66.02 |
| 8 | Предобученная с добавлением классификационных слоев | 1.3 | 76.77 |
| 9 | Предобученная с добавлением классификационных слоев планировщика StepLR | 1.14 | 79.63 |
| 10 | Предобученная с добавлением классификационных слоев и планировщика ExponentialLR | 1.09 | 78.98 |

7<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=cCk7DMzk-Uqw> 8<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=kLIPUzAqA5GI>

Оценка успешности классификатора F-1 Score для ResNet187 (5 эксперимент):

F-1 Score=0.63649.

Оценка успешности классификатора F-1 Score для VGG8 (10 эксперимент):

F-1 Score=0.59429.

9<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=MdPd8YCf08XW&line=10&uniqifier=1>

# ИТОГИ ОБУЧЕНИЯ МОДЕЛИ

Полученные результаты

В ходе проведенных экспериментов для каждой тестируемой модели были посчитаны следующие метрики: Значение функции потерь на валидационной выборке (Validation loss), Доля правильных ответов на валидационной выборке (Validation accuracy) и Оценка успешности классификатора (F-1 Score), представлены в таблице 1. Согласно этим метрикам лучшей моделью была признана ResNet1810: Validation loss = 0.72, Validation accuracy = 80.59, F-1 Score = 0,6364. Данная модель была предобучена с добавлением классификационных слоев и планировщика ExponentialLR входе 5-ого эксперимента.

На рисунках 6-8 представлены примеры тестирования модели ResNet1811.

Label: Chrysoberyl

Prediction: Chrysoberyl

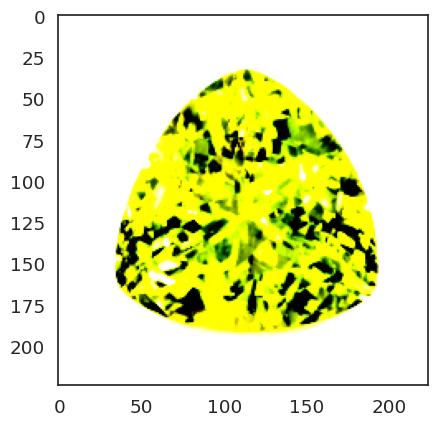


Рис. 6 – Chrysoberyl

10<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=yWL1eGXqAZjP>

11<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=UPRO_6WNXlLX>

Label: Zoisite

Prediction: Zoisite

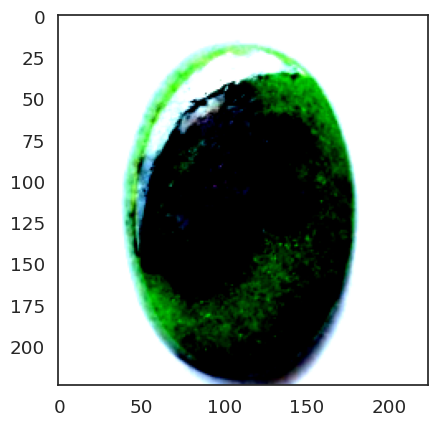


Рис. 7 - Zoisite

Label: Bloodstone

Prediction: Bloodstone

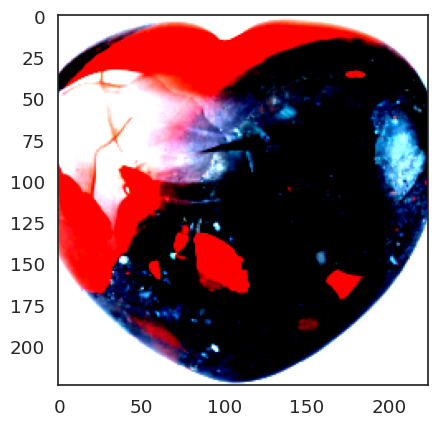


Рис. 8 – Bloodstone

# ВЫВОДЫ

В ходе выполненной работы была выявлена оптимальная модель из всех проведенных экспериментов. Это модель ResNet1812, которая была предобучена с добавлением классификационных слоев и планировщика ExponentialLR. Данная модель имеет следующие целевые метрики:

- значение функции потерь на валидационной выборке - 0.72;

- доля правильных ответов на валидационной выборке – 80.59;

- оценка успешности классификатора F-1 Score – 0.6363.

Сравнивая полученные результаты с моделями, предлагаемыми на платформе Kaggle13, можно утверждать, что представленная в данной работе модель демонстрирует хороший результат (validation accuracy - 80.59, validation loss - 0.72, F-1 score – 0.6363), не все представленные на платформе Kaggle работы достигли вышеприведенных значений метрик.

При совместном использовании ResNet18 и VGG16 качество успешности классификатора увеличивается до F-1 Score до 0,7143.

В дальнейшем данную задачу можно реализовать с помощью применением других архитектур сверточных нейросетей и оценить полученные метрики, например Inception v3(v4) или GoogLeNet.

12<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=p4reOU2N--Ho>

13<https://www.kaggle.com/datasets/lsind18/gemstones-images>

1. <https://www.kaggle.com/datasets/lsind18/gemstones-images>

   2<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=LEhmc-ef7RH9> [↑](#footnote-ref-1)
2. 3<https://d2l.ai/chapter_convolutional-modern/resnet.html>

   4<https://neurohive.io/ru/vidy-nejrosetej/vgg16-model/>

   5<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=cCk7DMzk-Uqw>

   6<https://colab.research.google.com/drive/1kAwhEWwhwvV7WdAxbbOBR73k66f2isgW#scrollTo=kLIPUzAqA5GI> [↑](#footnote-ref-2)