

Определение степени архитектурной ценности здания по фото

Иванова Анна 5030102/00101

Используемые данные

- За основу была взята архитектура Санкт-Петербурга.
- Датасет взят с сайта <https://kontikimaps.ru/how-old/saint-p>
- В датасете имеется более 15000 строк данных о зданиях Санкт-Петербурга. Большая часть данных в датасете ссылается на реестр зданий города, но есть и сведения, добавленные пользователями сайта.
- Из имеющихся данных были взяты URL-адреса фотографий, год постройки здания и архитектурный стиль
- Из-за малого количества объектов некоторых архитектурных стилей, пересечений некоторых классов («Рядовая Застройка» и «Современный»), плохого качества части фотографий, а также отсутствия единого стиля разметки, пришлось провести обработку данных.

Итоговый датасет

После обработки датафрейм содержит 10617 строку. Распределение данных по классам выглядит следующим образом. Как видно из гистограммы, преобладающим классом является «Рядовая застройка»

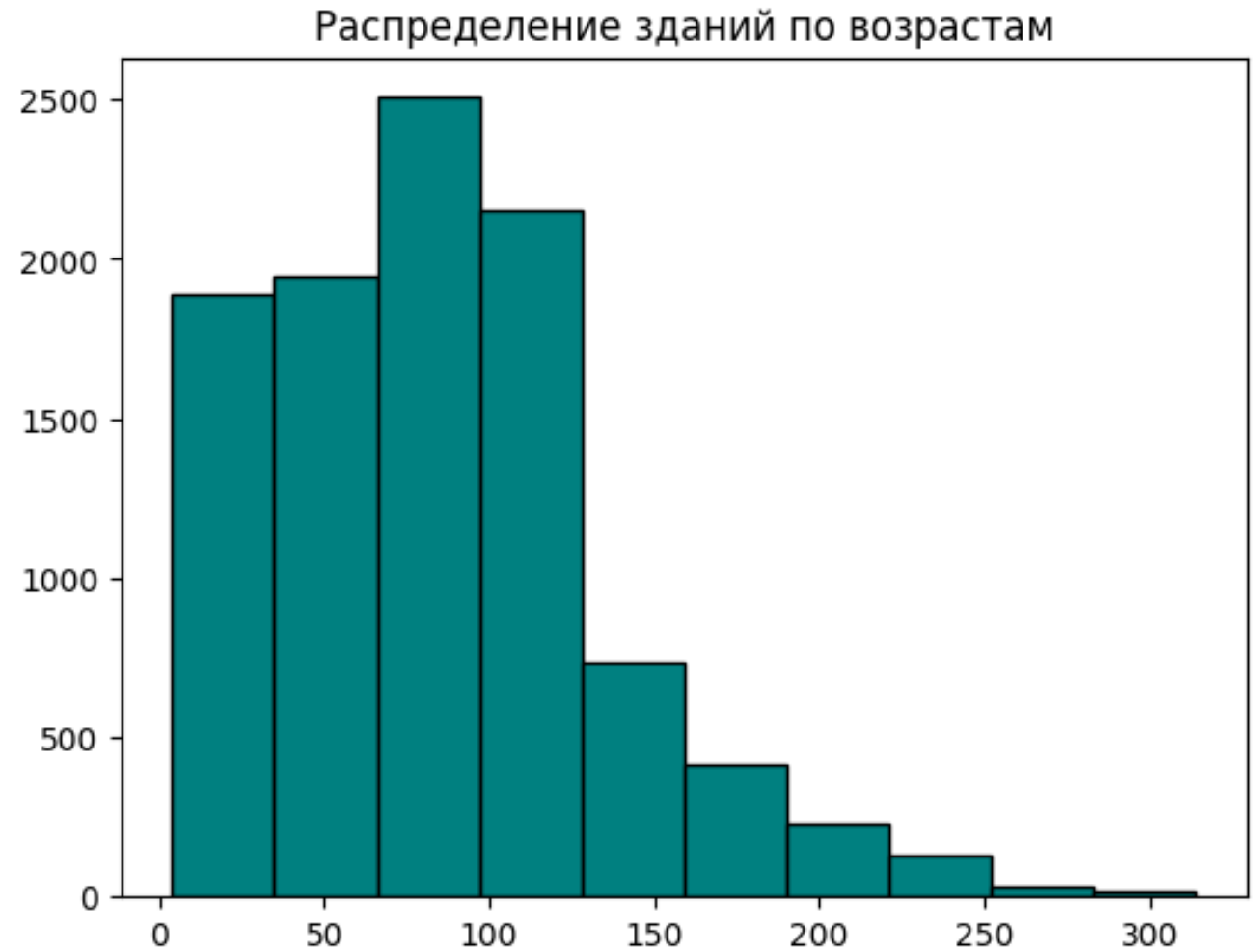


0	Кирпичный
1	Классицизм
2	Конструктивизм
3	Модерн
4	Неоклассицизм
5	Рядовая Застройка
6	Сталинский Неоклассицизм
7	Эклектика

Итоговый датасет

Распределение по возрастам зданий выглядит следующим образом

Как видно из гистограммы, преобладают здания возрастом до ста лет



Описание алгоритма

- Фотографии преобразовываются до размера 150 на 150 пикселей, формат – трехканальный RGB
- Выборка перемешивается и делится на тестовую и обучающую (размер тестовой выборки — 20% от общего объема данных)
- Для написания и обучения нейросети использовались средства библиотеки Keras

Модели различаются между собой целевыми функциями и метриками

Модель для определения архитектурного стиля

- Целевая функция **sparse_categorical_crossentropy**
- Метрика **accuracy** (измеряет количество верно классифицированных объектов относительно общего количества всех объектов)

Модель для определения возраста

- Целевая функция **mean_squared_error** (средне-квадратичная ошибка)
- Метрика **mae** (измеряет среднюю сумму абсолютной разницы между фактическим значением и прогнозируемым значением)

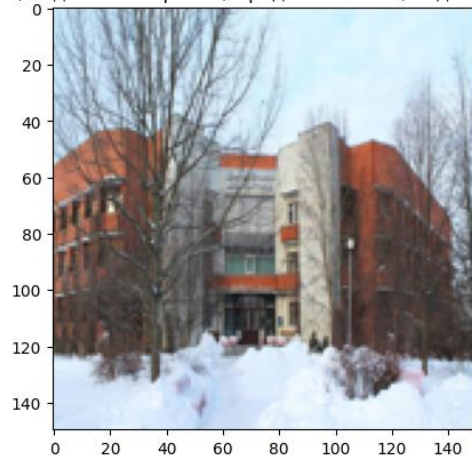
Результаты обучения нейросетей

Для задачи классификации зданий по архитектурному стилю была достигнута точность 94% на тренировочной и 62% на тестовой выборке

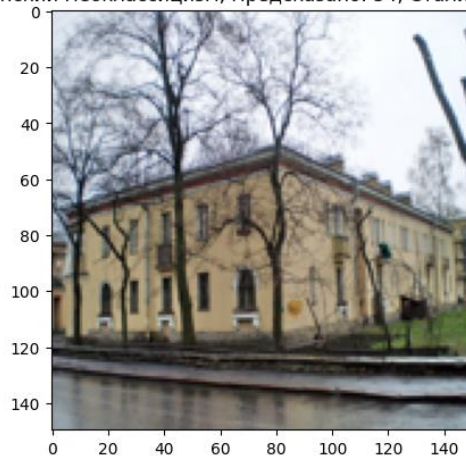
- Для задачи определения возраста ошибка составляет в среднем 23 года на тренировочной и 30 лет на тестовой выборке

«Хорошие» примеры результатов обучения

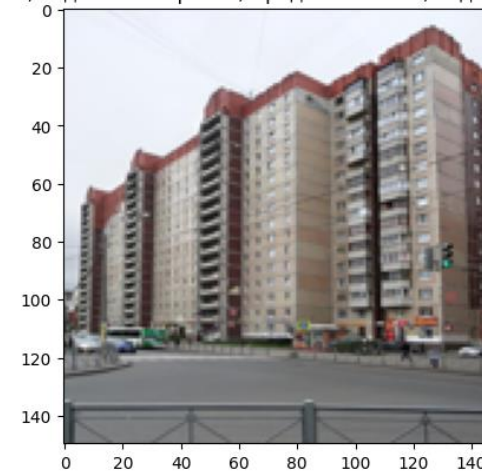
Точно: 38, Рядовая Застройка, Предсказано: 36, Рядовая Застройка



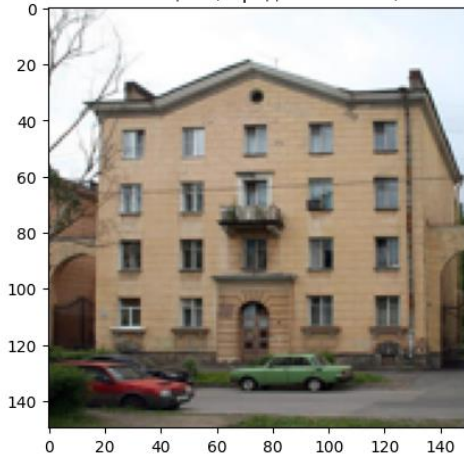
Точно: 76, Сталинский Неоклассицизм, Предсказано: 54, Сталинский Неоклассицизм



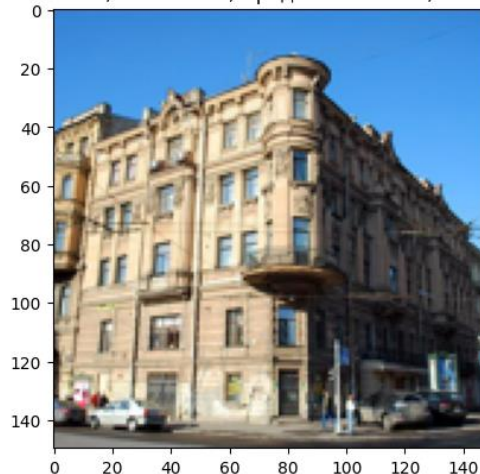
Точно: 28, Рядовая Застройка, Предсказано: 24, Рядовая Застройка



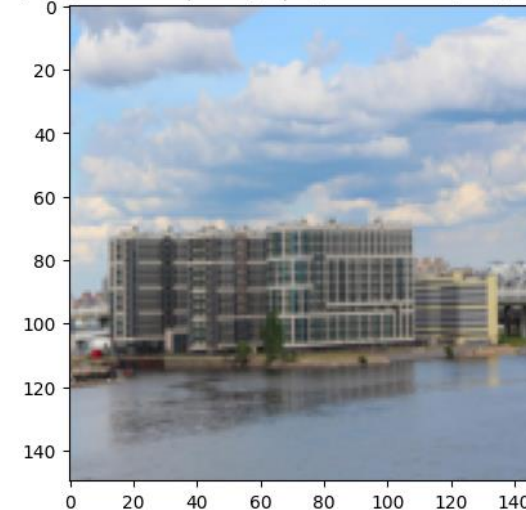
Точно: 74, Сталинский Неоклассицизм, Предсказано: 89, Сталинский Неоклассицизм



Точно: 154, Эkleктика, Предсказано: 114, Эkleктика

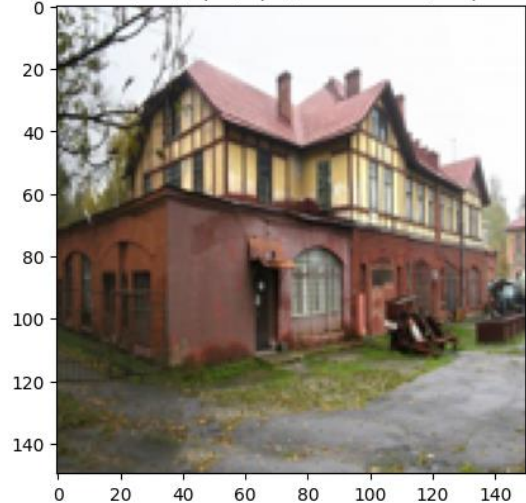


Точно: 7, Рядовая Застройка, Предсказано: 13, Рядовая Застройка

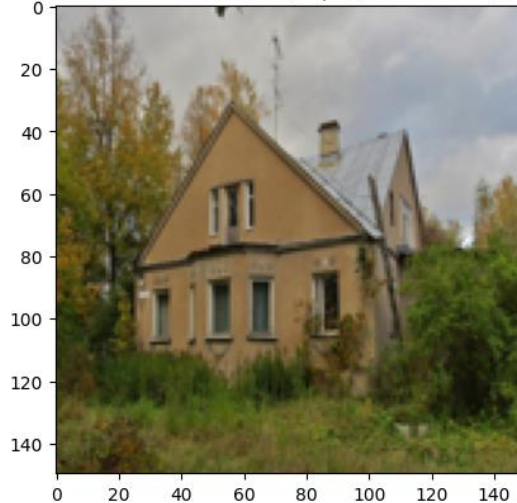


«Плохие» примеры результатов обучения

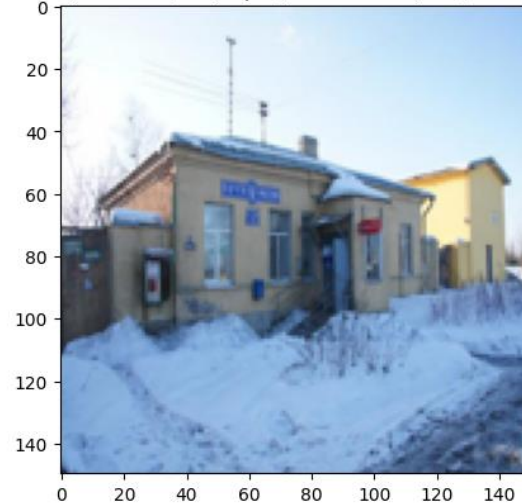
Точно: 117, Модерн, Предсказано: 73, Кирпичный



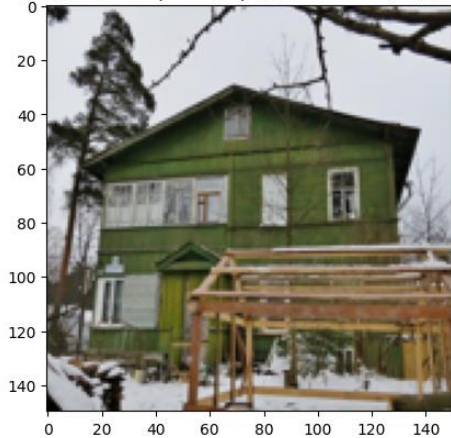
Точно: 64, Сталинский Неоклассицизм, Предсказано: 53, Рядовая Застройка



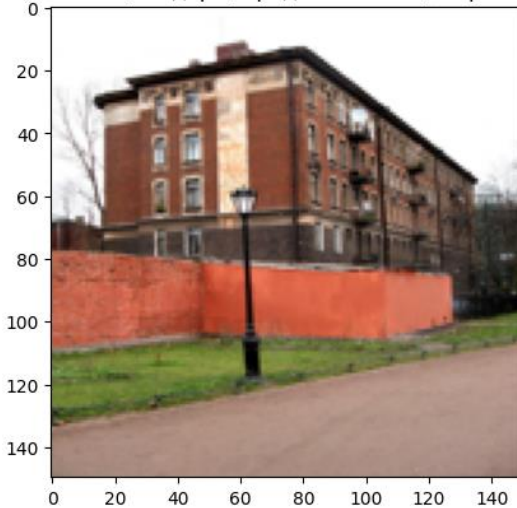
Точно: 199, Классицизм, Предсказано: 32, Рядовая Застройка



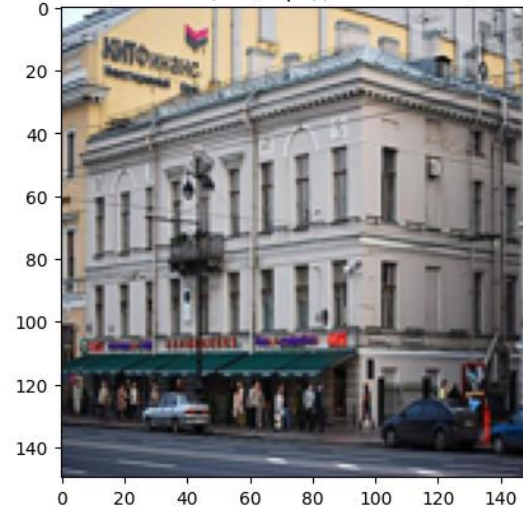
Точно: 123, Рядовая Застройка, Предсказано: 66, Рядовая Застройка



Точно: 120, Модерн, Предсказано: 80, Кирпичный



Точно: 226, Классицизм, Предсказано: 107, Эkleктика



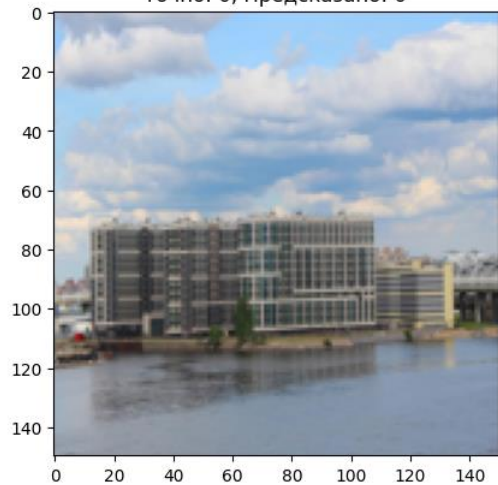
Определение степени архитектурной ценности

На основе некоторых критериев государственных ведомств и субъективных оценок было реализован алгоритм определения архитектурной ценности на основе возраста здания и архитектурного стиля. Некоторые общие критерии:

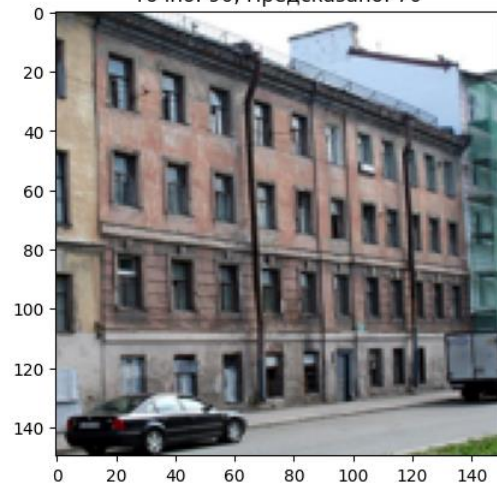
- Чем больше возраст здания — тем выше оценка.
- Оценка архитектурного стиля зависит от наличия/отсутствия «архитектурных излишеств»
- Оценка варьируется от 0 до 100

Некоторые результаты

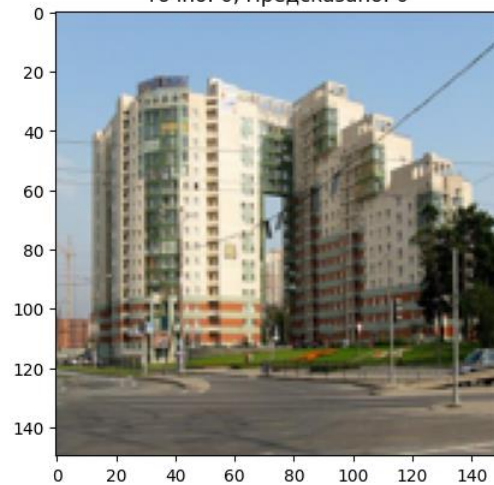
Точно: 0, Предсказано: 0



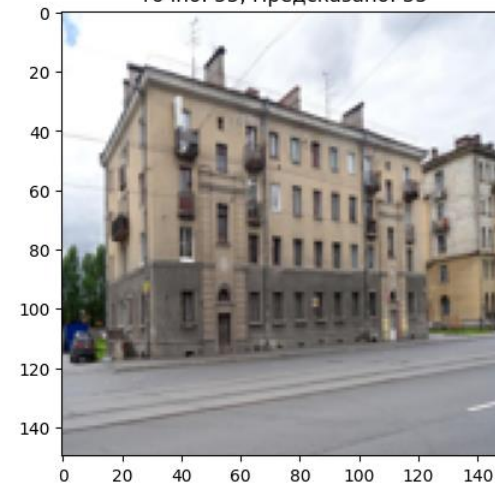
Точно: 90, Предсказано: 70



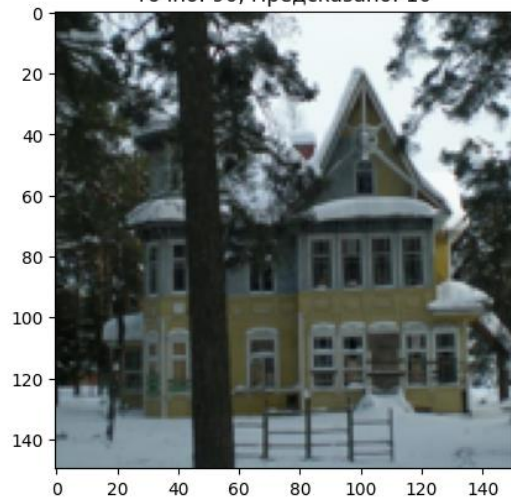
Точно: 0, Предсказано: 0



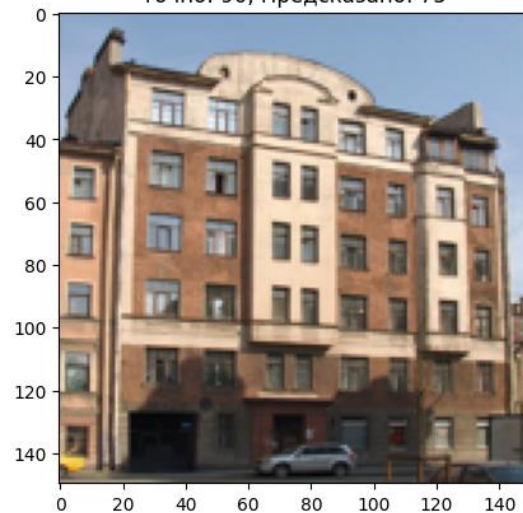
Точно: 55, Предсказано: 55



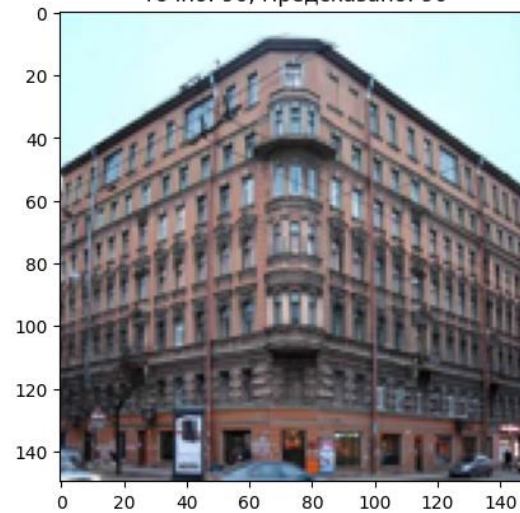
Точно: 90, Предсказано: 10



Точно: 90, Предсказано: 75



Точно: 90, Предсказано: 90



Точно: 50, Предсказано: 70

