

Подсистема «Нейросетевой блок распознавания фотоизображений объектов»

Проведение анкетирования

Перед разработкой нейронной сети было проведено анкетирование, в котором приняло участие 138 респондентов, из которых 57 мужчин и 81 женщина. Основные возрастные категории, которые приняли участие в анкетировании, это люди от 18 до 24 лет и от 50. Результаты представлены на рисунках 26 и 27.

Укажите ваш пол
138 ответов

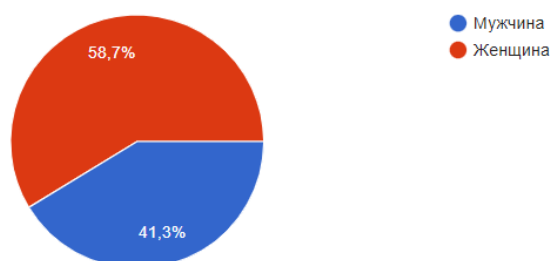


Рисунок 26 – Распределение пола респондентов

Укажите вашу возрастную категорию
138 ответов

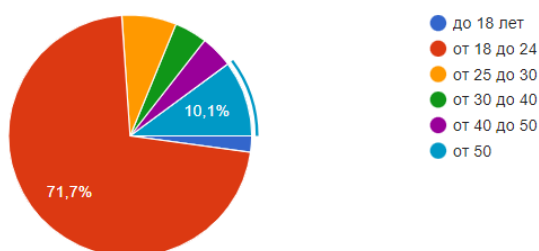


Рисунок 27 – Возрастные категории респондентов

Также, анкетлируемые написали наборы техники и мебели, которая должна быть в арендуемом объекте недвижимости (рисунок 28). Среди ответов были выбраны наиболее часто встречающиеся предметы интерьера и техники (рисунок 29), которые были выбраны для дальнейшего обучения нейросети.

Напишите набор техники и мебели, которая по вашему мнению должна быть в арендуемом(ой) доме/квартире:

138 ответов

Ванна, двуспальная кровать, стол для работы/учёбы, стиральная машинка, плита, кондиционер, холодильник

Холодильник, стиральная машина, микроволновая печь, плита, утюг, гладильная доска, телевизор, кровать , стол , стулья, шкаф для одежды.

Диван/кровать, бытовая техника (чайник, духовка, микроволновка,плита, вытяжка), кухонный гарнитур, место для хранения (шкафы, полки, комод и др.), рабочий стол, телевизор

Рисунок 28 – Варианты ответов анкетированных



Рисунок 29 – Наиболее востребованные предметы интерьера и техники

Блок распознавания фотоизображений объектов интерьера и бытовой техники дополнит данные из объявлений, путем анализа графической информации (фотоизображений), определяя наличие и состав бытовой техники и мебели на фотоизображениях. Полученные данные используются для поиска объявлений с наличием тех или иных предметов интерьера и бытовой техники, необходимых арендатору.

Разработка собственной нейронной сети

Для распознавания были определены следующие объекты интерьера: холодильник, диван, кровать, бойлер, ванна, душевая кабина, электрическая плита двух видов, газовая плит, духовка, телевизор.

Это было сделано для того, чтобы в дальнейшем выбирать на фото только актуальные предметы, которые чаще всего хотелось бы видеть арендаторам в российских квартирах.

Чтобы добиться планируемого результата были поставлены следующие задачи:

Произвести разметку фотографий полученных после парсинга.

Разработать программу для вырезания размеченных объектов на фотографиях.

Разработать, обучить и протестировать нейронную сеть на полученной коллекции изображений.

Разметка фотографий

Для разметки объектов на изображениях была выбрана программа VGG Annotator (инструмент маркировки изображения) позволяющая обработать изображения и выделить все предметы для дальнейшего преобразования [10]. Пример разметки представлен на рисунке 31. Разметка предметов была сохранена в json-файл, в котором находилось название изображений и координаты выделенных предметов.

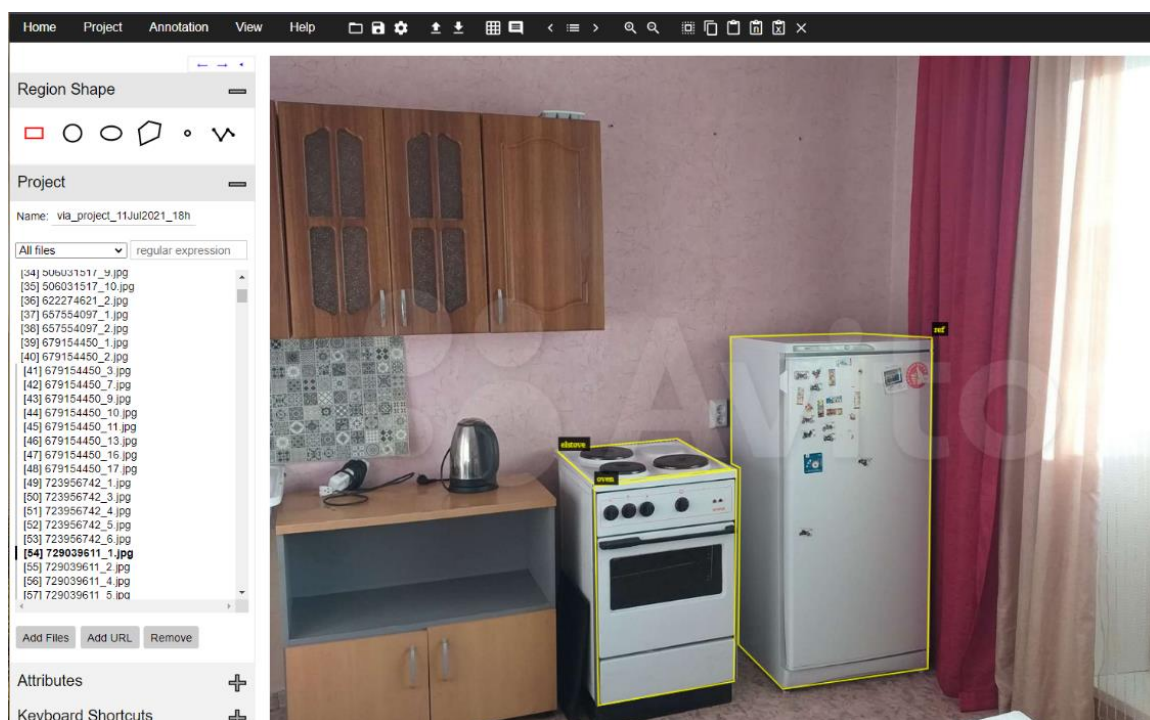


Рисунок 31 – Пример размеченного изображения

Получение размеченных объектов на фотографиях

Парсером было получено 4000 фотографий из объявлений по аренде недвижимости. Из общей коллекции выделено 13 классов, в которых по 100 фотографий для равномерного обучения. Таким образом, нейронную сеть обучили и протестировали на 1300 объектов. Примеры изображений 13-ти классов представлены на рисунке 32.



Рисунок 32 – Пример изображений объектов из датасета

Так как одна фотография могла содержать несколько объектов, а для обучения нейронной сети это не подходило, было принято решение вырезать каждый объект в соответствии с его координатами. Для этого были созданы программы для обработки изображений. Для обработки изображений были использованы библиотеки для компьютерного зрения, машинного обучения, работы с растровой графикой и визуализации данных: OpenCV, Matplotlib, Python Imaging Library, NumPy.

Работа программы начинается с запуска json-файла, из которого берется вся необходимая информация об объектах на фото, то есть происходит декодирование для получения наименований фото, выделенных регионов и их названий, а также координаты объектов,

которые находятся на фото. После этого происходит выделение рамок для будущих объектов в соответствии с их размерами.

Разработка, обучение и тестирование нейронной сети

Перед обучением были сформированы обучающие и тестовые выборки из имеющихся коллекций данных. Тестовые выборки составили 20% от датасета. Далее все фотографии из коллекций подверглись нормализации, при которой значения признаков во входном векторе приводятся к некоторому заданному диапазону $[0,1]$. Если не провести нормализацию, то входные данные будут оказывать дополнительное влияние на нейроны, что приведет к неверным решениям.

Когда имеется небольшое количество обучающих примеров, модель иногда учится на шуме или нежелательных деталях обучающих примеров - до такой степени, что это отрицательно влияет на производительность модели на новых примерах. Для борьбы с переобучением искусственно увеличили количество данных.

После того, как были подготовлены данные, создаем функцию, которая отвечает за обучение модели: считает частные производные, градиенты, потерю, точность, корректирует веса, формирует статистику.

Во время компиляции модели были использованы метод стохастической оптимизации Adam и кросс-энтропийная функция потерь [11]. Размер батча 20 элементов, так как должен быть кратен значению обучаемой выборки, скорость обучения равна 0,001, обучение нейросети проходило на 20 и 25 эпохах с разными архитектурами, для сравнения точности обучения.

Для дальнейшей работы была выбрана модель ResNet-50, которая прошла обучение на 25 эпохах, так как показала лучший результат по точности на обучающей и тестовой наборах данных.

Выбор архитектуры обоснован в таблицах 2 и 3. График обучения для модели ResNet-50 представлен на рисунке 33.

Таблица 2 – Результаты обучения сетей на 20 эпохах

Тип выборки	ResNet-18	ResNet-34	ResNet-50	ResNet-101	ResNet-152
Обучающая выборка, %	97,31	97,50	98,27	98,75	98,85
Тестовая выборка, %	92,69	93,08	93,85	92,31	91,92

Таблица 3 – Результаты обучения сетей на 25 эпохах

Тип выборки	ResNet-18	ResNet-34	ResNet-50	ResNet-101	ResNet-152
Обучающая выборка, %	97,50	97,98	99,13	99,04	98,75
Тестовая выборка, %	93,46	91,92	94,23	92,69	93,85

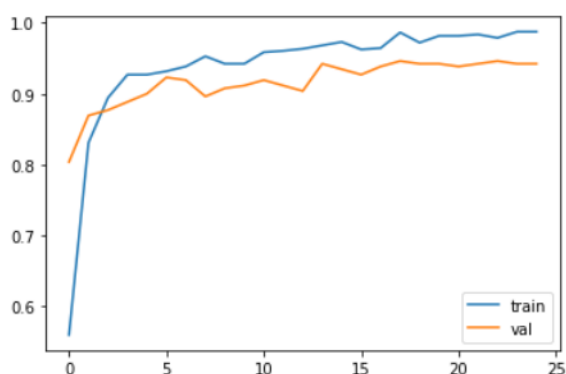


Рисунок 33– График точности обучения и проверки модели ResNet-50

На данный момент в новой нейронной есть все классы, которые необходимы арендаторам. Несмотря на эти преимущества, наша нейронная сеть будет дорабатываться, так как имеет множество перспектив развития. Такие как определение вида за окном или оценка ремонта в квартире.