

## **2.1. Лабораторная работа №1 «Создание датасета и аннотирование изображений»**

### **2.1.1. Цель и задания лабораторной работы № 1**

В ходе первой лабораторно работы рассматриваются особенности создания датасетов, используемых в задачах компьютерного зрения. Аннотации или разметка изображений — это первый шаг в создании алгоритмов распознавания изображений и моделей глубокого обучения. В качестве инструмента создания и разметки датасета могут быть использованы различные фреймворки по предпочтению студента (например, MatLab 2020a/2023b, OpenCV, SuperVisely, CVAT, графические редакторы и др.)

**Цель работы лабораторной работы № 1:** Создание датасетов двумерных растровых изображений и их разметка для решения задач компьютерного зрения

#### **Задание к лабораторной работе №1:**

- 1) Изучить теоретический материал, представленный в разделе 1 данного лабораторного практикума, а также методические рекомендации к данной лабораторной работе.
- 2) Развернуть фреймворк (по собственному выбору) для создания и разметки (аннотирования) датасетов. При выборе фреймворка следует учитывать имеющиеся в наличии вычислительные мощности, а также особенности изображений в соответствии с индивидуальным вариантом заданий.
- 3) Создать (пополнить имеющийся) датасет в соответствии с индивидуальным заданием (табл. 2.1) объемом не менее 100 изображений для каждого класса (например, для рукописных цифр общий объем минимального датасета должен составлять 100 изображений для каждой цифры \* 10 цифр = 1000 объектов).
- 4) Выполнить разметку созданного (дополненного) датасета для решения задачи многоклассовой классификации.
- 5) При необходимости использовать аугментацию данных.

- 6) Создать таблицу с характеристиками итогового датасета (количество объектов датасета каждого класса, количество классов, разрешение изображений, формат файла изображения, способ аннотации, формат файла аннотации)
- 7) Ознакомиться и реализовать удаленный доступ к одному из стандартных датасетов по распознаванию лиц, упомянутым в методических рекомендациях к данной лабораторной работе.
- 8) Добавить в стандартный датасет собственное изображение лица.
- 9) Аннотировать стандартный датасет для решения задачи обнаружения объектов.
- 10) Уметь продемонстрировать удаленный доступ к выбранному стандартному датасету и навигацию по нему с просмотром аннотаций при защите лабораторной работы.
- 11) Создать таблицу с характеристиками выбранного стандартного датасета.
- 12) Оформить отчет по лабораторной работе.

Таблица 2.1.

Варианты индивидуальных заданий к лабораторной работе № 1

№ варианта	Объект датасета (элементарное изображение)	Пример объекта датасета	Разрешение изображения (в пикселях)	Формат файла с объектом датасета
1	Рукописные арабские цифры		28x28	.bmp
2	Рукописные буквы русского алфавита		32x32	.bmp
3	Рукописные буквы английского алфавита		32x32	.bmp
4	Рукописные буквы латинского алфавита		32x32	.bmp
5	Римские цифры		28x28	.bmp
6	Рукописные смайлики		32x32	.bmp
7	Рукописные китайские иероглифы		32x32	.bmp

8	Изображения флагов стран		32x32	.bmp
9	Рукописные корейские иероглифы		32x32	.bmp
10	Рукописные знаки зодиака		32x32	.bmp
11	Логотипы web-браузеров		64x64	.png
12	Логотипы мессенджеров		64x64	.png
13	Логотипы производителей смартфонов		64x64	.png
14	Логотипы графических редакторов		64x64	.png
15	Логотипы легковых автомобилей		64x64	.png

### 2.1.2. Методические рекомендации к выполнению лабораторной работы №1

Аннотация изображений — это процесс маркировки изображений набора данных для обучения модели машинного обучения. Задача аннотации обычно включает ручную работу, иногда с помощью графических редакторов. Инженер по машинному обучению предварительно определяет метки, известные как «классы», и предоставляет информацию, относящуюся к изображению, в модель компьютерного зрения. После того, как модель обучена и развернута, она будет предсказывать и распознавать заранее определенные функции в новых изображениях, которые еще не были аннотированы.

Популярными наборами данных аннотированных изображений являются набор данных Microsoft COCO (Common Objects in Context) с 2,5 миллионами помеченных экземпляров в 328 тыс. изображений и набор данных OID (Open Images Database) Google с примерно 9 миллионами предварительно аннотированных изображений.

Поиск датасета (как аннотированного, так и неаннотированного) с соответствующим контекстом удобно производить с помощью ресурса Dataset Search [8].

Аннотирование видео основано на концепции аннотирования изображения. Для видеоаннотации функции вручную помечаются на каждом видеокadre (изображении). Следовательно, набор данных для модели обнаружения видео состоит из изображений для отдельных видеокadров.

Для разметки изображений можно использовать любой открытый или бесплатный инструмент аннотирования данных. Инструмент аннотирования компьютерного зрения (CVAT), вероятно, является самым популярным инструментом аннотирования изображений с открытым исходным кодом.

При работе с большим объемом данных для аннотирования изображений потребуется обученный персонал. Компании используют своих собственных специалистов по данным для маркировки изображений, но более сложные, реальные проекты часто требуют найма поставщика услуг аннотирования видео с помощью ИИ. Инструменты аннотирования предлагают различные наборы функций для эффективного аннотирования одного или нескольких кадров. Метки применяются к объектам с использованием любого из методов аннотирования, описанных ниже в изображении; количество меток на каждом изображении может варьироваться в зависимости от варианта использования.

Аннотация изображения включает следующие шаги:

Шаг № 1: Подготовка набора данных изображений.

Шаг № 2: Присвоение метки классов объектов для обнаружения.

Шаг № 3: На каждом изображении обведение рамкой (прямоугольной или растровой маской) объекта, который вы хотите обнаружить.

Шаг № 4: Выбор метки класса для каждой нарисованной рамки.

Шаг № 5: Экспорт аннотации в нужном формате (COCO JSON, YOLO и т. д.)

Бесплатные инструменты для аннотирования:

**Supervisely** (<https://supervisely.com/>) – платформа для компьютерного зрения, которая предоставляет полный конвейер обработки данных. Важным преимуществом является то, что экземпляр Supervisely можно развернуть автономно в клиентской инфраструктуре и распределить по разным серверам. Это помогает безопасно обрабатывать конфиденциальные данные. Имеет мощные инструменты маркировки, дополненные поддержкой искусственного интеллекта и пользовательскими интерфейсами. Процесс маркировки интуитивно понятен.

**Makingense.ai** (<https://www.makesense.ai/>) — бесплатный онлайн-инструмент для маркировки изображений, не требующий установки программного обеспечения. Можно использовать его с браузером. Он построен на движке TensorFlow.js, одном из самых популярных фреймворков для обучения нейронных сетей. Хотя этот инструмент предоставляет базовые функции, к которым легко получить доступ, он представляет собой хорошую альтернативу для быстрого тестирования аннотаций изображений и небольших проектов. Установка не требуется; инструмент полностью работает онлайн. Makesense.ai поддерживает несколько форм аннотаций.

**CVAT** (<https://viso.ai/computer-vision/cvat-computer-vision-annotation-tool/>) - разработанный исследователями Intel, является популярным инструментом с открытым исходным кодом для аннотирования изображений. Этот инструмент аннотации требует ручной установки, так как он основан на Github. После настройки он предоставляет больше инструментов и функций, чем другие, например, ярлыки и средство создания формы метки. CVAT поддерживает такие надстройки, как TensorFlow Object Detection и Deep Learning Deployment Toolkit. Платформа приложений компьютерного зрения Viso Suite включает версию CVAT для бизнеса.

**LabelImg** (<https://viso.ai/computer-vision/labelimg-for-image-annotation/>) — это популярный (написанный на Python) простой графический инструмент для аннотирования изображений. Установка относительно проста и обычно выполняется через командную строку/терминал. Инструмент аннотирования

изображений отлично подходит для наборов данных до 10 000 изображений, поскольку он требует много ручного взаимодействия и создан, чтобы помочь аннотировать наборы данных для моделей обнаружения объектов. Простой интерфейс упрощает использование, что делает его хорошим инструментом для начинающих программистов машинного обучения с множеством хорошо документированных руководств.

Цели аннотирования изображений (классификация изображений, обнаружение объектов и т. д.) требуют различных методов аннотирования изображений для разработки эффективных наборов данных.

1) Для *классификации* изображений требуется, чтобы изображения имели одну метку для идентификации всего изображения. Процесс аннотирования изображений для моделей классификации изображений направлен на распознавание наличия похожих объектов на изображениях набора данных. Он используется для обучения модели ИИ идентификации объекта на немаркированном изображении, который похож на классы на аннотированных изображениях, которые использовались для обучения модели. Обучающие изображения для классификации изображений также называются тегами. Таким образом, классификация изображений направлена на то, чтобы просто идентифицировать наличие определенного объекта и назвать его предопределенный класс.

2) *Обнаружение объектов и распознавание объектов.* Для этого типа модели процесс аннотации изображения требует, чтобы вокруг каждого обнаруженного объекта на каждом изображении были проведены границы, что позволяет определить точное положение и количество объектов, присутствующих на изображении. Следовательно, основное отличие состоит в том, что классы обнаруживаются внутри изображения, а не все изображение классифицируется как один класс (классификация изображений). Расположение класса является параметром в дополнение к классу, тогда как при классификации изображений расположение класса в изображении не имеет значения, поскольку все изображение идентифицируется как один

класс. Объекты могут быть аннотированы на изображении с помощью меток, таких как ограничивающие рамки или многоугольники.

3) *Сегментация изображения* Сегментация изображения — это тип аннотации изображения, который включает разделение изображения на несколько сегментов. Сегментация изображения используется для обнаружения объектов и границ (линий, кривых и т. д.) на изображениях. Она выполняется на уровне пикселей, присваивая каждый пиксель изображения определенному объекту или классу. Обычно сегментация используется для проектов, требующих более высокой точности классификации входных данных. Сегментация изображения далее делится на следующие три класса: *Семантическая сегментация* изображает границы между подобными объектами (рис. 2.1). Этот метод используется, когда требуется высокая точность в отношении наличия, местоположения, размера или формы объектов на изображении. *Сегментация экземпляра* определяет наличие, расположение, количество, размер или форму объектов на изображении. Таким образом, сегментация экземпляров помогает маркировать присутствие каждого отдельного объекта в изображении. *Паноптическая сегментация* сочетает в себе как семантическую, так и сегментацию экземпляров. Соответственно, паноптическая сегментация предоставляет данные, помеченные для фона (семантическая сегментация) и объекта (сегментация экземпляра) в изображении.

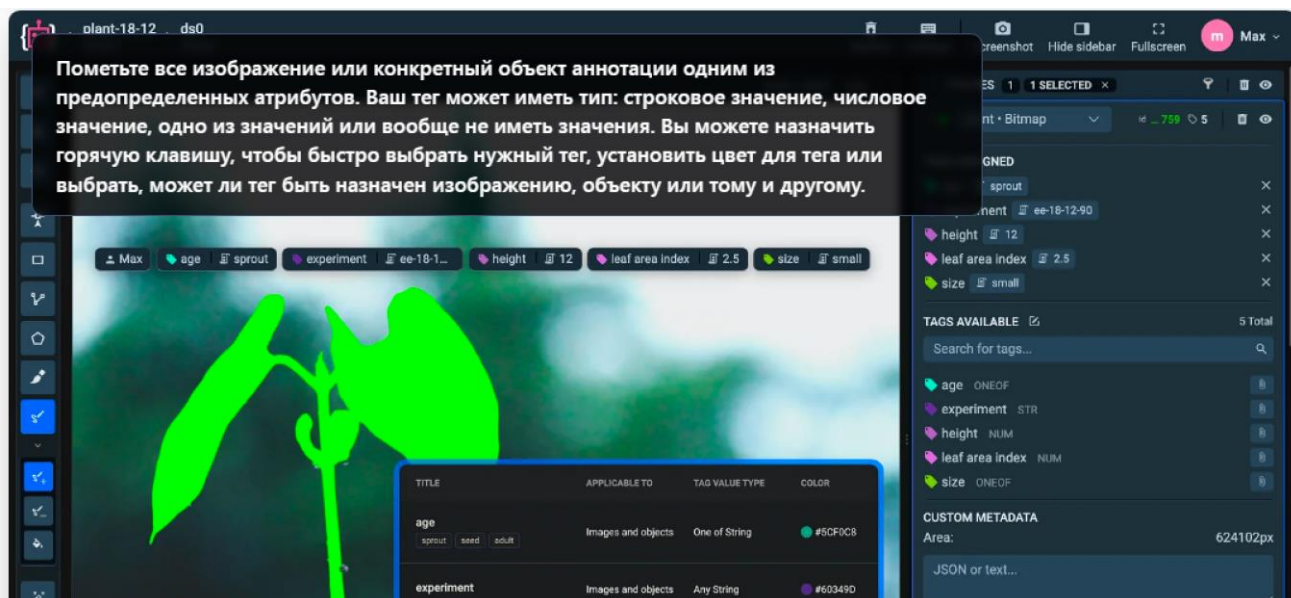


Рис. 2.1 – Интерфейс Supervisely в режиме семантической разметки

4) *Граничное распознавание.* Этот тип аннотации изображения идентифицирует линии или границы объектов на изображении. Границы могут включать в себя края конкретного объекта или области топографии, представленные на изображении. После того, как изображение правильно аннотировано, его можно использовать для выявления похожих шаблонов в изображениях без аннотаций. Распознавание границ играет важную роль в безопасной эксплуатации беспилотных автомобилей

В аннотации изображения используются различные *формы аннотаций* для аннотирования изображения на основе выбранного метода.

В дополнение к фигурам для аннотирования изображений также можно использовать такие средства аннотаций (рис. 2.2), как линии, сплайны и ориентиры. Ниже приведены популярные методы аннотирования изображений, которые используются в зависимости от варианта использования.

1) *Ограничивающие рамки* Ограничивающая рамка — наиболее часто используемая форма аннотации в компьютерном зрении. Ограничивающие рамки — это прямоугольные рамки, используемые для определения местоположения объекта на изображении. Они могут быть двухмерными (2D) или трехмерными (3D).



2) *Полигоны*. Полигоны используются для аннотирования «неправильных» объектов на изображении. Они используются для обозначения каждой из вершин предполагаемого объекта и аннотирования его краев.

3) *Ориентир* используется для определения основных точек интереса на изображении. Такие точки называются ориентирами или ключевыми точками. Ориентиры важны для распознавания лиц.

4) *Линии и сплайны*. Линии и сплайны аннотируют изображение прямыми или изогнутыми линиями. Это важно для распознавания границ, чтобы аннотировать тротуары, дорожные знаки и другие указатели границ.

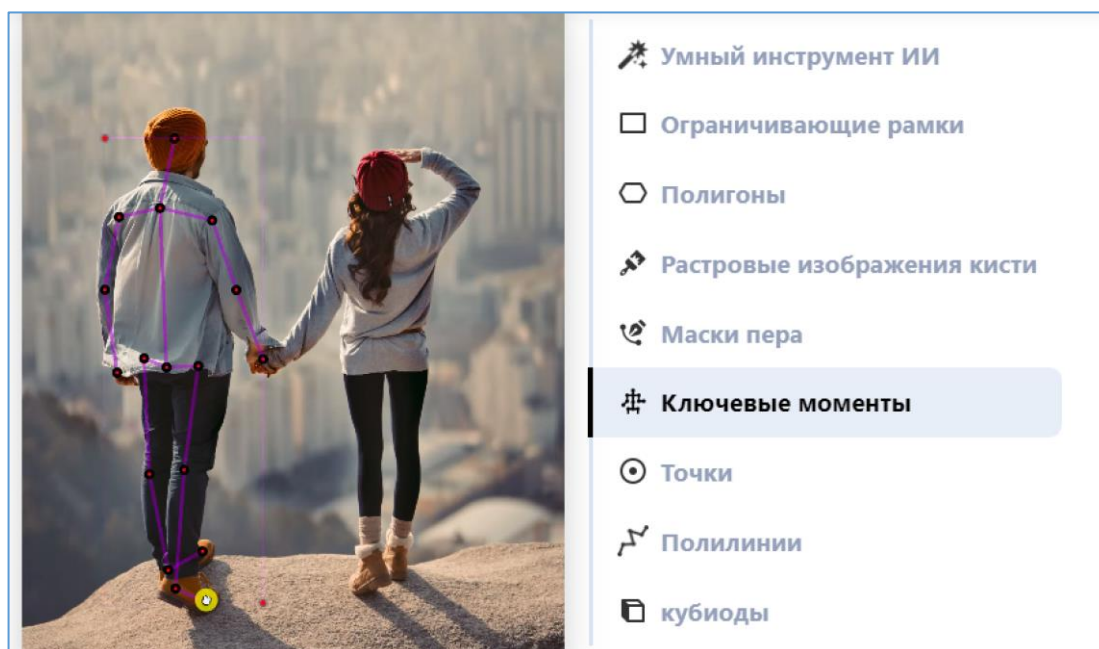


Рис. 2.2 – Средства аннотирования в Supervisely

Помимо обычных средств, таких как прямоугольник или кисть, Supervisely поставляется с «умными» инструментами маркировки, основанными на наборе нейронных сетей, не зависящих от класса.