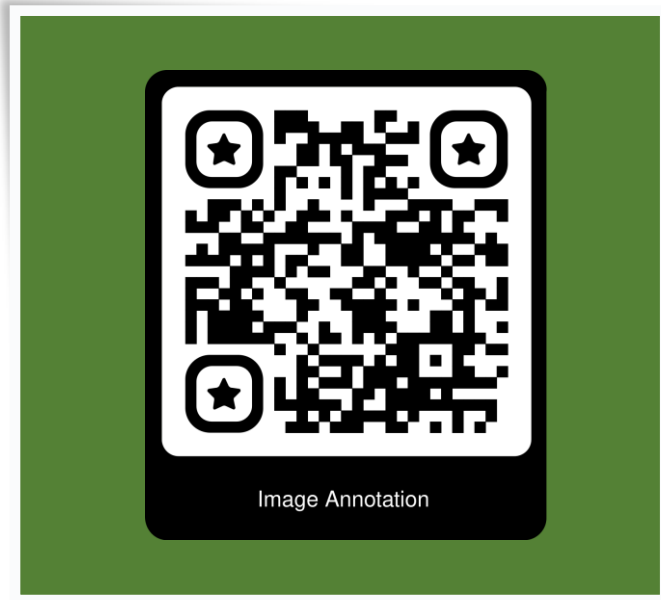




Image Annotation and Labeling

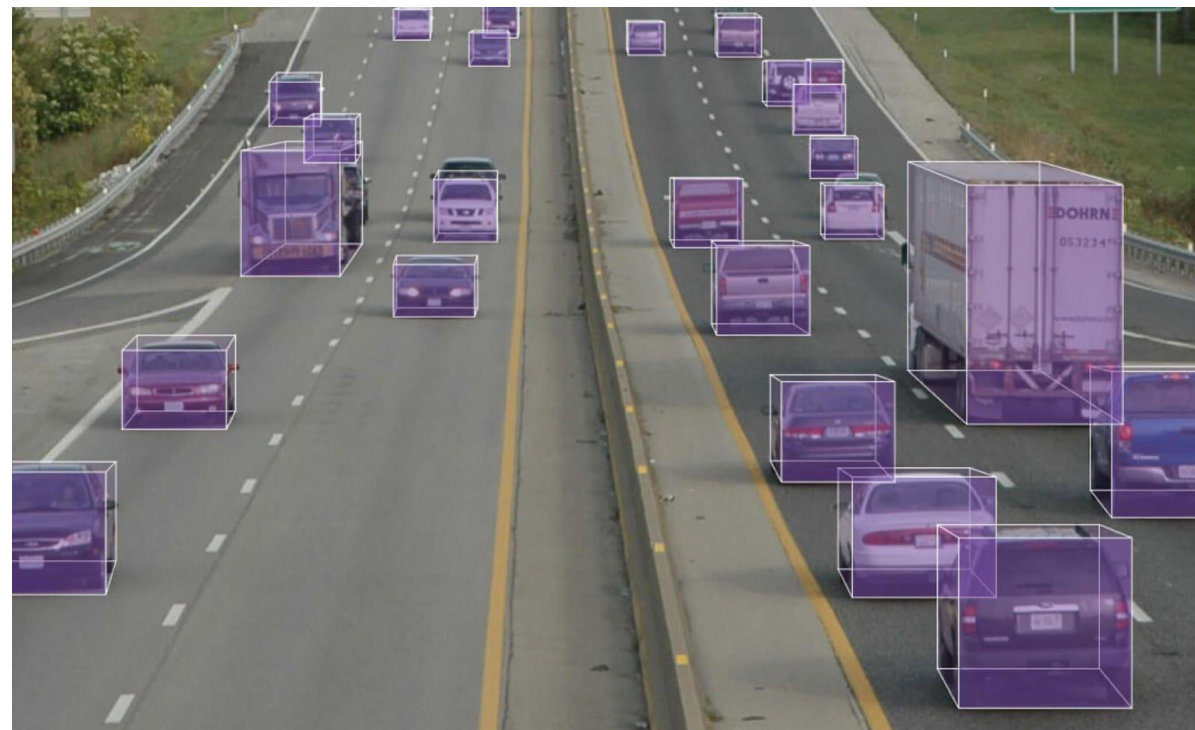
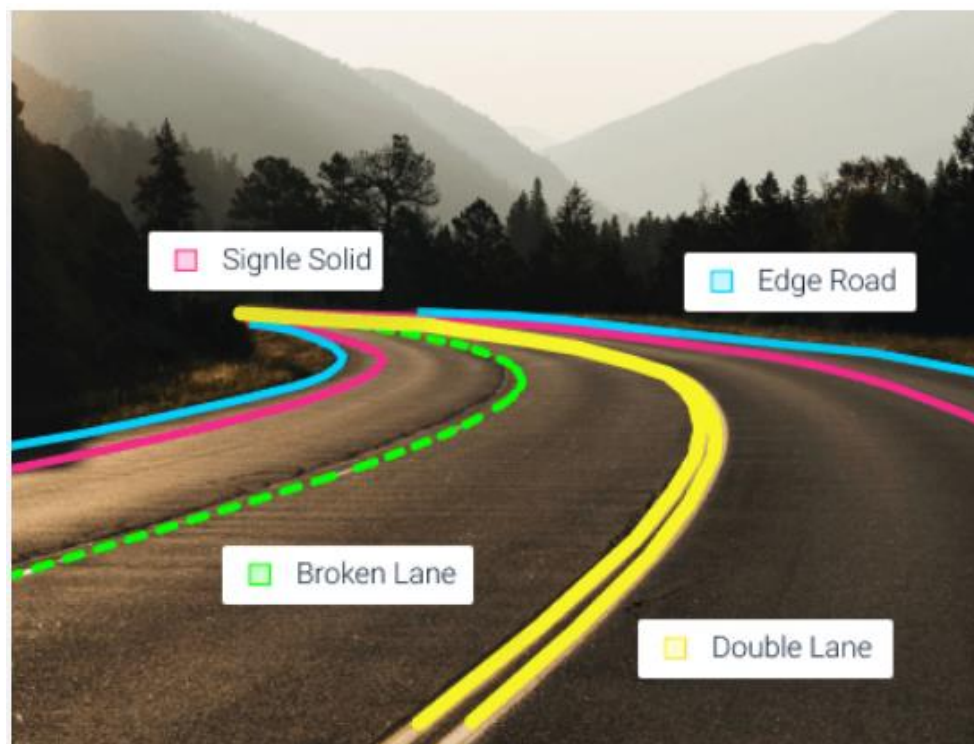
Shanshan Wang, PhD Student
School of Public Health
University of North Texas Health Science Center
ShanshanWang@my.unthsc.edu
October 28, 2023

Materials

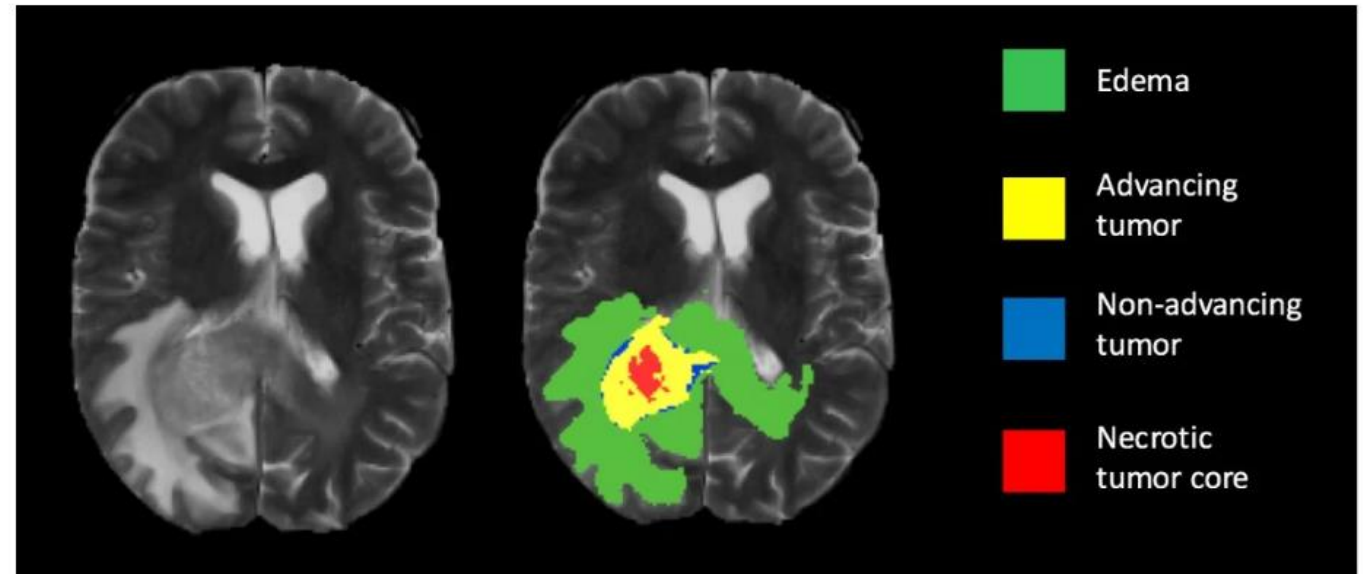
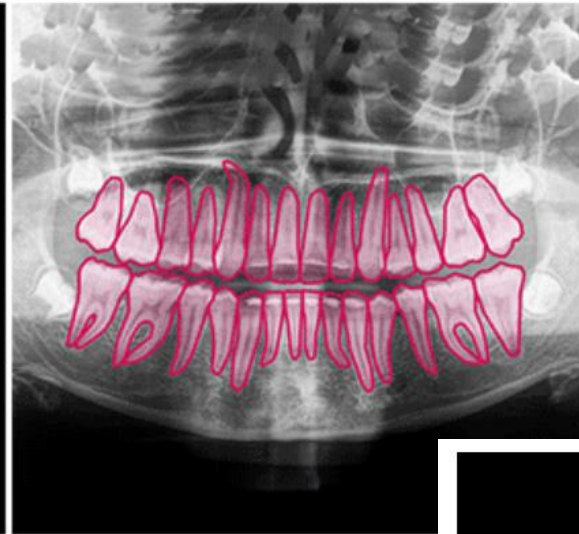
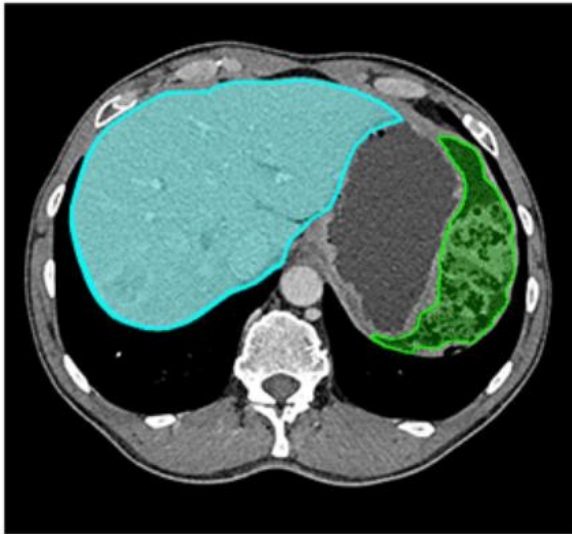


- Slides
- Data and Codes

Autonomous vehicles



Medicine



Surveillance



CONTENTS

01 What is Image Annotation?

02 Image Annotation Tools

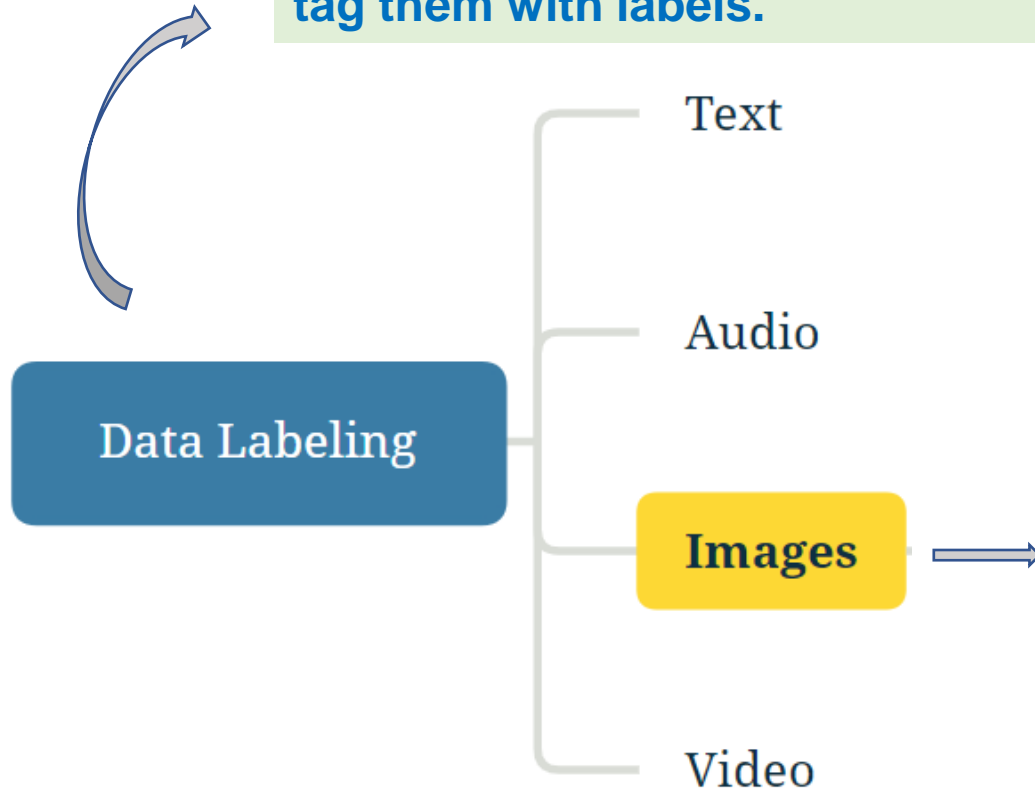
03 Practice



What is Image Annotation?

- **Computer Vision**
- **Annotation types**

Data labeling aims to **identify objects in raw data** (such as *images*, video, audio, or text) and **tag them with labels**.



- **Image annotation** is the human-powered task of **annotating an image with labels**.
- Those labels are **predetermined by the AI engineer** and are chosen to give the computer vision model information about what is shown in the image.

COMPUTER VISION (CV)

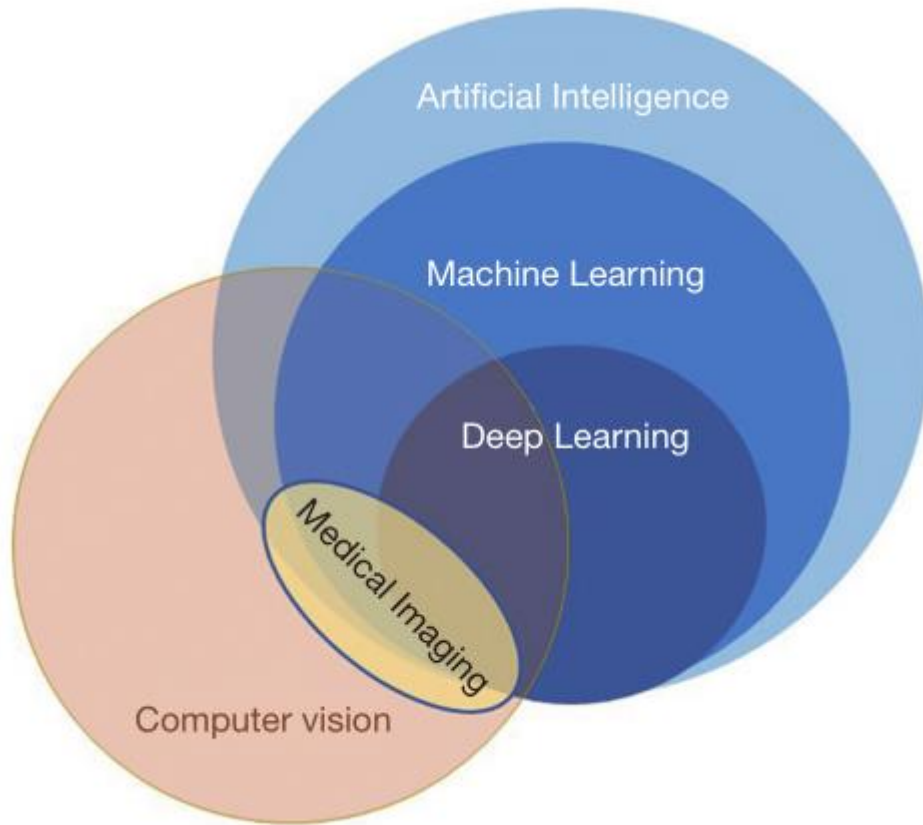
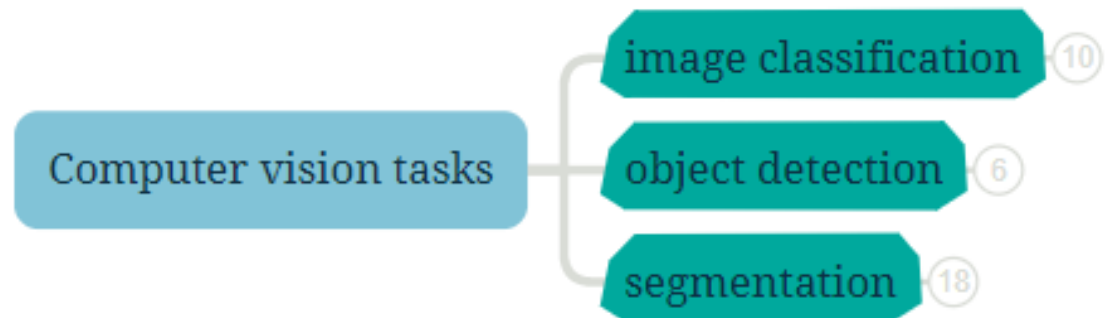
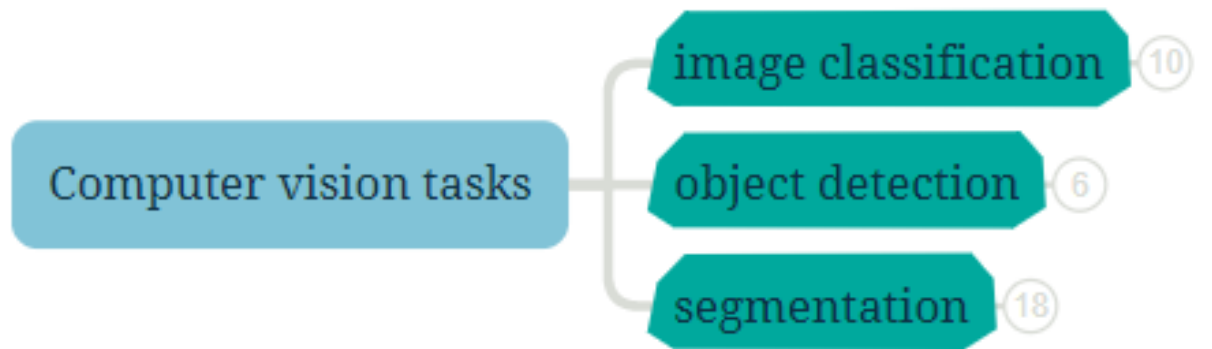


Figure. Relation between computer vision and artificial intelligence.

- CV deals with a large range of problems such as image segmentation, object recognition, detection, reconstruction, etc.
- It aims at modeling and understanding the visual world by extracting useful information from digital images, often inspired by complex tasks of human vision.
- Although it exists since the 1960s, it remains an unsolved and challenging task to the extent that only recently computers have been able to provide useful solutions in different application fields.
- **It is a multidisciplinary subject closely related to AI.**





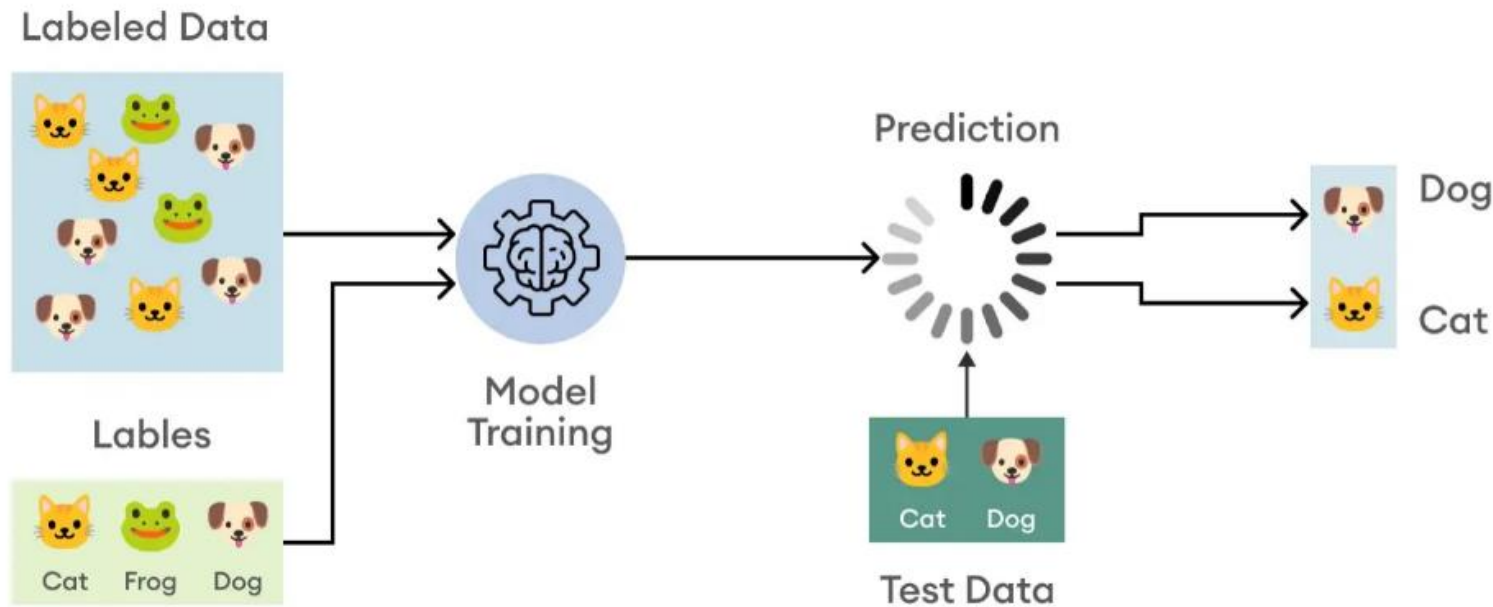
Computer vision tasks

Object detection



- The goal of **object detection** is to identify the presence of objects within an image or video and to determine their spatial location and extent within the image.

Image classification



- **Image classification** is the process of categorizing an image into one or more predefined classes or categories.

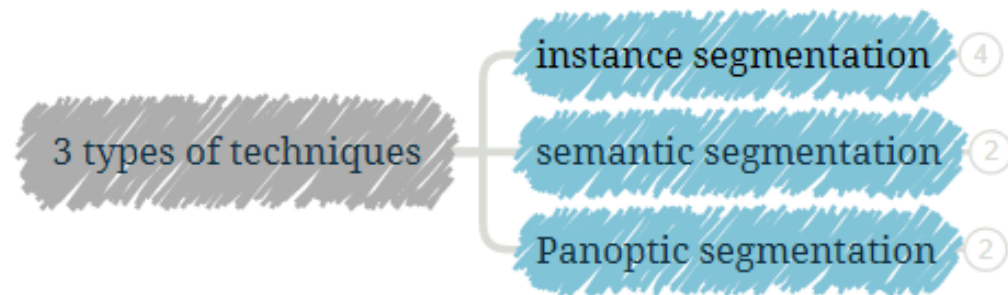
<https://encord.com/blog/image-annotation-guide/>









<https://www.superannotate.com/blog/image-classification-basics>

Image segmentation

A **segmentation mask** is a specific portion of an image that is isolated from the rest of an image.

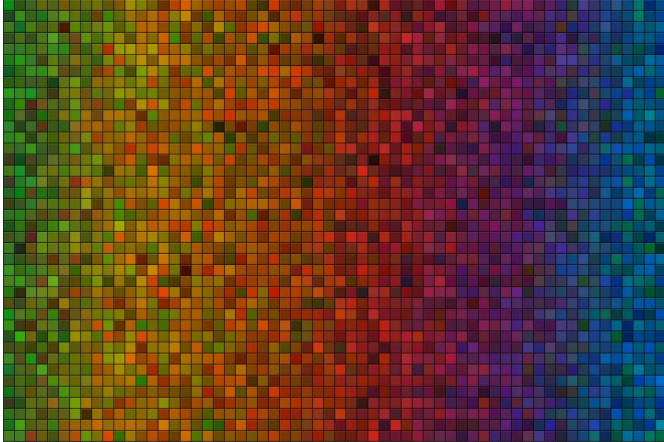
- **Segmentation** is a way of assigning a class to each pixel (or group of pixels) within images using **segmentation masks**.
- The main goal of image segmentation is to simplify and/or change the representation of an image into something more meaningful and easier to analyze.



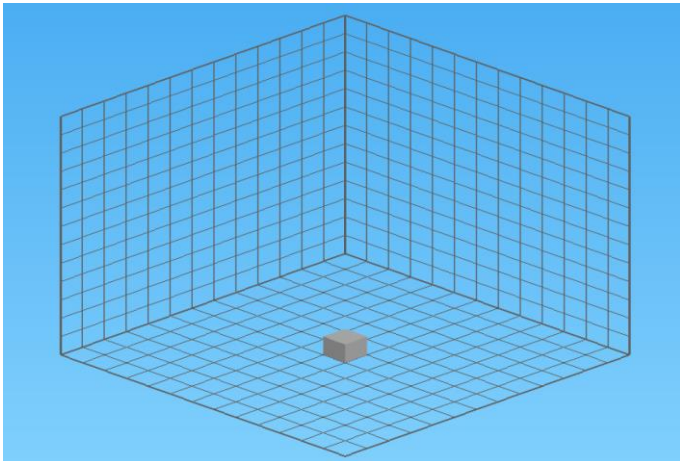
| | | | |
|--|--|--|---|
|  Road |  Sidewalk |  Building |  Fence |
|  Pole |  Vegetation |  Vehicle |  Unlabel |

Semantic segmentation

Pixel and Voxel



- In digital imaging, **a pixel** (abbreviated px), pel, or **picture element** is the smallest addressable element in a raster image, or the smallest addressable element in a dot matrix display device.
- In most digital display devices, **pixels are the smallest element** that can be manipulated through software.



- In 3D computer graphics, a voxel represents a value on a regular grid in three-dimensional space.
- **Voxels** are essentially 3D pixels, but instead of being squares, they are perfect cubes.

Image segmentation

- It involves **labeling each pixel** in an image with a specific class or category, such as “person”, “cat”, or “unicorn”.
- **Don't differentiate instances, only care about pixels.**



Instance segmentation



Semantic segmentation



Panoptic segmentation

- It is a technique that involves **identifying and delineating individual objects** within an image.
- **Every instance** of an object is **uniquely identified**, and **each pixel** in the image is assigned to a specific instance.

- It is a **hybrid of instance and semantic segmentation**, where the goal is to assign every pixel in an image to a specific instance or semantic category.

Annotation types

bounding boxes

keypoint ④

polygons

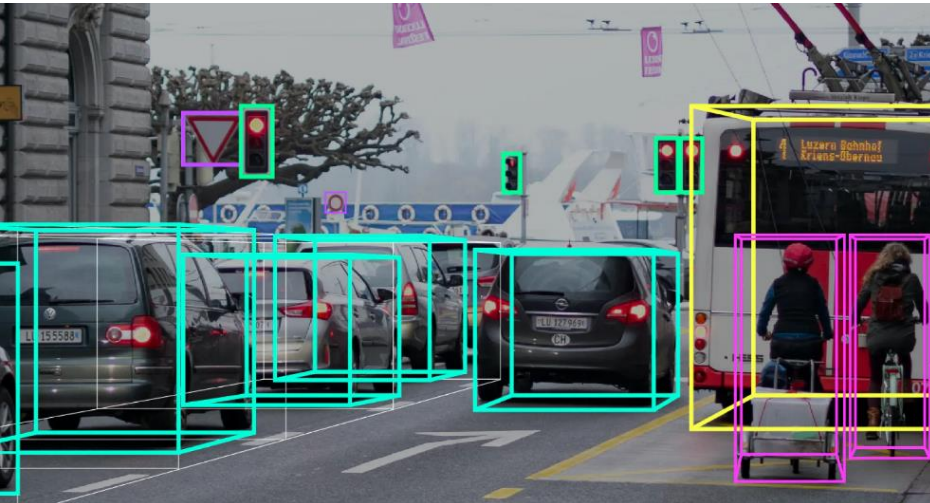
polylines

Annotation types

Bounding boxes



2D Bounding boxes

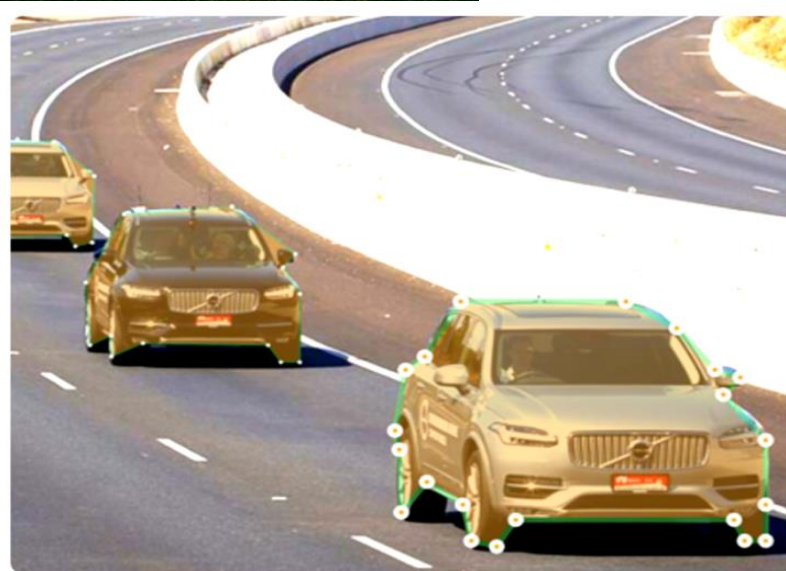


3D Bounding boxes

-**Cuboids**

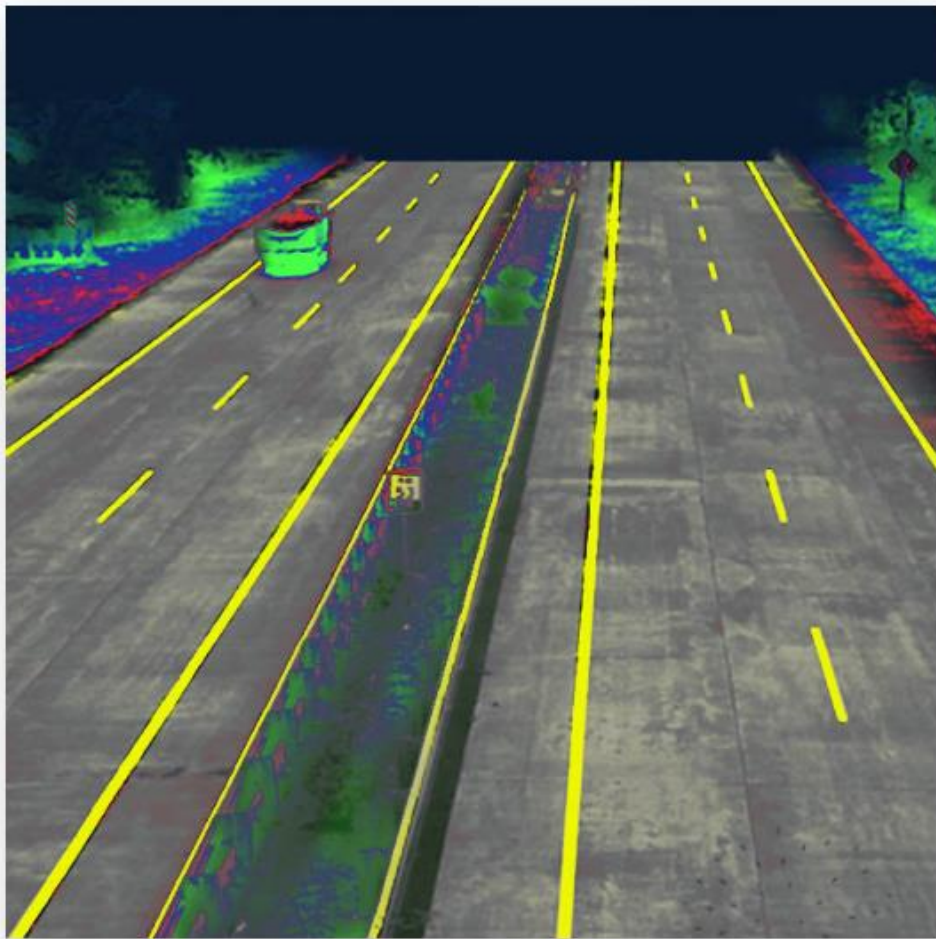
- Draw rectangular boxes around any object, and then apply a label to that object
- To define the spatial extent of the object and to provide a visual reference for machine learning models that are trained to recognize and detect objects in images

Polygons



- Polygon annotation captures more lines and more angles when compared to bounding box annotation.
- Outline complex shapes
- More precise

Polylines



- The labeling of straight or curved lines on images
- A way of annotating and labeling something static that continues throughout a series of images, such as a road or railway line
- Help autonomous vehicles detect street lanes across roads on cities and highway

Keypoints



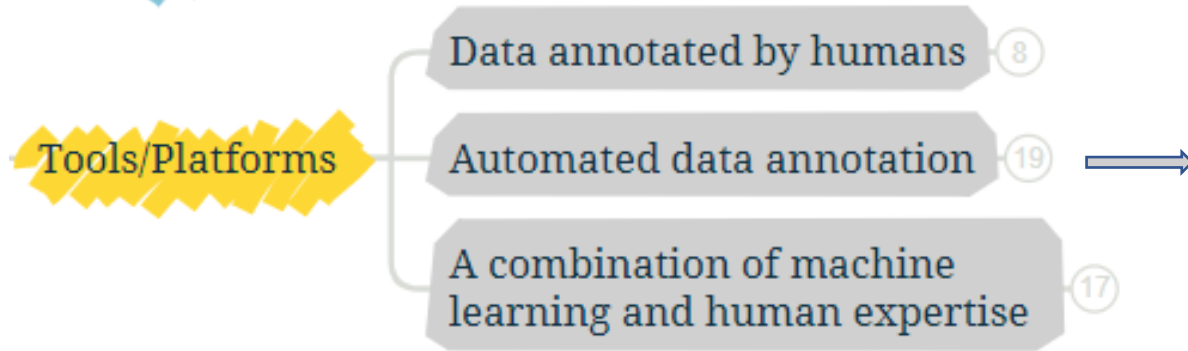
- Are typically important features or landmarks, such as the corners of a building or the joints of a human body
- Human pose estimation
- Skeleton



Image Annotation Tools

- **Commercial**
- **Open source**

Commercial: Amazon SageMaker Data Labeling



Commercial: Hive AI, Appen, Clarifai, Labellerr, Supervisely, Plainsight's vision AI platform, SuperAnnotate

Open source: **CVAT (Computer Vision Annotation Tool)**, Meta's Segment Anything Model, MONAI Label

Commercial: Scale AI, Labelbox, Encord Annotate, CloudFactory, Playment, IO. Annotator, Dataloop, V7 Labs

Open source: **VoTT, Labelstudio**

Human-in-the-Loop (HITL): In machine learning and computer vision training, HITL is a concept whereby **humans** play an interactive and iterative role in a model's development.

Commercial

- Amazon SageMaker Data Labeling: SageMaker offers two options, **Amazon SageMaker Ground Truth Plus** and **Amazon SageMaker Ground Truth.** (create and manages data labeling workflows and the workforce on your behalf, give you the option to use human annotators through Amazon Mechanical Turk, third-party vendors, or your own private workforce)
- Hive AI: popular with companies like Reddit, Quora, NETFLIX, Walmart, etc. (Try a demo)
- Scale AI: works with Microsoft, OpenAI, Meta. (<https://dashboard.scale.com/rapid>)
- Labelbox: works with Walmart, Adobe, etc.
- Encord Annotate: informative
(examples of annotation types: <https://encord.com/image/>)

Open source

- **The Computer Vision Annotation Tool (CVAT):** an open-source project supported by *Intel*, under the OpenCV umbrella (language: Typescript, React, CSS, Python)
- **VGG Image Annotator (VIA):** Visual Geometry Group, Department of Engineering Science, University of Oxford.
- **MONAI Label:** <https://github.com/Project-MONAI/MONAILabel>
- **LabelMe:** emerged from the MIT Computer Science and Artificial Intelligence Laboratory, the ability to outsource data labeling through Amazon Mechanical Turk
<http://labelme.csail.mit.edu/Release3.0/>
- **RIL-Contour:** using iterative deep learning (IDL), <https://bio.tools/RIL-Contour>
- **Sefexa:** an open-source image segmentation tool, <http://www.fexovi.com/sefexa.html>
- **3D Slicer:** <https://www.slicer.org/>
- **Labelimg:** <https://github.com/HumanSignal/labelImg>
- **MAKE SENSE:** <https://www.makesense.ai/>

Open source

- **Visual Object Tagging Tool (VoTT)**: developed by **Microsoft**, praised for its tags and assets export to Tensorflow (PascalVOC) and YOLO format (language: Typescript)
- **Labelstudio**: praised for its manual annotation process capabilities (language: Python)

CVAT



OpenCV

- **OpenCV (Open Source Computer Vision Library)** is an open source computer vision and machine learning software library.
- The library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms.
- Free course: https://opencv.org/university/free-opencv-course/?utm_source=opcv&utm_medium=menu&utm_campaign=obc

Datasets

- <https://www.kaggle.com/datasets/alessiocorrado99/animals10/data>
- Animals-10
- A subset of 100 images

The Computer Vision Annotation Tool (CVAT)

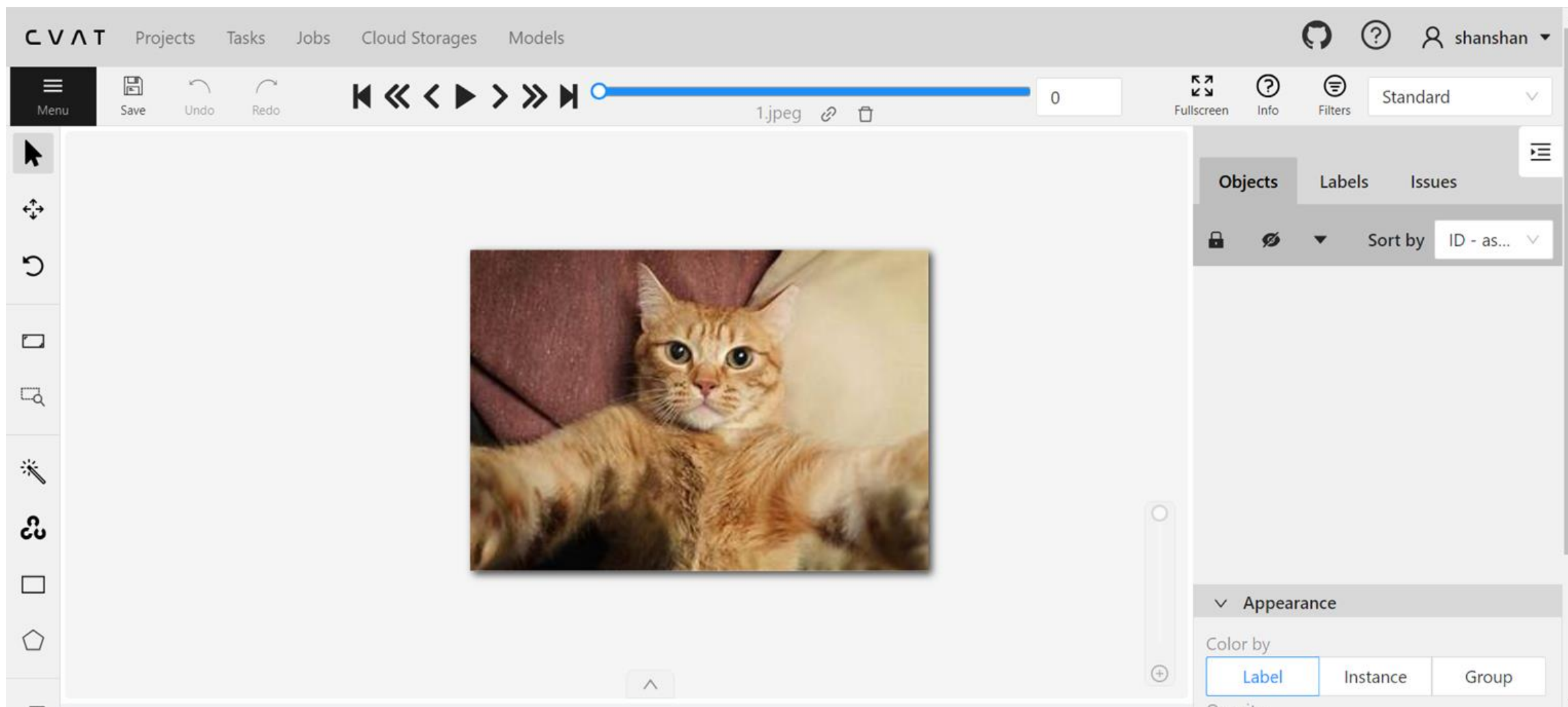
- **Interface**
- **<https://app.cvat.ai/projects?page=1>**

1. **Create a new project**
2. **Create a new task**
3. **Define labels**
4. **Manual vs. AI-assisted**
5. **Save annotations**
6. **Export dataset**

1. Create ground truth task (quality control, honeypot)
 2. Bounding box/cuboid, Polygon, Polyline, points, brush (mask); video
 3. Change job status
 4. AI-assisted, OpenCV
 5. Automatic annotation (upgrade)
 6. Image classification
 7. Analytics and monitoring, assignee, stage
 8. Top panel: project, task, jobs, **models**
 - <https://universe.roboflow.com/>
 - <https://huggingface.co/models>
- **AWS, Azure**

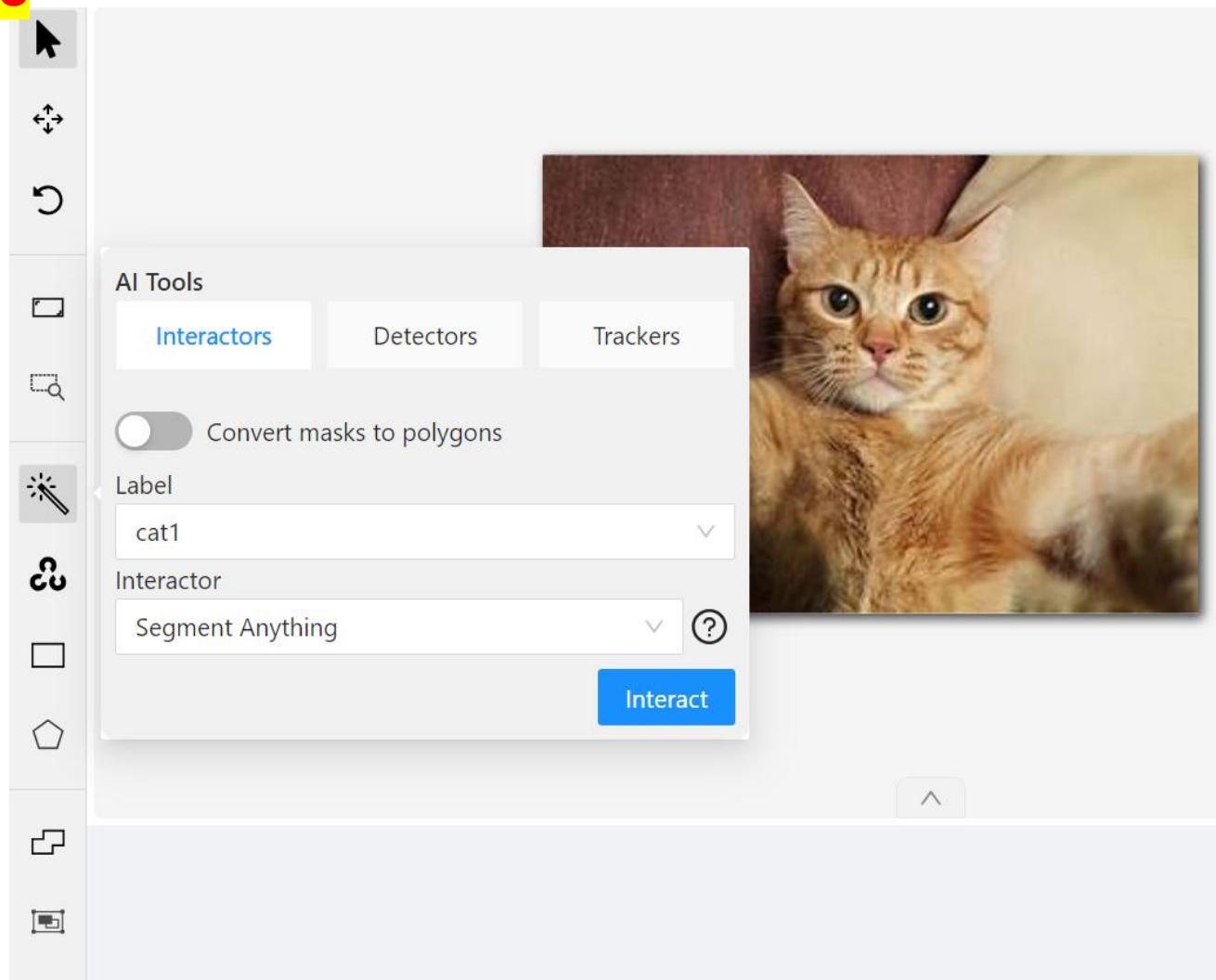
Manual

CVAT User Interface



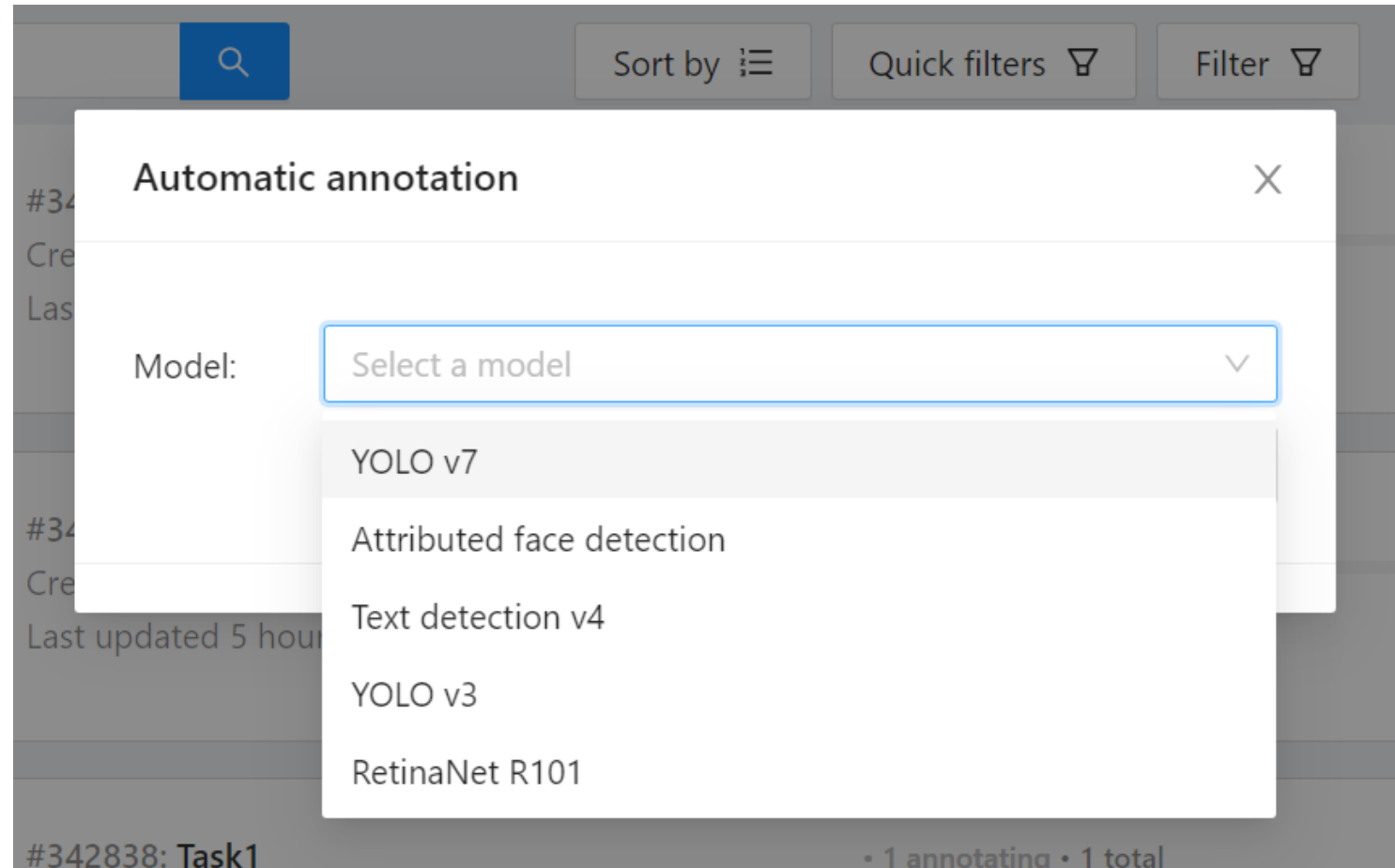
AI-assisted,
Semi-automatic

CVAT User Interface



Automatic annotation (upgrade)

CVAT User Interface



CVAT User Interface

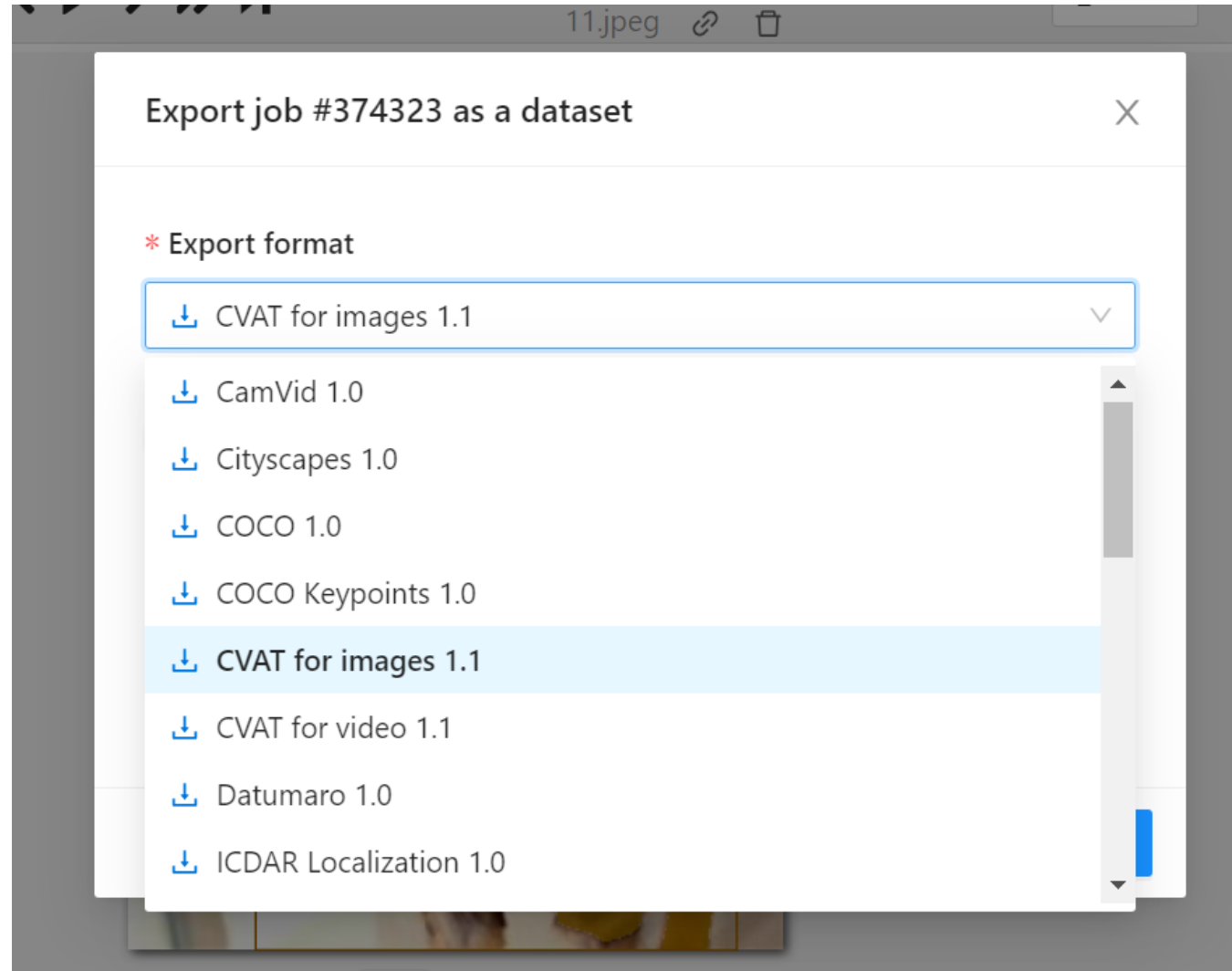
Model type:

- **Detector** - used for automatic annotation (available in [detectors](#) and [automatic annotation](#))
- **Interactor** - used for semi-automatic shape annotation (available in [interactors](#))
- **Tracker** - used for semi-automatic track annotation (available in [trackers](#))
- Reid - used to combine individual objects into a track (available in [automatic annotation](#))

<https://opencv.github.io/cvat/docs/manual/advanced/ai-tools/>

<https://opencv.github.io/cvat/docs/manual/advanced/automatic-annotation/>
(details about the models)

CVAT User Interface



Data export formats

- The choice of export format depends on the type of annotation as well as the intended future use of the dataset.

<https://opencv.github.io/cvat/docs/manual/advanced/formats/>

Popular Formats by Task and Use Case:

Classification: ImageNet, OpenImages

Detection: YOLO, COCO, PASCAL VOC

Segmentation: PASCAL VOC, CamVid, COCO, MOTS, Cityscapes

Keypoints: COCO keypoints

Video/Tracks: MOT, CVAT for Video, MOTS

Universal: CVAT, Datumaro

Tool

Compatibility: LabelMe, TFRecord, Supervisely Point Cloud

3D: KITTI, Supervisely Point Cloud

API

SDK (Python library) API includes several layers:

1. Low-level API with REST API wrappers.

- Located at `cvat_sdk.api_client`.
- The low-level API is useful if you need to work directly with REST API, but want to have data validation and syntax assistance from your code editor. The code on this layer is autogenerated.

2. High-level API.

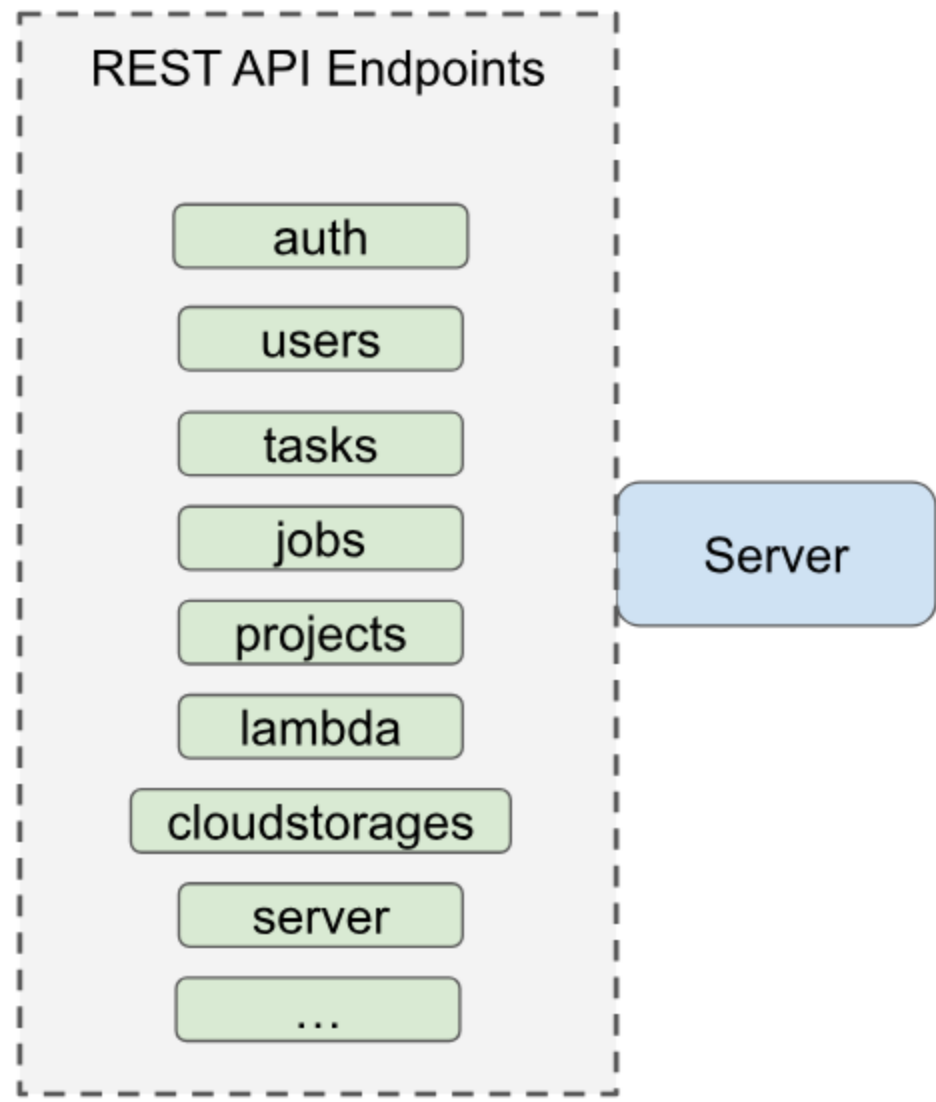
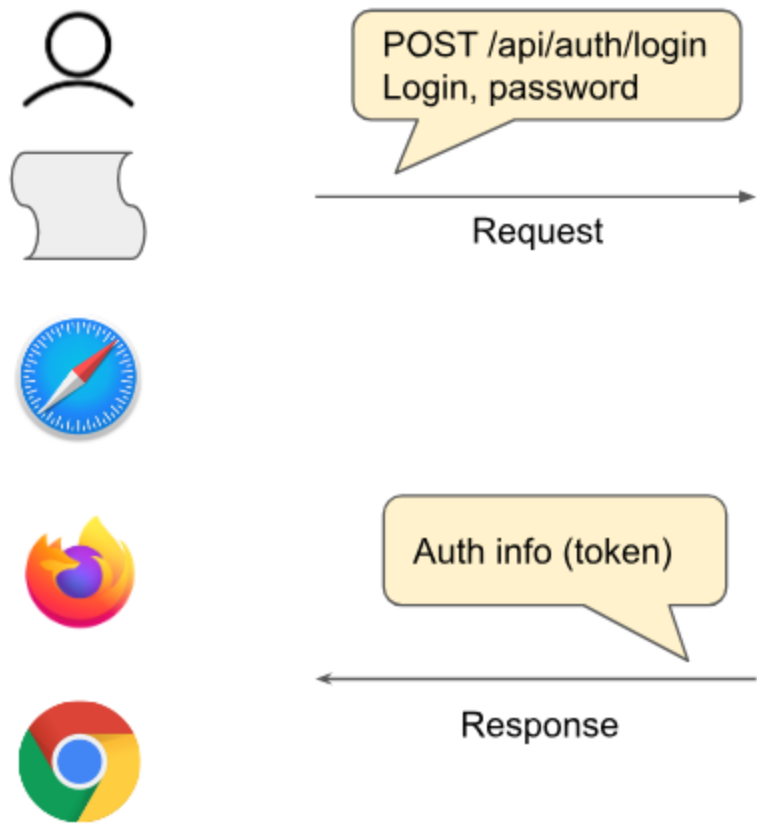
- Located at `cvat_sdk.core`.
- This layer provides high-level APIs, allowing easier access to server operations.

3. PyTorch adapter.

- Located at `cvat_sdk.pytorch`.
- This layer provides functionality that enables you to treat CVAT projects and tasks as PyTorch datasets.

4. Auto-annotation API.

- Located at `cvat_sdk.auto_annotation`.
- This layer provides functionality that allows you to automatically annotate a CVAT dataset by running a custom function on your local machine.



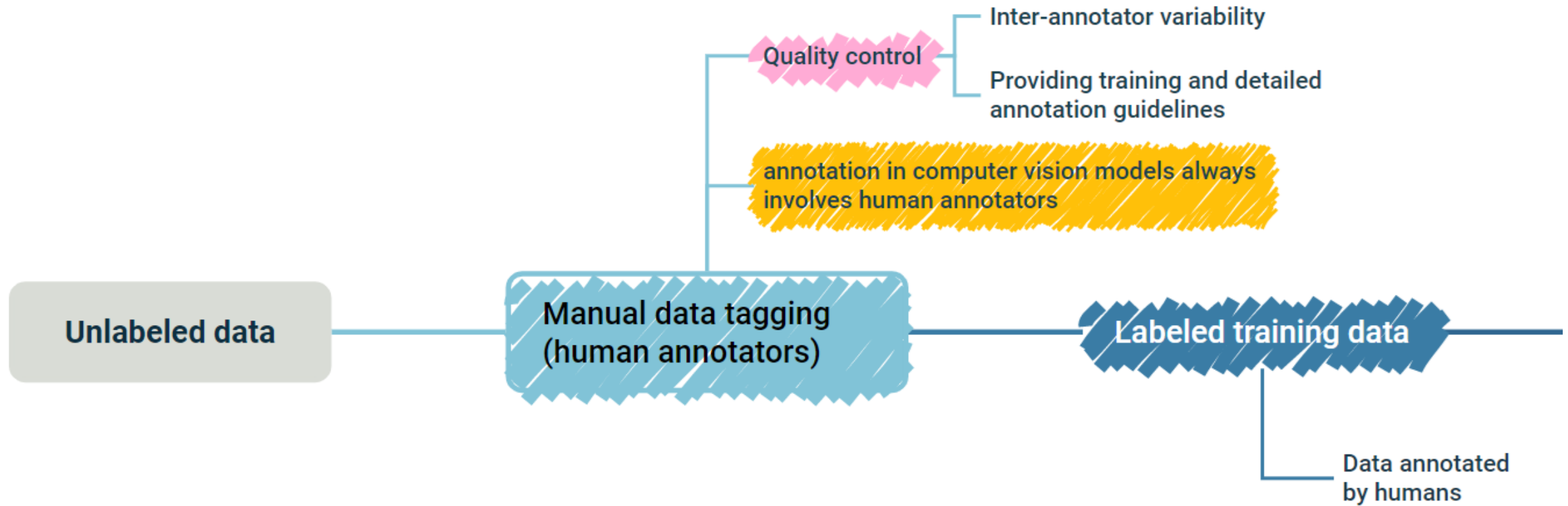
The Computer Vision Annotation Tool (CVAT)

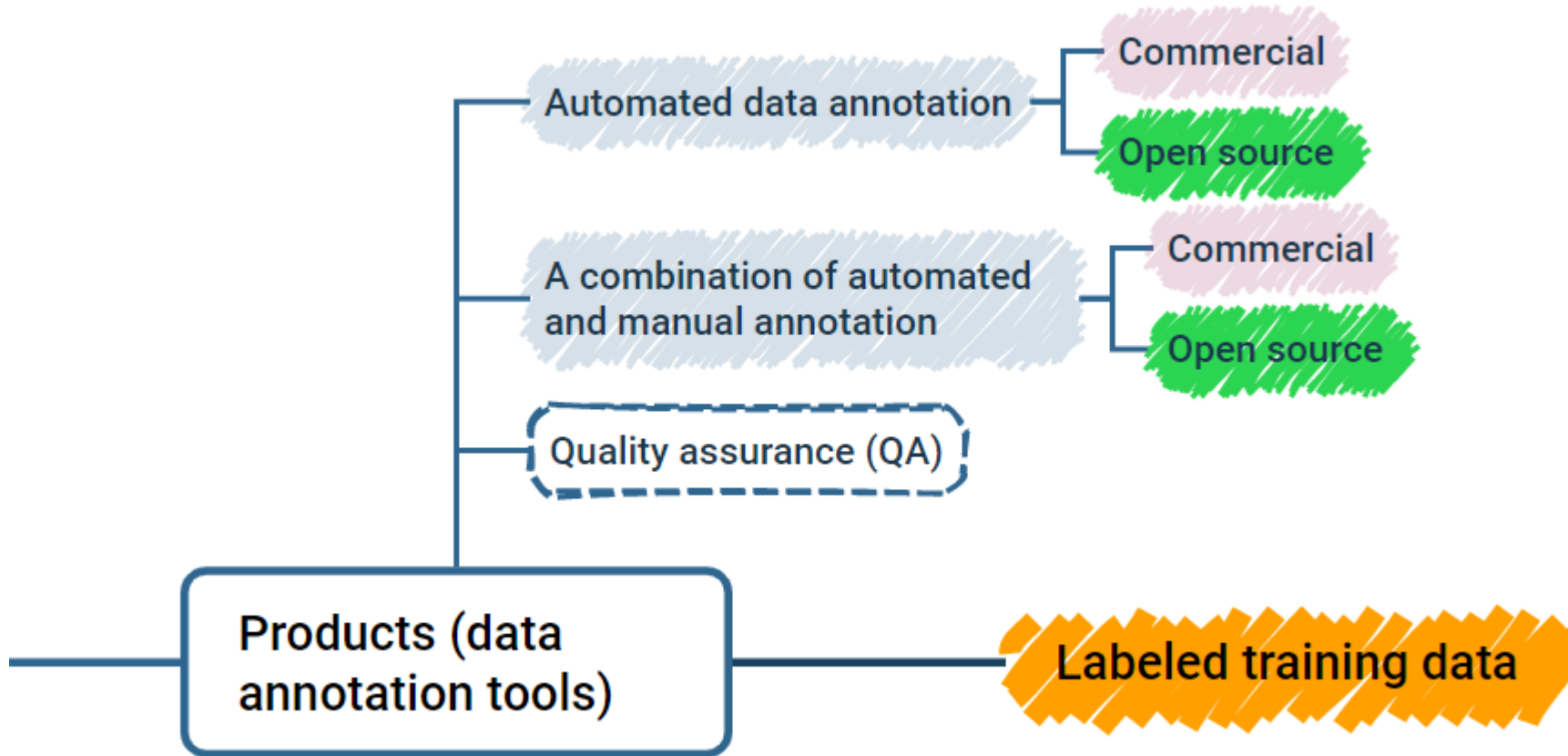
Reference codes:

- https://opencv.github.io/cvat/docs/api_sdk/sdk/auto-annotation/

Why do we use it?

- Researchers

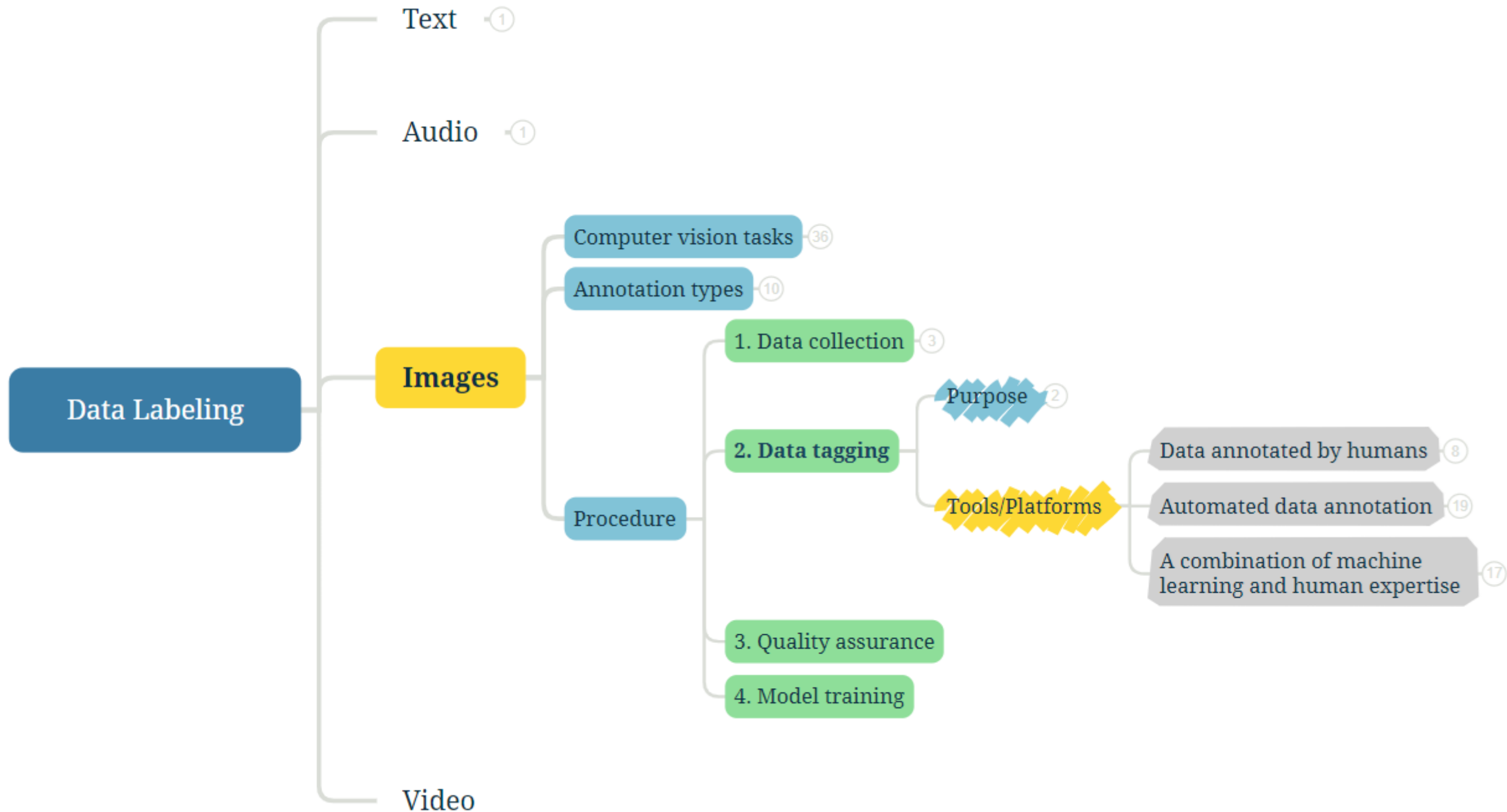


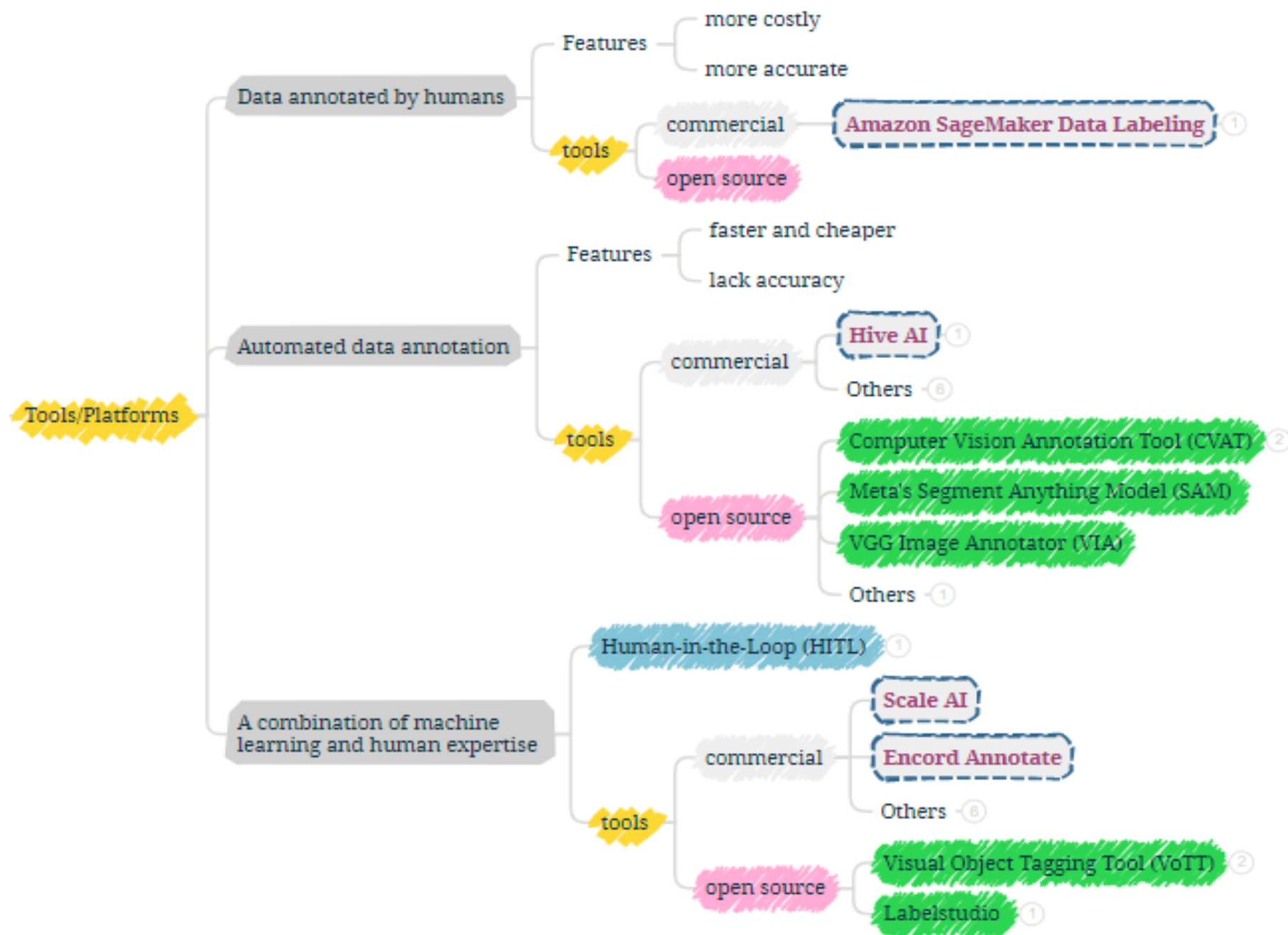


- Save time on data preparation
- Always involve human annotators

Recap







REFERENCES

- 10 Best Image Annotation Services Providers 2023-Reviewed <https://www.labellerr.com/blog/10-best-image-annotation-services-providers-of-2023-reviewed/>
- 9 Best Image Annotation Tools for Computer Vision <https://encord.com/blog/9-best-image-annotation-tools-for-computer-vision/>
- The Complete Guide to Image Annotation for Computer Vision <https://encord.com/blog/image-annotation-guide/>
- Complete Guide to Open Source Data Annotation <https://encord.com/blog/best-open-source-annotation-tools/>
- 10 best annotation tools for computer vision applications
<https://www.youtube.com/watch?v=LQe7XpIKfcE&list=PL11b9pg0T7WFDCmJgpqZ1bTzFtrvRSA52&index=5>
- Top 10 Open Source Data Labeling/Annotation Platforms in 2023
<https://research.aimultiple.com/open-source-data-labeling/>
- 10 of the best open-source annotation tools for computer vision <https://humansintheloop.org/10-of-the-best-open-source-annotation-tools-for-computer-vision/>

THANK YOU!



Question