**# YOLO Model Training on Custom Vehicle Data**

**Contents:**

yolo\_training\_project/

├── dataset/

│ ├── train/images/ # (put your training images here)

│ ├── train/annotations/ # COCO JSON: instances\_train.json

│ ├── val/images/ # (put your validation images here)

│ └── val/annotations/ # COCO JSON: instances\_val.json

├── configs/

│ └── custom\_data.yaml # autogenerated inside script

├── models/ # optional model configs

└── yolo\_train.py # full OOP training + ONNX export

**How to use on Google Colab:**

1. Upload and unzip:
2. !unzip yolo\_training\_project.zip
3. %cd yolo\_training\_project
4. Train:
5. !python yolo\_train.py --data\_dir dataset --model yolov5s --epochs 100 --batch\_size 8 --img\_size 640
6. Download model from:
7. yolov5/runs/train/vehicle\_detector/weights/best.onnx

**COCO annotation format** using a practical example with:

**✅ Setup:**

* **3 classes**: "car", "bus", "truck"
* **Images**: 5 training images and 5 validation images
* **Folder structure**:

dataset/

├── train/

│ ├── images/ # image\_1.jpg to image\_5.jpg

│ └── annotations/

│ └── instances\_train.json

└── val/

├── images/ # image\_6.jpg to image\_10.jpg

└── annotations/

└── instances\_val.json

**🔍 COCO JSON Format Overview**

Each .json annotation file (e.g., instances\_train.json) has:

* images: metadata for each image (filename, size, ID)
* annotations: list of bounding boxes for all objects
* categories: mapping of class names to class IDs

**🧾 Sample instances\_train.json (5 annotated training images)**

{

"info": {

"description": "Vehicle Dataset COCO Format"

},

"images": [

{"id": 1, "file\_name": "image\_1.jpg", "width": 1280, "height": 720},

{"id": 2, "file\_name": "image\_2.jpg", "width": 1280, "height": 720},

{"id": 3, "file\_name": "image\_3.jpg", "width": 1280, "height": 720},

{"id": 4, "file\_name": "image\_4.jpg", "width": 1280, "height": 720},

{"id": 5, "file\_name": "image\_5.jpg", "width": 1280, "height": 720}

],

"annotations": [

{"id": 1, "image\_id": 1, "category\_id": 1, "bbox": [100, 200, 200, 100], "area": 20000, "iscrowd": 0},

{"id": 2, "image\_id": 2, "category\_id": 2, "bbox": [300, 150, 250, 120], "area": 30000, "iscrowd": 0},

{"id": 3, "image\_id": 3, "category\_id": 3, "bbox": [500, 300, 180, 90], "area": 16200, "iscrowd": 0},

{"id": 4, "image\_id": 4, "category\_id": 1, "bbox": [50, 250, 190, 95], "area": 18050, "iscrowd": 0},

{"id": 5, "image\_id": 5, "category\_id": 2, "bbox": [220, 280, 240, 110], "area": 26400, "iscrowd": 0}

],

"categories": [

{"id": 1, "name": "car", "supercategory": "vehicle"},

{"id": 2, "name": "bus", "supercategory": "vehicle"},

{"id": 3, "name": "truck", "supercategory": "vehicle"}

]

}

**🧾 Sample instances\_val.json (5 annotated validation images)**

{

"images": [

{"id": 6, "file\_name": "image\_6.jpg", "width": 1280, "height": 720},

{"id": 7, "file\_name": "image\_7.jpg", "width": 1280, "height": 720},

{"id": 8, "file\_name": "image\_8.jpg", "width": 1280, "height": 720},

{"id": 9, "file\_name": "image\_9.jpg", "width": 1280, "height": 720},

{"id": 10, "file\_name": "image\_10.jpg", "width": 1280, "height": 720}

],

"annotations": [

{"id": 6, "image\_id": 6, "category\_id": 3, "bbox": [600, 200, 200, 100], "area": 20000, "iscrowd": 0},

{"id": 7, "image\_id": 7, "category\_id": 1, "bbox": [100, 180, 220, 110], "area": 24200, "iscrowd": 0},

{"id": 8, "image\_id": 8, "category\_id": 2, "bbox": [150, 250, 230, 100], "area": 23000, "iscrowd": 0},

{"id": 9, "image\_id": 9, "category\_id": 3, "bbox": [200, 300, 210, 90], "area": 18900, "iscrowd": 0},

{"id": 10, "image\_id": 10, "category\_id": 1, "bbox": [120, 100, 200, 120], "area": 24000, "iscrowd": 0}

],

"categories": [

{"id": 1, "name": "car", "supercategory": "vehicle"},

{"id": 2, "name": "bus", "supercategory": "vehicle"},

{"id": 3, "name": "truck", "supercategory": "vehicle"}

]

}

**📌 Key Format Notes**

| **Field** | **Description** |
| --- | --- |
| image\_id | Matches the id from the images section. |
| category\_id | Integer ID referring to the class (e.g., car=1). |
| bbox | [x\_min, y\_min, width, height] in pixels. |
| area | Area of the bounding box (optional but good to include). |
| iscrowd | 0 = individual object, 1 = grouped objects (e.g. crowd). |

Here's a visualization of a sample COCO annotation on a dummy image.

A screenshot of a computer

AI-generated content may be incorrect.

* The gray background represents a placeholder image (sample\_image.jpg).
* The red rectangle shows a detected object (vehicle) using the bounding box provided in the COCO annotation.
* The label vehicle is drawn above the bounding box.