# How to train a YOLO11n detector for Suction Cup Marks

Below are the complete step-by-step instructions for dataset creation and model training process. You may wish to skip dataset creation steps, download ready-to-use dataset in YOLO format from [here](https://epam.sharepoint.com/:f:/s/EPAMFutureDial/Eik6k9IUGpVBnfiliFmM_owBvZZaCALLF6qY7CDlXUeHRw?e=Rb85M6) and proceed directly to the model training step. However, to update that dataset with the new data collected, you will need to follow the instructions below.

## Prerequisites:

1. Install the ultralytics package from PyPI:

**pip install ultralytics**

## Original dataset preparation:

1. You need to have [Batch1](https://www.dropbox.com/scl/fo/xzl1zr4pk4fg871x6qoiv/AGj9Z0DgN5pgxDojSYgf-08?rlkey=4d4t1uk36sfaa2lkm8tgexhuo&e=2&st=edvyme15&dl=0) and [Batch2](https://www.dropbox.com/scl/fo/04aclwat417w30ybo75ri/AEeWK5rdjLHhBiL8EiutVlw?rlkey=gohtrn7ukrfguh68gq6jh94om&e=2&st=76rqqkl6&dl=0) downloaded on the local machine.
2. Download and unzip files with corresponding markup labels from [dataset\_labels](https://epam.sharepoint.com/:f:/s/EPAMFutureDial/EsIDjg979G5AtsqO4g-ISaMB7xolDAyoKglzTyRgygvQbA?e=zphXso).
3. Merge folders containing labels into corresponding dataset folders. You will get coco.json files in every sample folder and two .txt files located nearby source .bmp files: A screenshot of a computer

   AI-generated content may be incorrect.
4. The dataset is ready to be converted to YOLO format.

## Dataset conversion:

1. Prepare the script **YOLO\_dataset\_creator.py** to utilize this dataset, update parameters section in the file, so that BASE\_PATH = '…\\Downloads\\Batch1\\Suction Cup Marks\\' points to the correct folder on your machine.
2. Set TARGET\_DATA\_PATH = … to the path where converted images and annotations will be stored.
3. Run Python script without any arguments given.
4. Check TARGET\_DATA\_PATH folder, it should contain small **bmp** suction cup mark images, png images with labels and corresponding txt files with labels.
5. Repeat the process for …Batch2\\Suction Cup Marks\\ folder. If everything is correct, you will get 44 samples of data for training int the TARGET\_DATA\_PATH folder.
6. Create folder where YOLO dataset will be placed and create subfolder structure in it as indicated in Ultralytics documentation: A white background with black lines

   AI-generated content may be incorrect.
7. Copy all created **bmp** images into train and val images folders.
8. Copy all txt files with labels into train and val labels folders.
9. Add coco.yaml file to the root dataset folder, **path** variable in it should point to the dataset root folder you created: A screen shot of a computer

   AI-generated content may be incorrect.
10. The dataset is ready to use, to make it fully compatible with the one provided from us, you need to add one extra false negative class sample image to train folder and corresponding empty txt file. For that, you need to repeat steps 1 and 2 for the folder **..Batch3\SuctionCupMarks\** and add small generated copy of **353516542025760\_21\_3\_1\_055525265.bmp** image to the dataset/train/images folder. Then add empty **353516542025760\_21\_3\_1\_055525265.txt** file to the dataset/train/labels folder. This step is optional and helps get higher precision for the model. You may wish to add other extra samples (both positive and negative) to the dataset if needed.

## Model training

1. Open the script **YOLO\_suction\_cup\_mark\_detector.py**. Modify parameters section so that TRAINING = True, DATASET\_PATH points to the root folder of the prepared dataset. You may wish to also change PROJECT\_NAME variable to fit your needs.
2. Run Python script without any arguments given.
3. Training for 250 epochs takes some time, at the end in the PROJECT\_NAME/train2 subfolder you will find complete set of artifacts including the models in weights subfolder: A screenshot of a computer program

   AI-generated content may be incorrect.

## Model testing and usage

1. Open the script **YOLO\_suction\_cup\_mark\_detector.py**. Modify parameters section so that TRAINING = False, BASE\_PATH = … so that it points to the Dataset subfolder which will be used for testing, and optionally, OUTPUT\_PATH = … where model predictions will be stored.
2. IMPORTANT: To get correct predictions, modify CONFIDENCE\_THRESHOLD in the parameters section with optimal value obtained during model training. Value can differ from the one obtained in previous experiments. Value can be found in F1 graph plot: A screen shot of a graph

   AI-generated content may be incorrect.

Futher steps for model optimization and conversion to the OpenVINO format will be provided later.