

Name	Roll No
AU2140002	Parth Joshi
AU2140220	Shubham Apat
AU21	Shubham Shah

#### Report:

We filtered the whole data and have divided the data in train and validation. We are ready with the code to apply yolo-v8 on the dataset and train it. The problem currently is that there is someone working on the super computer and we are not given the access to fully use the GPU and so the training needs to hold till it is free.

We faced several issues this week. Firstly, we had trouble in getting the internet access on the param shavak which took two days to get access and then download the libraries required for the dataset and model to run.

We developed a fully new virtual environment for python on the param shavak. Param shavak only had python2.7 and to run yolo-v8 we require ultralytics library which only works on the python3.4 or higher version. We had set up a full environment of python 3.6 and downloaded all necessary libraries and sorted data.

We had an issue of space in the param shavak so that thing also took a lot of time to get solved. Our code is ready to train on yolo. On batch size of 1 the model is fully functional but for this large dataset higher batch size is required and for that GPU is required at full potential which is not available. We have attached the code to sort the dataset and rewire it to have the folder name in the xml file along with naming the file convention appropriately as per dataset. We have also converted the format of .XML files into YOLO txt format. We have attached the python script for that too.

There is a yml file required to do the task for yolo-v8 which we have created. We are attaching that too. It does the following of path and then mapping the folders from 1 to 160 to '1' to '160'.

\*(All the code and files mentioned are attached in github)

### Next Week Goal:

For next week's plan, hopefully we would get access to GPU storage to use it at its optimum level and start the training as soon as possible. Based on the size of the dataset and considering all the complexities in training the dataset, the amount of time training would be done will be a minimum 1 to 2 weeks.