**Written By :** Meet Parmar, 2023A7PS0406G

**Process and Reasoning of Collecting Data To Train Model:**

**AIM :**

Gathering a dataset of images to train a model that segments the images. The segmented images are going to help the autonomous vehicle to drive by identifying roads/footpath. Our aim was to get a varied dataset having images from different angles , different light conditions , different atmospheric conditions.

**Process:**

Aligning to our aim , I gathered 30 frames per video that were located 10 frames away from each other. The parameters that is max\_frames\_per\_video and the distance between each frame can be altered later on to obtain a slightly better dataset. I will later in this document explain why the current parameters have been chosen and try to come to a conclusion what can be improvised and implement the improvisations too.

So the first goal was to shoot videos of the D side road in BITS Goa Campus in varying light conditions, different angles , different environment conditions and during different time of the day.

The biggest challenge in shooting the video was that of course you cannot get all of the variations in one go , it took me 3-4 days to shoot videos of varying conditions like one being during the day, one during the night , one during rainy roads ,one during cloudy weather , one during shiny weather , one having glare from the sunlight , one having glare from the break lights of the car ahead , one having occlusions so to find these conditions it took me 3-4 days to come up with the videos.

Then my next goal was to ensure that the videos I shot were as varied as possible so that the model can be rigorously be trained in different conditions. So from 15 videos that I shot only 8 made it to the final stage because rest of them had overlapping/similar conditions.

Then my next goal was to extract images from the videos that made it to the final stage. Since I had to collect many number of images from a particular video and there were many videos there was a lot of manual labour involved in this so , I automated this task with the help of a python script which can be found in the GitHub repo under the branch MeetParmar\_2023A7S0406G. In that script I read all the 8 videos in an object and extracted 30 frames per video and saved them in a directory. Now this task of collecting videos and naming the videos was also a lot of manual labour so I also automated this by writing a python script using the os module I renamed all the videos and numbered them from 1 to 8 and loaded them one by one in the cap named object. Then to save the chosen frames I also automated that using the same os module and also created a directory to save the extracted the images and the format of the images is video-number\_frame-number.jpg.

Hence till now I collected the images to be used as the dataset.

**Reasoning**

In this section I will explain why did I do, what I did . So first of all I collected 15 videos reason being it was important to have a good amount of initial raw data to begin with and then remove the unwanted data.

Then I mentioned that I took this 15 videos over 4 days , reason being that the dataset must be as varied as possible so that our model can be trained on varied data and it was impossible to have varied conditions on one day at the same time.

Then I mentioned some of the videos couldn’t make it to the final stage reason being some of them were shot in similar conditions hence would have overlapped with the rest.

I chose the same roads during a different time in the day reason being to get varied road conditions and varied light conditions moreover I also shot the videos of same road in different atmospheric conditions from different angles to have some variations.

Then the reason I automated this process was because It would have been a tedious task to get 200+ images form 8 videos and name them.

Why did choose the parameters in my code that I chose . I will have an elaborate discussion on this . This is something where I am still working on currently I have chosen 30 frames per video distanced at 10 frames from each other. I choose 30 but it can also be increased to any number in fact its always better to have more number of frames as dataset but there also should be a limit otherwise handling all these data can be very computationally expensive and also the data will overlap/ not have much changes in them.

Then I decided to keep the gaps between the selected frame to be of 10 the reason being to get some changes in the selected frames. This is something in which changes can and should be made because the frames are a bit similar it can be changed to 20 frames gap. But this can be changed easily just by making a few line changes in the python script variables.

And finally the reason why we need a varied dataset is so that we can train the model in different variables that are angles , lightning condition , atmospheric conditions so that it gives us accurate results which can be used in path planning for the auto vehicle that we aim to build in this project.

Also I will attach the use case diagram separately.