

Image annotation

The objective of creating this dataset to train AI models for traffic analysis or autonomous driving. This dataset contains 150 images which were taken from YouTube. This video showed the busy streets of Chawk bazar, Dhaka, Bangladesh during day time. The images were annotated with **Computer Vision Annotation Tool (CVAT)** with labels such as- **rickshaw, car, motorcycle, pedestrian, leguna, truck and bus.**

Methodology:

Dataset:

The link of the video is given below:

 [4K Bangladeshi Walking Tour Dhaka City 2023 || Dhaka, Chawk bazar 4K Walking Tour 2...](#)

The video shows the busy streets of Chawk bazar, Dhaka, Bangladesh which are suitable for building traffic control or autonomous driving systems. The resolution of the video is 4K. 150 high quality images were extracted. The images were extracted in 1080p resolution.

Tool:

I used **CVAT** for annotation. The reason is that this tool is full of useful features that can handle annotations efficiently. After annotation, we can export the annotated dataset to any format compatible for any AI model.

Labels:

I chose seven labels for annotation. The reason behind choosing these labels is because they are a common scenario in the streets of Bangladesh. The labels are given below:

1. Rickshaw
2. Car
3. Motorcycle
4. Pedestrian
5. Leguna
6. Truck
7. Bus

Annotation:

Bounding boxes were drawn around each object in every image. Tight bounding boxes were used around the object so that the AI model can easily understand the object. Finally, annotated images were exported in **Pascal VOC format.**

Challenges and its solution:

The objects frequently overlapped with each other. So, I have to annotate carefully to separate the objects correctly. Secondly, I tried to use the python library to download the youtube video. The plan is to convert video into frame and extract 150 images. But the plan didn't work because I couldn't download the video higher than 360p. As a result, I took a screenshot from the video and annotated it.