# Project Overview

This project will be executed in two steps, focusing on vehicle detection in real-time videos and images using object detection techniques.

Dataset is Canadian Vehicle Dataset (Canadian road condition videos collected by Thales, Canada in different weather conditions so much robust and heterogeneous as compared to publicly available vehicle dataset.

## Step 1: Data Preparation (Due by April 1)

Team Task (5 members/team):

* Each team will annotate one video (or less).

Process:

1. Convert video into images (frames)-script will be shared.

2. Use the defined classes and numbering provided in the document attached.

3. Annotate images using any annotation tool (e.g., CVAT, LabelImg).

4. Ensure quality and completeness of annotations by having consistency in class id.

Grading Criteria for Step 1:

Team contribution should be clearly outlined

Annotations should follow given label/class structure

Submission: A Google Drive link will be provided soon for uploading your annotated data submission

## Step 2: Project Implementation (Starts 1st Week of April)

Apply deep learning-based object detection techniques to detect and classify vehicles in real-time road traffic videos and images using the annotated dataset from Step 1.

**Deep Learning Models for Object Detection & Classification in Road Traffic Images:**

* YOLO (You Only Look Once) Family and its recent version (Available on Ultralatics platform)

Versions: **YOLOv9, YOLOv10, YOLOv11, Yolo12.** Real-time performance, widely used in autonomous vehicles. Suitable for both image and video-based vehicle detection.

* SSD (Single Shot MultiBox Detector)

Lightweight, fast inference. Good for edge devices and mobile deployment.

* Faster R-CNN

High accuracy, suitable for detailed vehicle classification. Can be used if real-time speed is not critical.

* EfficientDet

Scalable, balances speed and accuracy. Useful for detecting vehicles of different scales and sizes.

* **DETR (Detection Transformer) E.g., CLIP, DERT, SWIN transformers etc.**

Transformer-based architecture, good for complex scenes. Suitable for traffic scenes with occlusions or overlapping vehicles.

* CenterNet or YOLOX

Anchor-free models, good performance on crowded road conditions. Detect center points and dimensions directly.

* You can add some more deep learning-based techniques which can be applied in object detection and classification in Canadian Vehicle dataset (CVD) if you are familiar with.

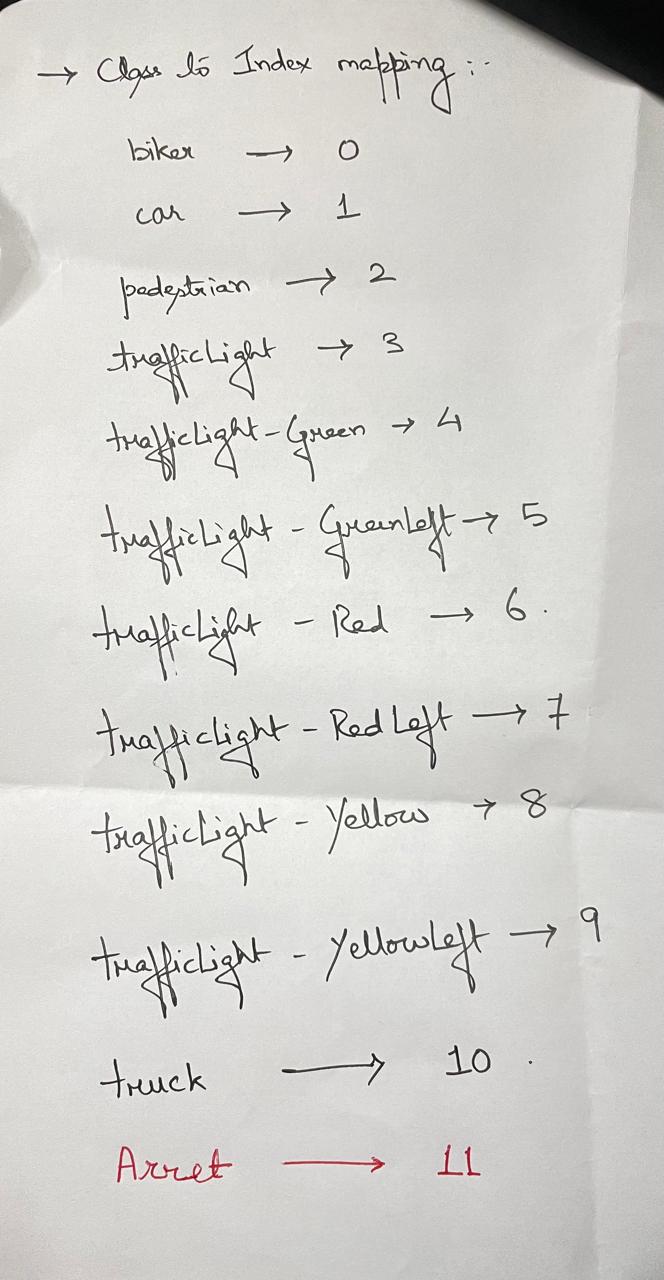
Meanwhile, start reading and implementation of object detection algorithms on image data on YouTube, GitHub and other sources.

(One of my Article link is shared to get basic knowledge about the project- <https://ieeexplore.ieee.org/abstract/document/10399478>

**Deadline:** April 30 (The project will be verified in your CV, and it will be graded (depending on completion of the project 10 marks).

Class Labels details are below:

Everyone has to follow the same labeling (save labeled images in YOLO format)



**Script to convert video into Images**

A screenshot of a computer

AI-generated content may be incorrect.