MINOR PROJECT REPORT

Group-15

Automatic Road Traffic Signs Detection using Object-Localization and Classification Model

Under the guidance of:-

Dr Rupali Verma

Dr Satnam Kaur

Submitted by:-

Aryan Arora (21103007)

Kamalpreet Singh (21103029)

Aditya Gupta (21103019)

Prajwal Sharma (21103020)

Department of Computer Science and Engineering

Punjab Engineering College, Chandigarh

DECLARATION

We hereby declare that the project work entitled "Automatic Detection and Road Traffic Signs Categorization using Object-Localization and Classification Model" is an authentic record of our work, carried out at Punjab Engineering College, Chandigarh as per requirements of "Minor Project" for the award of the degree of B.Tech. Computer Science Engineering, Punjab Engineering College (Deemed to be University), Chandigarh under the guidance of Dr. Rupali Verma and Dr.Satnam Kaur.

We further declare that the information has been collected from genuine & authentic sources and we have not submitted this project report to this or any other university for the award of diploma or degree of certificate examination.

Aditya Gupta (21103019)

Kamalpreet Singh (21103029)

Aryan Arora (21103007)

Prajwal Sharma (21103020)

CERTIFICATE

This is to certify that the project entitled "Automatic Detection and Road Traffic Signs Categorization using Object-Localization and Classification Model" by Aditya Gupta, Kamalpreet Singh, Aryan Arora and Prajwal Sharma is an authentic record of the work carried out under the supervision of Dr Rupali Verma and Dr Satnam Kaur, Computer Science and Engineering Department, Punjab Engineering College, Chandigarh in the partial fulfilment of the requirements as a part of Minor Project for the award of 03 credits in semester 5 of the degree of Bachelor of Technology in Computer Science and Engineering.

I certify that the above statement made by the students is correct to the best of my knowledge and belief.

Dr. Rupali Verma

Dr. Satnam Kaur

ACKNOWLEDGEMENT

We would like to take this opportunity to thank our college Punjab Engineering College, Chandigarh and the Department of Computer Science and Engineering for allowing us to work on this project.

We are immensely grateful to our project mentors Dr. Rupali Verma (Department of Computer Science and Engineering) and Dr. Satnam Kaur (Department of Computer Science and Engineering) whose continuous guidance, technical support, and moral support at times of difficulty helped us to achieve milestones in the given time. They have been a great source of knowledge and with them, this project was made.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support, and selfless help throughout the project work. It is a great pleasure to acknowledge the help and suggestions we received from the Department of Computer Science and Engineering.

We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement several times.

Aditya Gupta (21103019)

Kamalpreet Singh (21103029)

Aryan Arora (21103007)

Prajwal Sharma (21103020)

EXPLORE INNOVATE EXCEL

ABSTRACT

This project introduces a novel Automatic Road Traffic Signs Detection system, integrating advanced object-localization and classification models using YOLO. The primary objective is twofold: to provide assistance to individuals with visual impairments and Attention Deficit Hyperactivity Disorder (ADHD) and to contribute to the automation of vehicles. The system employs real-time auditory and haptic feedback mechanisms, providing crucial information to users for safer navigation in urban environments.

Simultaneously, the model seamlessly integrates into autonomous driving systems, ensuring vehicles can effectively interpret and respond to various traffic signs. The study emphasizes the dual impact of this technology, offering valuable insights into the convergence of assistive technology and intelligent transportation systems.

The project findings underscore the significance of leveraging innovation to bridge the accessibility gap for individuals with specific needs while advancing the capabilities of automated transportation. By facilitating clear communication of detected signs through auditory and haptic channels, the proposed system enhances the travel experience for users. Additionally, its integration with autonomous vehicles contributes to improved road safety and compliance with traffic regulations. This project contributes to the ongoing discourse on the symbiotic relationship between assistive technology and the evolution of intelligent transportation systems, fostering a more inclusive and connected society.

Table of Contents

Table of Figures

EXPLORE INNOVATE EXCEL

List of Tables

EXPLORE INNOVATE EXCEL

Introduction

Motivation

The inspiration behind undertaking the project on **"Automatic Road Traffic Signs Detection using Object-Localization and Classification Model"** is deeply rooted in the need for technological solutions to users along with bringing improvement in the Automation of Sign Detection Vehicles.

Addressing Safety Concerns: Our project aims to tackle the hurdles faced by those with fully or partially visually impaired by maintaining their focus while driving by providing information when a crucial traffic sign is encountered/passed by during driving.

Contributing to Accessibility: By enhancing accessibility for individuals with visual impairment or ADHD, we aim to positively impact the broader community's safety and create a transportation ecosystem that accommodates diverse needs.

Automation for Vehicles: In essence, our model provides crucial information to the autonomous vehicle's control system, enabling it to interpret and respond to road signs and signals effectively. This contributes to safer, more efficient, and rule-compliant automated driving.

In essence, the motivation behind this project is to merge cutting-edge technology with societal well-being, addressing the unique challenges faced by individuals, who are visually impaired or suffer from ADHD and contribute to a safer, more inclusive road experience for everyone.

EXPLORE INNOVATE EXCEL

BACKGROUND

Background

In recent years, the integration of advanced technologies into transportation systems has become a focal point for enhancing road safety and addressing the diverse needs of drivers. One particular area that demands attention is the intersection between road safety concerns and the challenges faced by individuals with partial visual impairment or Attention Deficit Hyperactivity Disorder (ADHD) while navigating through traffic environments.

Current Limitations in Road Sign Recognition: Traditional approaches to road sign recognition often fall short in providing real-time and accurate results, especially in scenarios where rapid decision-making is crucial. The existing challenges are exacerbated for individuals with disabilities as mentioned above,who have some cognitive load associated with processing visual information that can impede their ability to respond promptly to changing road conditions.

Technological Advancements and the Role of YOLO: The advent of computer vision and deep learning technologies presents an opportunity to revolutionise road sign detection and classification. You Only Look Once (YOLO) stands out as a real-time object detection system that has proven efficacy in processing images swiftly and accurately. By applying YOLO to the specific context of traffic road sign recognition, the project aims to overcome the limitations of traditional methods and provide a responsive and reliable solution.

Rationale for Project Development: Motivated by the pressing need to enhance road safety, this project seeks to leverage YOLO's capabilities to create an intelligent system for the detection and classification of traffic road signs. The goal is to develop a technology-driven solution that not only addresses the challenges faced by individuals but also contributes to the broader realm of road safety, creating a more inclusive and adaptive transportation environment.

Alignment with Societal Impact: By addressing the intersection of users and road safety through technological innovation, this project aligns with broader societal goals of creating transportation systems that accommodate diverse abilities. The development of an efficient and accessible traffic road sign detection system has the potential to positively impact not only individuals with ADHD but also contribute to the overall safety and inclusivity of our roadways.

Leveraging the current technology: A traffic sign detection model provides crucial information to the autonomous vehicle's control system, enabling it to interpret and respond to road signs and signals effectively. This contributes to safer, more efficient, and rule-compliant automated driving

In summary, the background of this project underscores the urgency to address the challenges faced by visually impaired individuals or individuals with ADHD in the context of road safety, highlighting the role of YOLO as a transformative technology in creating a more responsive and inclusive transportation landscape.

PEC

EXPLORE INNOVATE EXCEL

PROPOSED WORK

ALGORITHM

The YOLO family is a classic object detector. Since the first edition of this algorithm was published in 2015, it has achieved leading efficiency with a single-stage framework and has quickly become a mainstream detection algorithm. Through continuous research and innovations, different versions of YOLO have been proposed. The latest version is the YOLOv8 algorithm, which was open-sourced by Ultralytics in January 2023. This algorithm introduces new features and improvements, becoming the best model in the YOLO family. YOLOv8 includes four parts: Input, Backbone, Neck and Output.

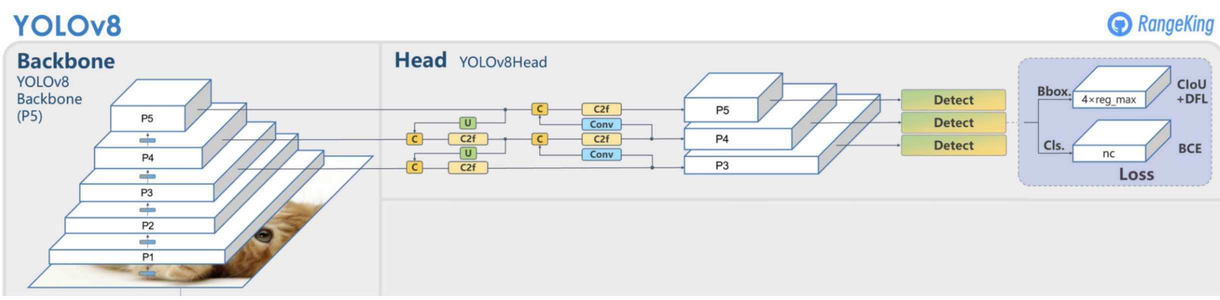


Figure 1: YOLOV8 Architecture

YOLO V8 ALGORITHMIC THEORY

Input mainly includes colour perturbation, spatial perturbation, Mosaic, and MixUp. Different numbers of pictures are spliced after processing a single one by combined data augmentation, which increases the multi-directional object perspective and also enriches the diversity of image backgrounds.

The backbone mainly includes the convolution layer, C2F layer and SPPF layer. The C2F structure is different from the C3 module in YOLOv5 which draws on the idea of ELAN (efficient layer aggregation network) in YOLOv7, which increases the efficiency of gradient propagation and enables the network to quickly converge. The SPPF layer maintains the design in YOLOv5.

Neck adopts a structure combining FPN (feature pyramid network) and PAN (path aggregation network). The features of adjacent layers are concatenated which are the input to the C2F module. As features are passed from top to bottom and bottom to top, high-level semantic features with underlying features are combined.

Output realises the decoupling of detection and classification. The bottom-level features are used to obtain information on small target objects, and the top-level features are the results of large target objects. Each detection layer outputs a result vector, which contains location and corresponding category information.

MODEL

Model Architecture: YOLOv8 (You Only Look Once version 8)

Task: Object Detection for Traffic Sign Recognition

Labels: The following labels were used for the detection task in this exact order

“Cycle_track”	(index 0)
“No_stopping”	(index 1)
“speedlimit_60_cars”	(index 2)
“pedestrian_crossing”	(index 3)
“Speedlimit_45_motorbikes”	(index 4)
“motor_vehicle_prohibited”	(index 5)
“no_parking”	(index 6)
“roundabout_ahead”	(index 7)

Figure 2: Roundabout Ahead

Figure 3: No Parking

Figure 4: Pedestrian Crossing

Figure 5: Cycle Track

Figure 6: No Parking

Figure 7: Motor Vehicle Prohibited

Figure 8: Speed Limit 45

Figure 9: Speed Limit 60 (Cars)

Dataset Information:

Training Dataset:
Number of Images and corresponding labels: 10,000

We have accumulated and annotated the entire dataset manually by using recording data from all over Chandigarh which we accumulated ourselves, annotation of bounding boxes was done using the aid of ***labelImg*** *software.*

Validation Dataset:
Number of Images and corresponding labels: 1000, annotated with the specified traffic sign labels

Data Augmentations:
Techniques like *Random Flip*, and *Random Crop* are used in the "config.yaml" file (used for model configuration setting) to augment the dataset and improve model generalisation. This allows for the model to learn the dataset with more variance

```yaml
# Data augmentation options
augment:
  # Enable or disable data augmentation
  enabled: True

  # Random image flipping
  flip: 0.5

  # Random image rotation
  rotate: 30

  # Random image brightness adjustment
  brightness: 0.5

  # Random image contrast adjustment
  contrast: 0.5

  # Random image saturation adjustment
  saturation: 0.5

  # Random image hue adjustment
  hue: 0.1

  # Random image crop
  crop: 0.5

  # Random image mosaic
  mosaic: 0.5

  # Random image mixup
  mixup: 0.5
```

Figure 10: Augmentation configuration in config.yaml

Data Preprocessing: After data accumulation, We also ensured uniformity in image sizes, aspect ratios, and labelling format.

Model Selection: Chose YOLOv8 for its efficiency in real-time object detection tasks and well-suited architecture for handling multiple object classes.

Model Training: Trained the YOLOv8 model using the prepared dataset. Configured the model to recognize the specified traffic sign labels.

Hyperparameter Tuning: The highlight adjusted hyperparameters are as follows

learning rate = 0.01
batch size = 16
optimizer = Adam (short for Adaptive Moment Estimation)
epochs =100

Validation: Utilised a separate validation dataset to assess the model's performance and prevent overfitting.

A train-test split of 10:1 was used for both effective training and validation of the model, and for evaluating mean average precision (MAP) performance for the trained YOLO model weights.

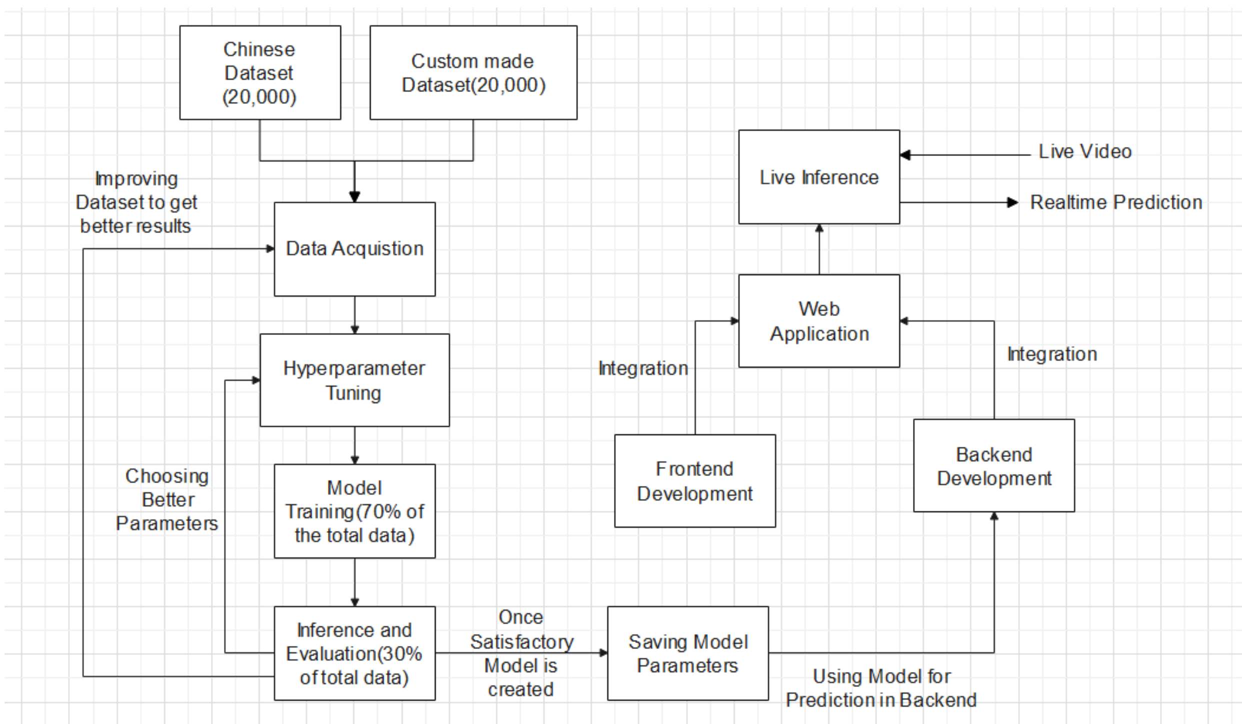
Fine-Tuning: Iteratively fine-tuned the model based on validation results to improve accuracy and generalisation.

Evaluation: Evaluated the model's performance using standard metrics (e.g., Mean Absolute Precision scores) to ensure it meets the project requirements.

Deployment: Deployed the trained YOLOv8 model for real-time recognition applications using a web app backend pipeline.

Documentation: Documented the entire process, including dataset details, model architecture, training strategy, and evaluation results.

Results: Achieved a robust object detection model capable of accurately recognizing specified traffic sign labels in diverse scenarios.



METHODOLOGY DIAGRAM

In Iteration 1, the YOLOv6 Model was used to gain understanding of the YOLO Model and a subset of about 1000 handpicked images from the Chinese traffic sign dataset (CCTSDB) consisting of roughly 20,000 images was used for both Training/Testing with a 70-30 ratio.

For the Hyperparameters value, the Model creation was done for the value of Batch size 16 and the Number of Epochs set to 50. At the end of the 50th Iteration of the Model, the important parameters value for the Model were as follows-

Epoch lr iou_loss dfl_loss cls_loss
49/49 0.0001976 0.2402 0 0.5126: 100%|██████████| 69/69 [00:26<00:00, 2.58it/s]

```
Results saved to runs/train/exp
Epoch: 47 | mAP@0.5: 0.3581632747157688 | mAP@0.50:0.95: 0.20223589596767047

     Epoch        lr  iou_loss  dfl_loss  cls_loss
     48/49 0.0002045    0.2387         0    0.5093: 100%|██████████| 69/69 [00:26<00:00,  2.57it/s]

     Epoch        lr  iou_loss  dfl_loss  cls_loss
     49/49 0.0001976    0.2402         0    0.5126: 100%|██████████| 69/69 [00:26<00:00,  2.58it/s]
Inferencing model in train datasets.: 100%|█████████████████████████| 10/10 [00:06<00:00,  1.49it/s]

Evaluating speed.
```

Figure 11: 50th epoch of Model 1

lr(Final Learning rate)=0.0001976
iou_loss=0.2402
dfl_loss=0
cls_loss=0.5126
The value of dfl_loss being 0 implies the model created was able to correctly identify the bounding box in the input image.

After being able to gain sufficient knowledge about working with YOLOv6 and able to train a model, we moved into 2nd Iteration.

EXPLORE INNOVATE EXCEL

In the 2nd Iteration, the YOLOv6 Model was used with the complete Chinese Traffic Sign Dataset(CATSDB) which consisted of roughly 20,000 images used for both Training/Testing in a 70-30 ratio.
To avoid overfitting of Model, the number of epochs was reduced to 4. But it has a disadvantage of its own,as if the number of epochs is too small, the model may not learn the underlying patterns in the data, resulting in underfitting.But if the number of epochs are too high,the model may overfit on the data.

So we wish to avoid Model Training to stop before the Model converges due to low epochs, and also avoid the problem of overfitting on data for high number of epochs,we tried to experiment with the epoch value of 5.

A good rule of thumb is to start with a value that is 3 times the number of columns in our data. Larger batch sizes result in faster progress in training, but don't always converge as fast. Smaller batch sizes train slower, but can converge faster.So for Model Training during this Iteration, standard batch size value of 16 was used.

```
Epoch       lr  iou_loss  dfl_loss  cls_loss
    4/4  0.0006785    0.4343        0    0.6843: 100% 1023/1023 [16:03<00:00,  1.06it/s]
```

```
Results saved to runs/train/exp
Epoch: 2 | mAP@0.5: 0.6463669430631847 | mAP@0.50:0.95: 0.351456980304035

     Epoch        lr   iou_loss  dfl_loss  cls_loss
       3/4   0.001114    0.4584         0    0.7062: 100% 1023/1023 [16:12<00:00,  1.05it/s]

     Epoch        lr   iou_loss  dfl_loss  cls_loss
       4/4  0.0006785    0.4343         0    0.6843: 100% 1023/1023 [16:03<00:00,  1.06it/s]
Inferencing model in train datasets.: 100% 47/47 [00:52<00:00,  1.11s/it]
```

Figure 12: 5th epoch of Model 2

At the end of the 5th Iteration of the Model,the important parameters value for the Model were as follows-
lr(Final learning rate)=0.0006785
iou_loss=0.4343
dfl_loss=0
cls_loss=0.6843

Now with the introduction of YOLOv8,due to new backbone network, a new anchor-free detection head, and a new loss function along with overcoming the YOLOv6 disadvantages of low confidence score for occurrence of multiple objects in an image,we decided to transition to YOLOv8 Model for our purpose.
Equipped with the knowledge of handling experience of YOLOv6 from the past two iterations, we still decided to feasibility study of whether we would be able to create a Model on YOLOv8.

So In third iteration, YOLOv8 Model was used with the handpicked dataset of 1100 images which were used in the 1st iteration. Hyperparameters Batch size and epoch value were set to 16 and 10 respectively.
The parameters value for YOLOv8 Model obtained after the 10th Epoch were as follows-

Epoch GPU_mem box_loss cls_loss dfl_loss Instances Size
10/10 0G 0.7219 0.6259 0.8474 16 640: 100%|██████████| 69/69 [12:23<00:00, 10.
Class Images Instances Box(P R mAP50 mAP50-95): 100%|██████████| 10/10 [00:55
all 309 738 0.567 0.194 0.221 0.118

```
Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
 9/10         0G     0.7541     0.6923     0.8468         17        640: 100%|██████████| 69/69 [22:08<00:00, 19.
               Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 10/10 [00:57
                 all        309        738      0.617      0.183      0.227      0.119

Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
10/10         0G     0.7219     0.6259     0.8474         16        640: 100%|██████████| 69/69 [12:23<00:00, 10.
               Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 10/10 [00:55
                 all        309        738      0.567      0.194      0.221      0.118
```

Figure 13: 10th epoch of Model 3

So here,the following values were obtained-
cls_loss=0.6259
dfl_loss=0.8474

The Model summary during the Testing was obtained as-

```
Validating runs\detect\train5\weights\best.pt...
Ultralytics YOLOv8.0.206  Python-3.10.6 torch-2.0.1+cpu CPU (11th Gen Intel Core(TM) i5-1135G7 2.40GHz)
Model summary (fused): 168 layers, 3006233 parameters, 0 gradients, 8.1 GFLOPs
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 10/10 [00:56
                   all        309        738      0.522      0.201      0.232      0.128
               Warning        309        148      0.235      0.027     0.0206     0.0123
           Prohibitory        309        571      0.742      0.156      0.276      0.147
             Mandatory        309         19      0.589      0.421      0.401      0.223
Speed: 2.0ms preprocess, 142.6ms inference, 0.0ms loss, 0.7ms postprocess per image
Results saved to runs\detect\train5
```

Figure 14: Summary Table for Model 3

Now in 4th Iteration, we created a custom dataset for our purpose by recording the roads of Chandigarh, converting the videos recorded into Image frames and then manually labelling the useful Images according to the Traffic sign present in the dataset.

For the 4th Iteration YOLOv8 was used with the default hyperparameter value for batch size being 16 and a custom value for epoch being 100 was set for Training the YOLOv8 Model.

But after the 9th Epoch, the Mean Average Precision value started to drop so to avoid the overfitting the Model Training was stopped.

RESULT AND DISCUSSION

Model Training :

The model is trained on the accumulated and prepared traffic-sign dataset, and then trained for multiple epochs to attain better precision cores. The epoch details are mentioned below.

```
[1]: from ultralytics import YOLO

[2]: import yaml                    # for yaml files
     import torch
     from PIL import Image
     import os
     import cv2
     import time

[8]: !pwd

     /Users/aryanarora/Desktop

[12]: !cd /Users/aryanarora/Desktop/data_yoloV8

[13]: !pwd

     /Users/aryanarora/Desktop

[25]: !yolo task=detect mode=train model=yolov8s.pt data=/Users/aryanarora/Desktop/data_yoloV8/custom.yaml epochs=100

     New https://pypi.org/project/ultralytics/8.0.222 available 😃 Update with 'pip install -U ultralytics'
     Ultralytics YOLOv8.0.221 🚀 Python-3.8.18 torch-2.1.1 CPU (Apple M1)
     engine/trainer: task=detect, mode=train, model=yolov8s.pt, data=/Users/aryanarora/Desktop/data_yoloV8/custom.yaml, epochs=100,
     patience=50, batch=16, imgsz=640, save=True, save_period=-1, cache=False, device=None, workers=8, project=None, name=train12, e
     xist_ok=False, pretrained=True, optimizer=auto, verbose=True, seed=0, deterministic=True, single_cls=False, rect=False, cos_lr=
     False, close_mosaic=10, resume=False, amp=True, fraction=1.0, profile=False, freeze=None, overlap_mask=True, mask_ratio=4, drop
     out=0.0, val=True, split=val, save_json=False, save_hybrid=False, conf=None, iou=0.7, max_det=300, half=False, dnn=False, plots
     =True, source=None, vid_stride=1, stream_buffer=False, visualize=False, augment=False, agnostic_nms=False, classes=None, retina
     _masks=False, show=False, save_frames=False, save_txt=False, save_conf=False, save_crop=False, show_labels=True, show_conf=Tru
     e, show_boxes=True, line_width=None, format=torchscript, keras=False, optimize=False, int8=False, dynamic=False, simplify=Fals
     e, opset=None, workspace=4, nms=False, lr0=0.01, lrf=0.01, momentum=0.937, weight_decay=0.0005, warmup_epochs=3.0, warmup_momen
     tum=0.8, warmup_bias_lr=0.1, box=7.5, cls=0.5, dfl=1.5, pose=12.0, kobj=1.0, label_smoothing=0.0, nbs=64, hsv_h=0.015, hsv_s=0.
     7, hsv_v=0.4, degrees=0.0, translate=0.1, scale=0.5, shear=0.0, perspective=0.0, flipud=0.0, fliplr=0.5, mosaic=1.0, mixup=0.0,
     copy_paste=0.0, cfg=None, tracker=botsort.yaml, save_dir=/Users/aryanarora/runs/detect/train12
     Overriding model.yaml nc=80 with nc=8
```

Figure 15: YOLOV8 training

Figure 16: Model Epochs

Results:

The model was then tested on a set of validation traffic sign data, unseen by the model. The resulting runs gave the following results.

```
  Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  6/100         0G      1.467     0.8796      1.204          4       640: 1
             Class     Images  Instances       Box(P          R      mAP50  m
               all       1059       1344      0.774      0.844      0.862      0.483

  Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  7/100         0G      1.451     0.8928      1.177          6       640: 1
             Class     Images  Instances       Box(P          R      mAP50  m
               all       1059       1344      0.857      0.867      0.883      0.519

  Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  8/100         0G      1.441     0.8593      1.187         11       640: 1
             Class     Images  Instances       Box(P          R      mAP50  m
               all       1059       1344      0.857      0.858      0.887      0.514

  Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
  9/100         0G      1.428     0.8175      1.188          5       640: 1
             Class     Images  Instances       Box(P          R      mAP50  m
               all       1059       1344      0.905      0.839      0.901      0.544

  Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
 10/100         0G      1.412     0.8049       1.18          8       640: 1
             Class     Images  Instances       Box(P          R      mAP50  m
               all       1059       1344      0.906      0.848      0.896      0.531

  Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
 11/100         0G      1.416     0.7692      1.158          6       640: 1
             Class     Images  Instances       Box(P          R      mAP50  m
               all       1059       1344      0.888      0.869      0.902      0.516

  Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
 12/100         0G      1.429     0.7955       1.16         39       640:  ^C
```

Figure 17: Run Directory / Folder Structure

- Here, best.pt is the best attained trained model out of all runs, based on its performance on the validation set.

- last.pt represent the most recently achieved updates in the model weights

The localization and classification results of the given images are displayed below:

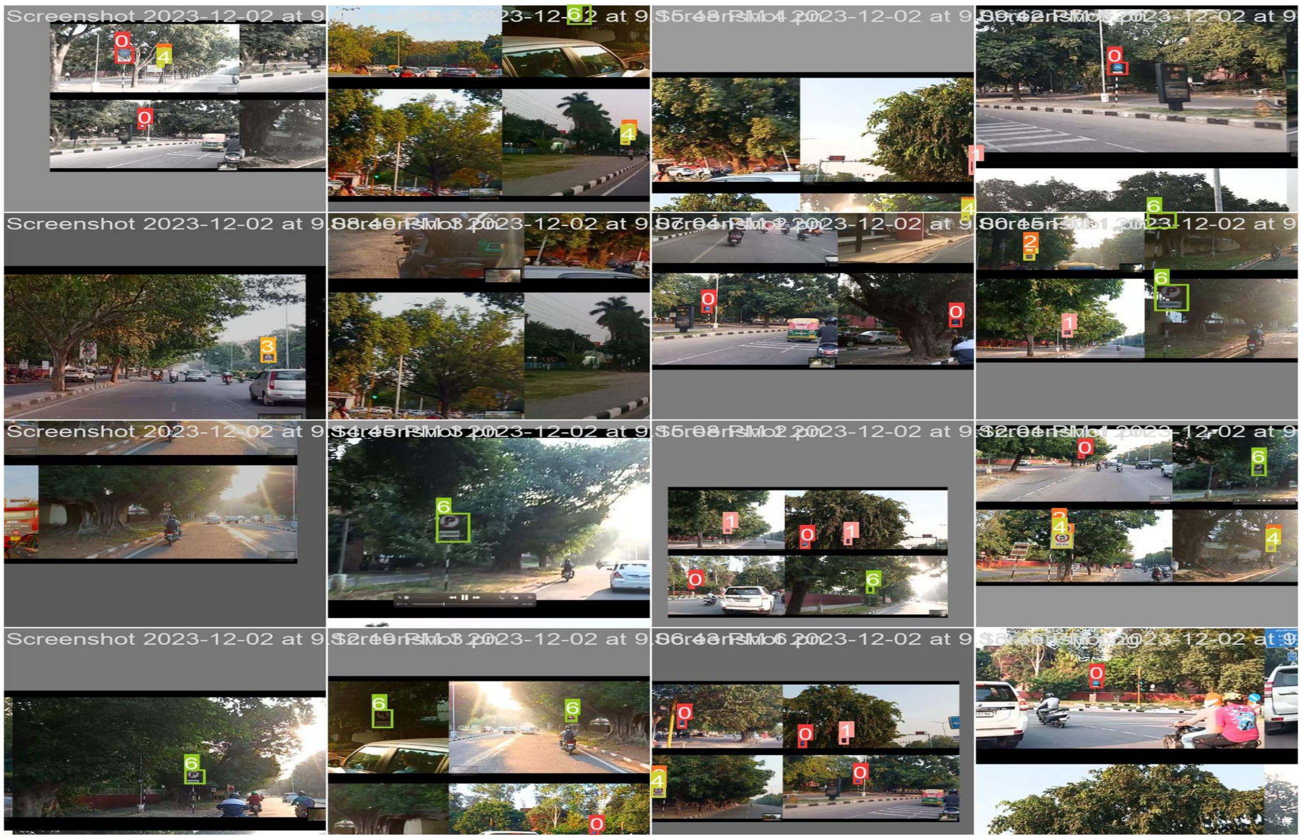


Figure 18: Model Inferences (1)

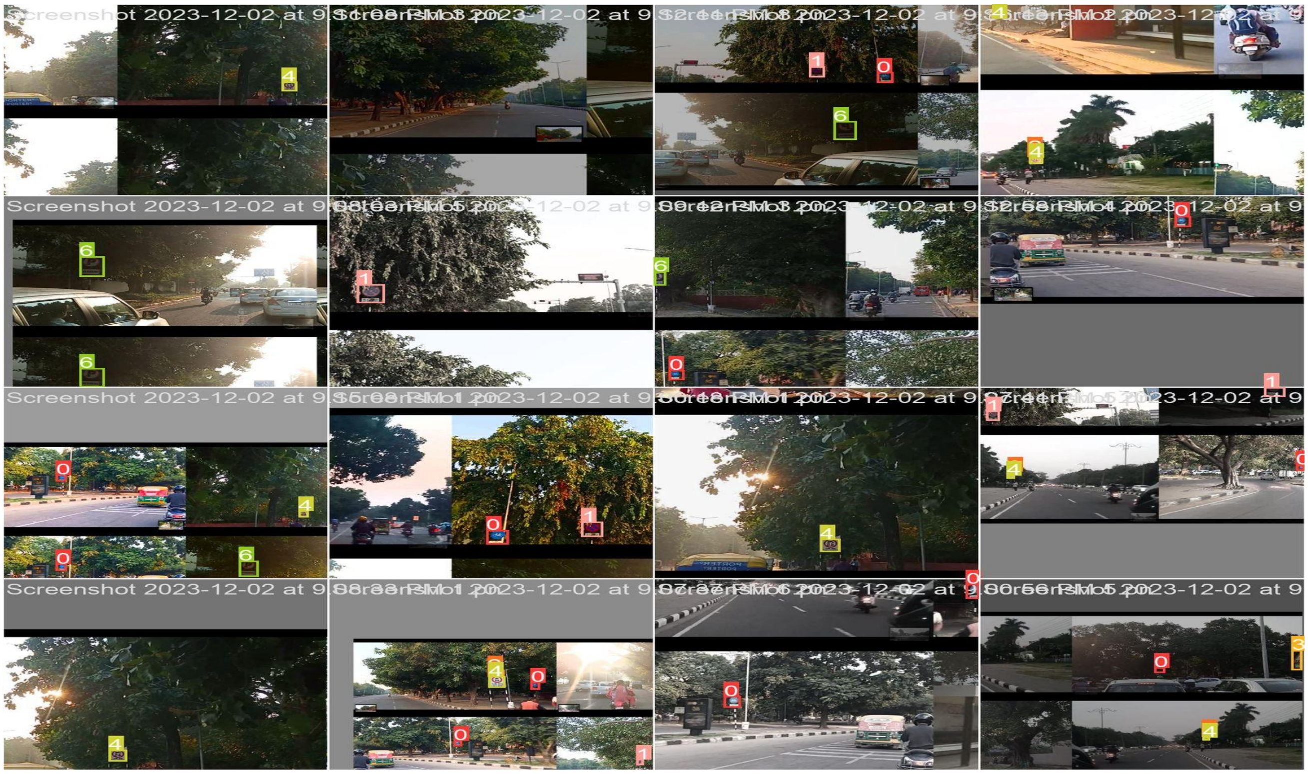


Figure 19: Model Inferences (2)

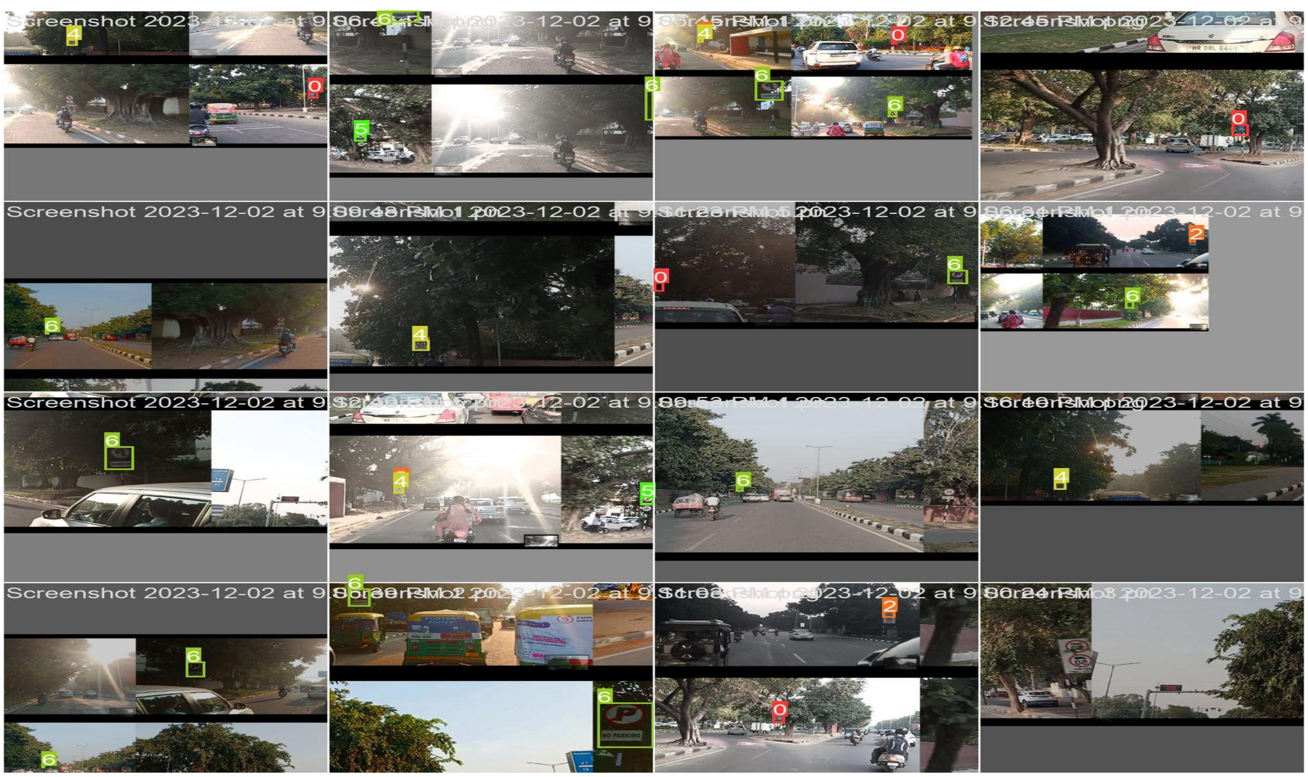


Figure 20: Model Inferences (3)

Model Validation and Evaluation:

Evaluation Metrics:

Following were the model evaluation statistics received from the trained model.

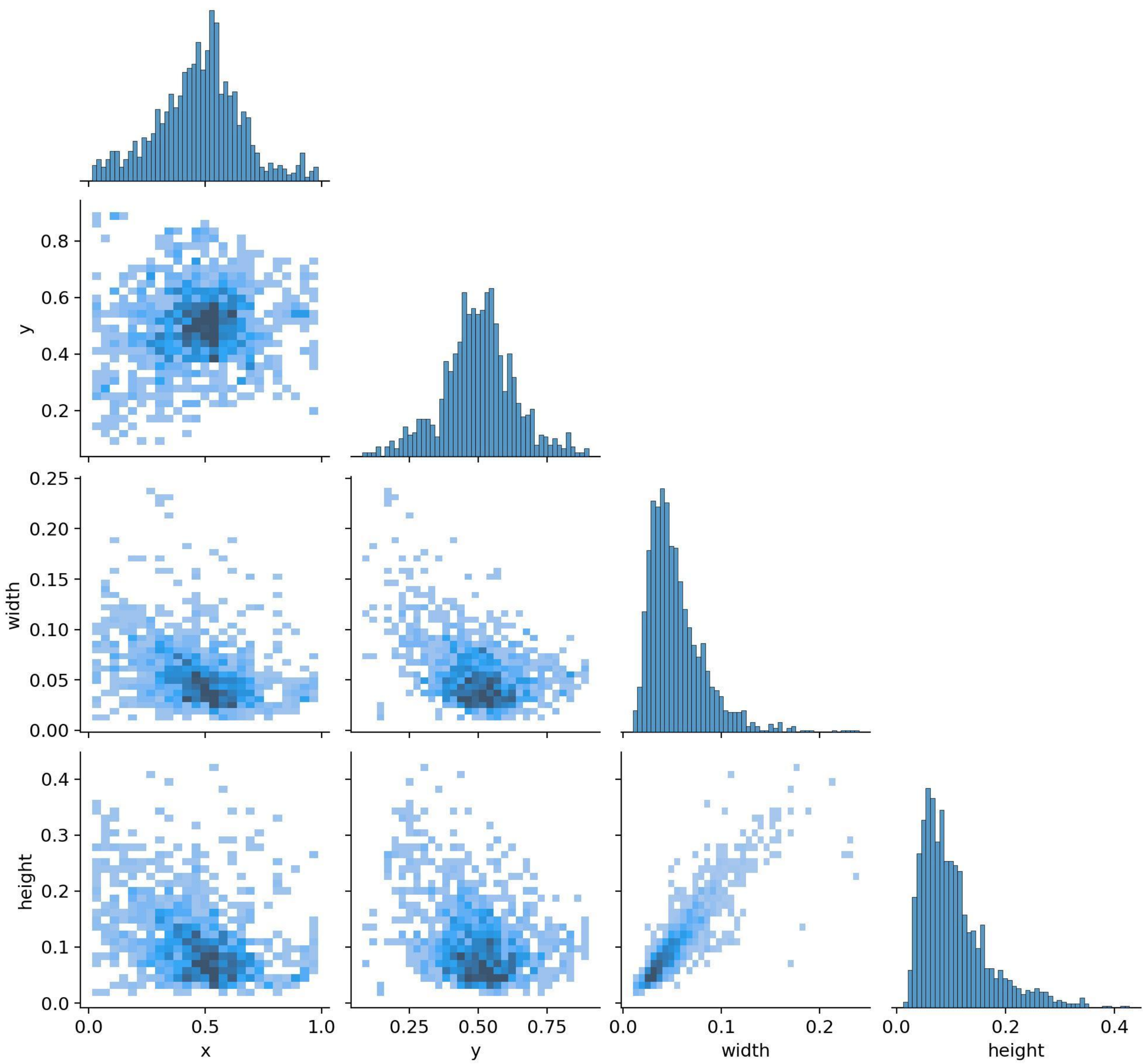


Figure 21: Labels Correlogram

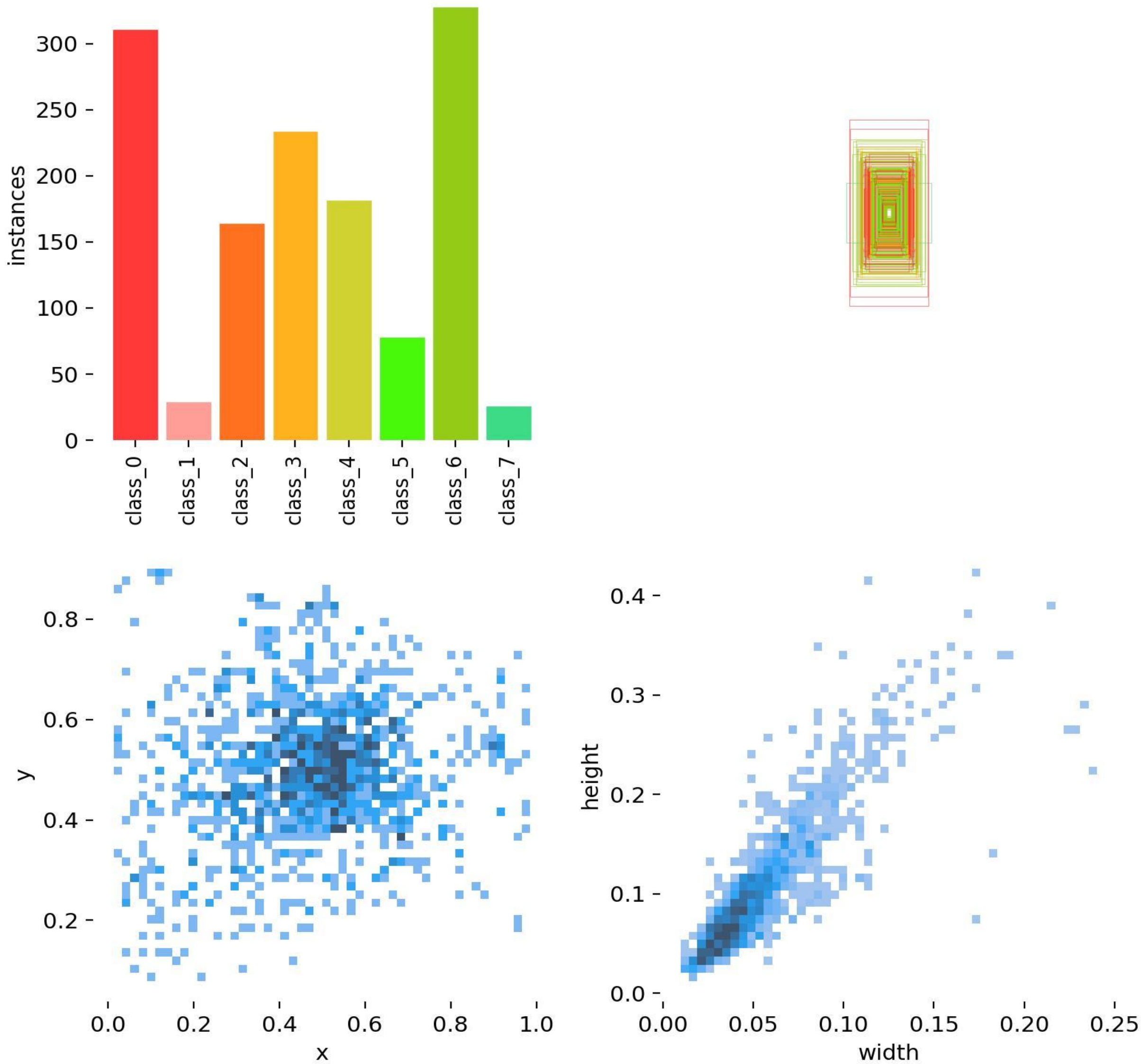


Figure 22: Labels Scatter Plot and Histogram Data

epoch	train/box_loss	train/cls_loss	train/dfl_loss	metrics/precision(B)	metrics/recall(B)	metrics/mAP50(B)	metrics/mAP50-95(B)
1	1.9616	6.4162	1.3935	0.68833	0.54239	0.60767	0.32918
2	1.5559	1.3512	1.2242	0.67584	0.68191	0.68739	0.36756
3	1.5385	1.2265	1.2315	0.78904	0.73153	0.7732	0.45647
4	1.5063	1.0268	1.2384	0.76558	0.79994	0.8296	0.47052
5	1.5241	0.99002	1.2052	0.81926	0.8134	0.84809	0.4847
6	1.4668	0.8796	1.2036	0.77406	0.84424	0.86222	0.48342
7	1.4509	0.89282	1.177	0.85667	0.867	0.88297	0.51911
8	1.4413	0.85926	1.1865	0.85727	0.85818	0.88742	0.5135
9	1.4278	0.81755	1.1885	0.90548	0.83896	0.90128	0.54398
10	1.4123	0.80489	1.1798	0.90632	0.84832	0.89629	0.53076
11	1.4159	0.76922	1.1579	0.88818	0.86899	0.90229	0.51571

Table 1: Final Evaluation Results (1)

val/box_loss	val/cls_loss	val/dfl_loss	lr/pg0	lr/pg1	lr/pg2
1.5508	1.7491	1.264	0.00027352	0.00027352	0.00027352
1.5366	1.2751	1.2677	0.00054573	0.00054573	0.00054573
1.4697	1.0768	1.256	0.00081244	0.00081244	0.00081244
1.47	0.95924	1.2414	0.00080826	0.00080826	0.00080826
1.4713	0.84942	1.2429	0.00080826	0.00080826	0.00080826
1.4469	0.87088	1.2225	0.00080001	0.00080001	0.00080001
1.4279	0.80227	1.2246	0.00079177	0.00079177	0.00079177
1.4015	0.76802	1.2103	0.00078352	0.00078352	0.00078352
1.3843	0.74588	1.2009	0.00077527	0.00077527	0.00077527
1.4288	0.73842	1.223	0.00076703	0.00076703	0.00076703
1.4179	0.71183	1.2087	0.00075878	0.00075878	0.00075878

Table 2: Final Evaluation Results (2)

FUTURE SCOPE

CONCLUSION

The YOLO model demonstrated remarkable performance in detecting and classifying traffic signs across various scenarios. Precision and recall scores were consistently improved with each epoch, indicating a high level of accuracy in both localization and classification tasks. The model successfully identified a wide range of traffic sign types, including stop signs, speed limits, and warning signs.

FUTURE WORK

While the current project achieves significant milestones, there is always room for improvement. This section discusses potential avenues for refining existing models, optimising hyperparameters,

Model Enhancement:

- Define better hyperparameters for YOLOv8 to increase accuracy.

Text-to-Audio Conversion:

- Implement NLP for real-time spoken descriptions.

Multilingual Recognition:

- Train on diverse language datasets for recognizing signs in multiple languages.
- Address linguistic diversity on Indian roads.

Dataset Variation:

- Diversify training datasets for robustness in varied scenarios.
- Capture weather, lighting, and infrastructure variations.

Collaboration with Traffic Systems:

- Integrate with traffic management systems for real-time updates.
- Enhance system relevance and accuracy with dynamic traffic information.

Website Integration:

- Extend accessibility by integrating with a dedicated website.
- Allow remote user access, expanding usability.

Continuous Improvement:

- Embrace emerging YOLO technologies for more enhancements

In summary, the project's future focuses on precision enhancement, inclusivity through audio integration, adaptability to diverse linguistic contexts, robust dataset utilisation, collaboration for real-time updates, accessibility features, efficient edge computing, user interface optimization, website integration, and a commitment to continuous improvement in the dynamic field of AI.

REFERENCES

[1] Srivastava, Vivek, Sumita Mishra, and Nishu Gupta. "Automatic Detection and Categorization of Road Traffic Signs using a Knowledge-Assisted Method." *Procedia Computer Science* 218 (2023): 1280-1287.

[2] Fuermaier, Anselm B M et al. "Driving and attention deficit hyperactivity disorder." *Journal of neural transmission (Vienna, Austria : 1996)* vol. 124,Suppl 1 (2017): 55-67. doi:10.1007/s00702-015-1465-6

[3] Huang, Zhongjie, et al. "Research on Traffic Sign Detection Based on Improved YOLOv8." *Journal of Computer and Communications* 11.7 (2023): 226-232.

[4] Avramović, Aleksej, Domen Tabernik, and Danijel Skočaj. "Real-time large scale traffic sign detection." *2018 14th Symposium on Neural Networks and Applications (NEUREL)*. IEEE, 2018.

[5] Y. Sun, P. Ge and D. Liu, "Traffic Sign Detection and Recognition Based on Convolutional Neural Network," 2019 Chinese Automation Congress (CAC), Hangzhou, China, 2019, pp. 2851-2854, doi: 10.1109/CAC48633.2019.8997240.

[6] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[7] XiongFei Liu and Fan Xiong 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* 787 012034
DOI 10.1088/1757-899X/787/1/012034