Walchand College of Engineering, Sangli

(An Autonomous Institute)

Department of Computer Science and Engineering

TY CSE Mini Project 1

Report on

Real-Time Traffic Signs Detection with Voice Alert

Submitted by

Mr. Anurag Takalkar	(21510016)
Mr. Onkar Yemul	(21510017)
Mr. Omkar Auti	(21510042)

Under the guidance of

Prof. N. L. Mudegol
Department of Computer Science and Engineering,
Walchand College of Engineering, Sangli

2023 - 2024

Walchand College of Engineering, Sangli
(An Autonomous Institute)

Department of Computer Science and Engineering

CERTIFICATE

This is to certify that the Project Report entitled, **"REAL TIME TRAFFIC SIGN DETECTION WITH VOICE ALERT"** submitted by Mr. Anurag Takalkar, Mr. Omkar Auti, Mr. Onkar Yemul to Walchand College of Engineering Sangli, India, is a record of Bonafide project work of course **Mini-Project I 6CS341** carried out under our supervision and guidance and is worthy of consideration for the **Mini-project I** in Computer Science & Engineering of the Institute.

Prof. N. L. Mudegol		**Dr. M. A. Shah**
Guide	**External examiner**	**Head Of Department**
Department of Computer Science and Engineering,		Department of Computer Science and Engineering,
Walchand College of Engineering, Sangli		Walchand College of Engineering, Sangli

Acknowledgement

We would like to express our sincere gratitude to all those who have contributed to the successful completion of this mini-project on "**Real-time Traffic Sign Detection with Voice Alert.**" This project has been an exciting and challenging journey, and We are deeply appreciative of the support and guidance received.

We extend our heartfelt thanks to our project guide **Prof. N. L. Mudegol,** whose expertise and encouragement played a pivotal role in shaping the project. Their valuable insights, continuous feedback, and unwavering support were instrumental in overcoming obstacles and achieving the project's objectives.

We are also thankful to **Department of Computer Science and Engineering, Walchand College of Engineering**, **Sangli** for providing the necessary resources, infrastructure, and environment conducive to learning and innovation.

We want to acknowledge the contributions of our peers and colleagues who provided valuable suggestions, constructive criticism, and moral support throughout the development process. Their collaboration has enhanced the overall quality of the project.

Last but not least, we would like to express our gratitude to our families and friends for their understanding, encouragement, and patience during the project's ups and downs.

This project would not have been possible without the collective effort and support of all those mentioned above. Thank you for being part of this journey.

Anurag Takalkar
Omkar Auti
Onkar Yemul

Department of Computer
Science and Engineering,

Walchand College of
Engineering, Sangli

Declaration

We hereby declare that work presented in this project report titled **"REAL TIME TRAFFIC SIGN DETECTION WITH VOICE ALERT"** submitted by us in the fulfillment of the **Mini- Project I** in the **Department of Computer Science & Engineering, Walchand College of Engineering, Sangli** is an authentic record of our project work carried out under the guidance of Prof. N. L. Mudegol.

Date:

Place: Sangli

Anurag Vijay Takalkar

Omkar Rajesh Auti

Onkar Anand Yemul

Table of Contents

Acronyms

- YOLO You Only Look Once
- TSDR Traffic Sign detection and recognition

Abstract

Every year, approximately 1.3 million lives are tragically cut short due to traffic accidents, with an additional 20 to 50 million individuals enduring non-fatal injuries, many of whom are left with long-lasting disabilities. These incidents not only inflict immeasurable human suffering but also inflict substantial economic burdens on individuals, families, and entire nations. The economic toll includes not only direct medical expenses but also the considerable cost of lost productivity for those killed or injured in accidents. Moreover, family members often need to take time off work or school to care for the injured, further compounding the economic impact.

Remarkably, road traffic accidents take a heavy toll on national economies, with most countries enduring losses equating to 3% of their gross domestic product (GDP). This staggering financial burden underscores the urgent need for innovative solutions to mitigate road accidents and enhance safety.

In light of these sobering statistics, our project aims to develop a real-time traffic sign detection system capable of processing both image and video feeds. This system will not only identify and interpret traffic signs but also provide textual descriptions and audio notifications in real time. By enhancing road safety and accessibility, our technology seeks to contribute to the reduction of accidents and the preservation of lives, while also addressing the economic repercussions of road traffic injuries on both individual and national scales.

Introduction

Introducing a revolutionary real-time traffic sign detection system with voice alerts – a cutting-edge solution poised to redefine road safety. Leveraging advanced computer vision technology, this system utilizes state-of-the-art algorithms to instantaneously recognize and interpret traffic signs, providing drivers with crucial information in the blink of an eye.
The primary objective of this innovation is to enhance driver awareness and compliance with traffic regulations. As the system identifies various road signs, such as speed limits, stop signs, and directional indicators, it promptly generates voice alerts, delivering real-time guidance to the driver. This dynamic interaction minimizes the risk of human error and ensures a heightened level of responsiveness on the road.
With seamless integration into vehicle systems, this technology operates in real-time, adapting to changing road conditions and signage. The voice alert feature adds an extra layer of safety, allowing drivers to maintain focus on the road while receiving essential information audibly. This system represents a paradigm shift in road safety, combining artificial intelligence, computer vision, and auditory feedback to create a comprehensive solution for safer and more informed driving experiences.

Related work

A significant amount of work has already been carried out in the past with respect to TSDR. A few of them have been listed below:

Jensen *et al.* [1] applied a state-of-the-art, to detect traffic lights, a real-time object detection system using You Only Look Once (YOLO) on the public LISA Traffic Light dataset which is accessible through the VIVA-challenge. For day sequences, the AUC of the YOLO object detection is 90.49 %. Santos *et al.* [2] implemented a real-time TSDR system. Their study offered a comparative analysis on detection and recognition performance of several algorithms based on accuracy and processing speed.

In [3] AlexNet structure of CNN is used in which the architecture contains eight layers. The first five layers are convolutional layers and the latter three are all connected layers. The accuracy of this architecture comes out to be 92.63%. Also, the GoogleNet architecture is implemented in [3] which helps in working with large data and a high number of parameters. However it comes with an issue that the large data causes network overfitting decreasing the accuracy to just 80.5%. VGG CNN is proposed in [4] which has a significantly superior performance as compared to other available architectures. The number of parameters in this approach is considerably reduced in order to optimise and speed up the calculation. The network also includes the BN (batch normalisation) and GAP (global average pooling) layers, which help to improve accuracy without increasing the number of parameters.

The position of the traffic sign can be determined with a 1 metre accuracy using a single colour camera and a high precision GNSS (global navigation satellite systems) receiver. Another use of GPS as suggested in [5], is determining the style of driving where GPS data is collected from the mobile phone of a person while also detecting the Traffic Signs in the locality. It helps in classifying the style of driving as safe or aggressive.

1 Problem statement

To develop a comprehensive real-time traffic sign detection system with voice alert in three languages (English, Hindi, Marathi) capable of processing both image and video feeds. This system should be equipped to identify and recognize traffic signs from live visual data streams, and it should provide textual descriptions and audio/voice notifications for increased safety and accessibility. The proposed system also contains a section where the vehicle driver is alerted about the traffic signs in the near proximity which helps them to be aware of what rules to follow on the route. The aim of this system is to ensure the safety of the vehicle's driver, passengers, and pedestrians.

2 Objectives

1. **Data Collection and Preprocessing:** To collect a diverse and curated dataset of Indian traffic sign images, covering variations in color, shape, size, and background and preprocess the dataset to normalize colors, resize images to a consistent format, and apply data augmentation techniques to enhance the model's robustness.

2. **Model Training and Optimization:** To split the dataset into training, validation, and testing sets to train and evaluate the model's performance and implement optimization techniques.

3. **Traffic Sign Detection System Implementation:** To develop a model that allows users to detect real time traffic signs which will be highly accurate, precise and robust.

4. **User Interface Design**: Create an intuitive and user-friendly interface that displays real-time information about detected traffic signs. Design the user interface to be easily understandable, with clear visuals and concise information, enhancing the overall user experience.

3 Methodology

Model's Working

The presented model has trained on the YOLO (Version 5) object detection algorithm, which was released by Ultralytics in June 2020 and currently being state of the art among other algorithms. YOLO (You only look once) is a creative convolutional neural network (CNN) for doing object detection in real-time with high accuracy. This algorithm applies a single neural network to the full image, and then divides the image into regions and predicts bounding boxes and probabilities for each part. The predicted probabilities weight these bounding boxes. The algorithm "you only look once" at the image in the sense that it requires only one forward propagation pass through the neural network to make predictions. After non-max suppression (which makes sure that the object detection algorithm only detects each object once), it then outputs recognized objects.

Capabilities:

1. Real-Time detection and classification upto 30 fps

2. Accuracy as high as 98%

3. Robust model - detects in all weather conditions

5. Able to detect a wide variety of signs (for ex. Indian, European, American traffic signs)

Traffic Signs detected by our model.

Compulsory Ahead, Compulsory Keep Left, Compulsory Left Turn, Compulsory Right Turn, Cycle Crossing, Compulsory Cycle Track, Dangerous Dip, Falling Rocks, Horn prohibited, Humpy Road, Left hair pin bend, Left hand curve, Left Reverse Bend, Left turn prohibited, Major road ahead, Men at work, Motor vehicles prohibited, Narrow bridge, Narrow road ahead, Straight prohibited, No parking, No stopping, One way sign, Overtaking prohibited, Pedestrian crossing, Pedestrian prohibited, Restriction ends sign, Right hair pin bend, Right hand curve, Right Reverse Bend, Right turn prohibited, Road wideness ahead, Roundabout, School ahead, Slippery road, Compulsory sound horn, Speed limit, Staggered intersection, Steep ascent, Steep descent, Stop, Truck prohibited, Compulsory turn left ahead, Compulsory right turn ahead, T-intersection, U-turn prohibited, Y-intersection, Compulsory Keep Right, Parking

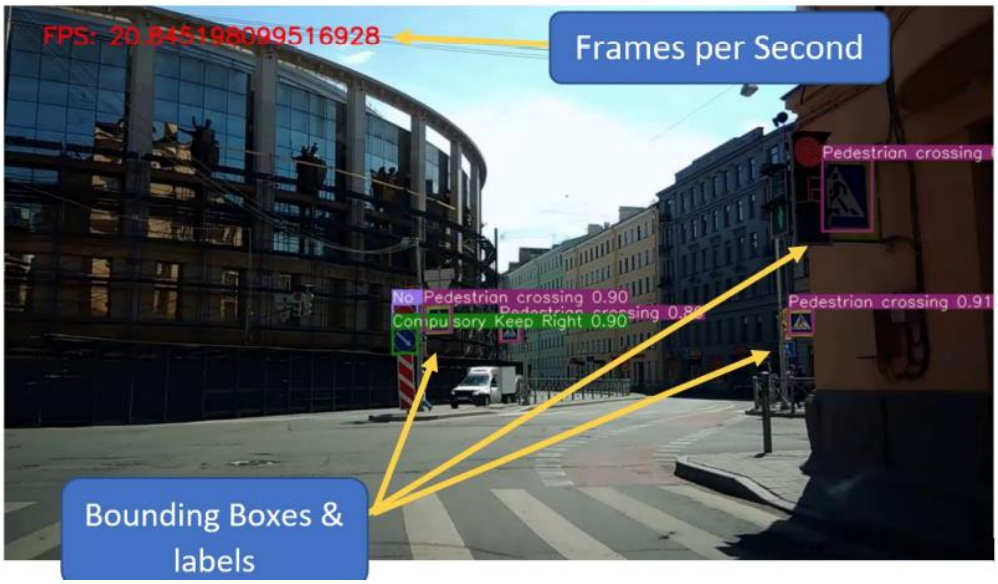
Specifications

YOLOv5 is released in four different versions namely- YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x. We have implemented our model using YOLOv5s algorithm because it is fast as well accurate enough to perform real time classification. We have collected our dataset from sources like Kaggle All the Traffic Signs are labelled by hand using the software named LabelImg. Firstly, all the images are converted into 640x640x3 pixels (3 represents the RGB channels) and after performing some normalization and pre-processing tasks these images are fed into the model for training.

Experimental Setups

The training and testing of the proposed real-time detection system challenge were implemented using the Python programming language.

Possible Output:

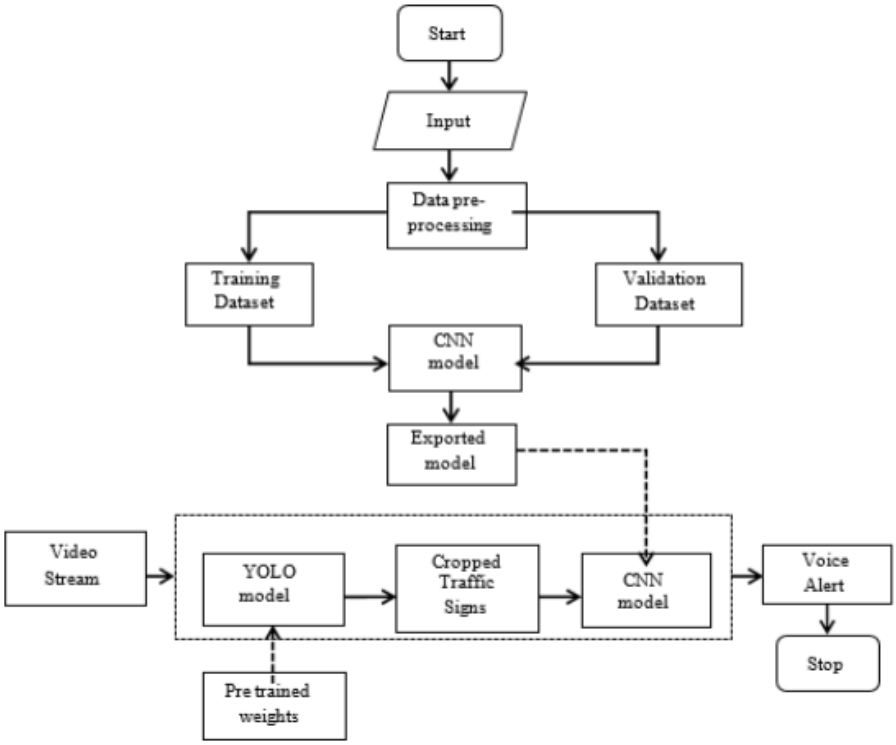


Dataset

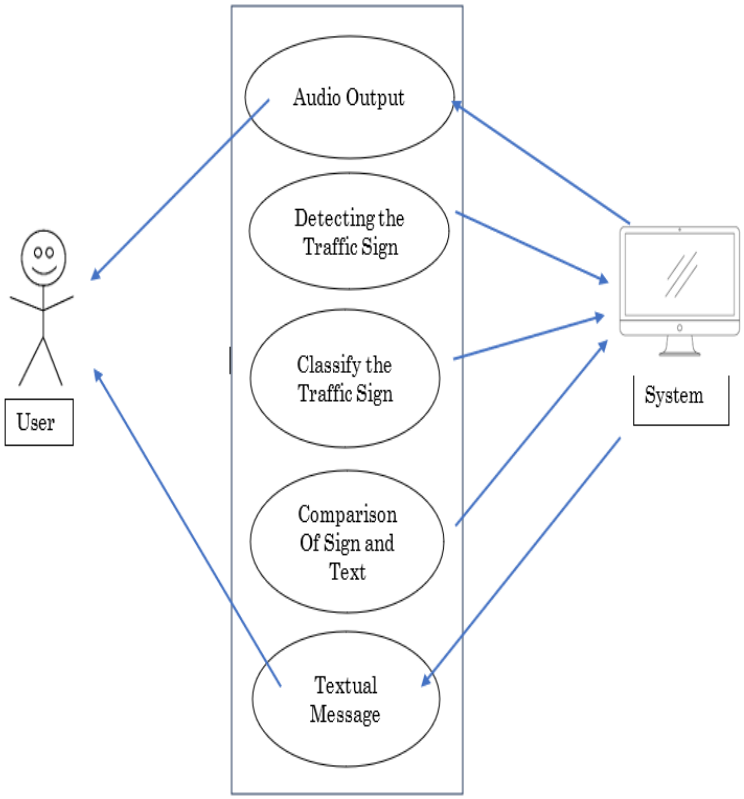
This dataset encompasses 877 images aimed at road sign detection, encompassing 61 distinct classes. Crucially, each image is meticulously annotated with bounding boxes, following the PASCAL VOC format. These annotations serve as invaluable training data for computer vision algorithms, enabling the development of robust models capable of recognizing and localizing different road signs. Such datasets are instrumental in advancing the accuracy and efficiency of traffic sign detection systems, contributing to enhanced road safety. all the images are converted into 640x640x3 pixels (3 represents the RGB channels) and after performing some normalization and pre-processing tasks these images are fed into the model for training.

4 Project diagrams

Functional block diagram - Flowchart



Use case Diagram



5 Results and Conclusion

1. Real-Time detection and classification upto 30 fps
2. Accuracy as high as 98%
3. mAP@0.5=91.75%
4. mAP@0.5:0.95=74.65%

Sr. No.	Traffic Sign Names	Laptop Camera		Mobile Camera		Wired Webcam	
		Length	Accuracy	Length	Accuracy	Length	Accuracy
1	Stop	30cm	95%	100 cm	87%	100cm	82%
2	Men At Work	30cm	93%	100 cm	86%	100 cm	80%
3	Cattle	30cm	90%	100 cm	82%	100 cm	79%
4	Pedestrian crossing	30cm	92%	100 cm	84%	100 cm	81%
5	Go Straight	30cm	91%	100 cm	80%	100 cm	78%

Sr. No.	Traffic Sign Names	Laptop Camera		Mobile Camera		Wired Webcam	
		Length	Accuracy	Length	Accuracy	Length	Accuracy
1	Stop	60cm	85%	200 cm	77%	200cm	72%
2	Men At Work	60cm	83%	200 cm	76%	200 cm	70%
3	Cattle	60cm	80%	200 cm	72%	200 cm	79%
4	Pedestrian crossing	60cm	82%	200 cm	74%	200 cm	71%
5	Go Straight	60cm	81%	200 cm	70%	200 cm	78%

Sr. No.	Traffic Sign Names	Laptop Camera		Mobile Camera		Wired Webcam	
		Length	Accuracy	Length	Accuracy	Length	Accuracy
1	Stop	90cm	75%	300 cm	67%	300cm	62%
2	Men At Work	90cm	73%	300 cm	66%	300 cm	60%
3	Cattle	90cm	70%	300 cm	62%	300 cm	69%
4	Pedestrian crossing	90cm	72%	300 cm	64%	300 cm	61%
5	Go Straight	90cm	71%	300 cm	60%	300 cm	68%

In conclusion, the real-time traffic sign detection with voice alert system represents a transformative solution poised to revolutionize road safety. Through cutting-edge computer vision algorithms, the system adeptly identifies and categorizes traffic signs in real time. The integration of voice alerts ensures instantaneous communication with the driver, fostering heightened situational awareness and reducing the potential for traffic violations. The user interface's clarity enhances user experience, presenting information seamlessly.

With adaptability for various vehicles and scalability for future advancements, the system demonstrates forward-thinking design. Compliance with traffic regulations solidifies its practicality, while user customization options prioritize individual preferences. This project stands at the intersection of innovation and safety, offering a comprehensive tool to mitigate risks and contribute to a more secure and efficient road environment. As the culmination of these efforts, the real-time traffic sign detection with voice alert system represents a significant advancement in intelligent transportation systems.

6 References

1. Morten B. Jensen, Kamal Nasrollahi and Thomas B. Moeslund, "Evaluating state-of-the-art object detector on challenging traffic light data", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 9-15, 2017.
https://openaccess.thecvf.com/content_cvpr_2017_workshops/w9/html/Jensen_Evaluating_State-Of-The-Art_Object_CVPR_2017_paper.html

2. Santos, Adonis, Patricia Angela R. Abu, Carlos Oppus and Rosula SJ Reyes, Real-Time Traffic Sign Detection and Recognition System for Assistive Driving, 2020.
https://archium.ateneo.edu/ecce-faculty-pubs/77/

3. S. Huang, H. Lin and C. Chang, "An in-car camera system for traffic sign detection and recognition", *2017 Joint 17th World Congress of International Fuzzy Systems Association and 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS)*, pp. 1-6, 2017.
https://ieeexplore.ieee.org/abstract/document/8023239

4. Z. Bi, L. Yu, H. Gao et al., "Improved VGG model-based efficient traffic sign recognition for safe driving in 5G scenarios", *Int. J. Mach. Learn. & Cyber*, 2020.
https://link.springer.com/article/10.1007/s13042-020-01185-5

5. M. Karaduman and H. Eren, "Deep learning-based traffic direction sign detection and determining driving style", *2017 International Conference on Computer Science and Engineering (UBMK)*, pp. 1046-1050, 2017.
https://ieeexplore.ieee.org/abstract/document/8093453

6. World Health Organization (WHO)", March 2022, [online] Available: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

7. GitHub link to the Ultralytics repository-
https://github.com/ultralytics/yolov5

8. Kaggle link of the dataset-
https://www.kaggle.com/datasets/ahemateja19bec1025/trafficsignlocalizationdetectionssdannotated