Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.1, May : 2024

TRAFFIC SIGN DETECTION AND VOICE ALERT SYSTEM USING DEEP LEARNING

Mrs. P.S.L. Sravani, Assistant Professor, Department of CSE, Raghu Engineering College
D. Iswarya, A. Ganesh Babu, A.N.S.S. Yaswanth, G. Manoj Kumar
Department of CSE, Raghu Engineering College

Abstract
Automation simplifies everything we do in today's world. Drivers often avoid road signs to focus on the road and this can be dangerous for themselves and others. To solve this problem, we have developed a system by using the latest version of YOLO (You Only Look Once) i.e., YOLOv8 in this project. YOLO is one of the fastest deep-learning models for real-time detection. Traffic sign detection is the task of detecting and identifying traffic signs in images or videos. Our system plans to detect nearby objects through the camera and send a voice alert to the driver so that the driver can notice traffic signs without any deviation. We have used a custom traffic sign dataset to train the model for the detection of signs. By using pyttsx3 (python text-to-speech) Library we have implemented the voice alert to the developed system.
Keywords: Object detection, YOLO v8, CNN, pyttsx3, Text to speech conversion, Voice Alert.

I. Introduction

According to statistics, there are around 400 road accidents in India every day. Road signs contribute to the safety of drivers and pedestrians by preventing accidents. Traffic signs also lessen the chance of infractions by ensuring that drivers follow the law. Road signals must take precedence over all other users of the road, including automobiles and pedestrians. People ignore traffic signs for many reasons such as emotional problems, fatigue, insomnia, Blindness, external influences, and environmental conditions. Many new machine-learning methods and algorithms have been developed to solve all these problems. Traffic sign detection was previously based on primitive search methods. Traffic sign detection pipelines typically use a handcraft to remove positive local code and then use a classifier to filter out faulty code. In recent years, deep learning has become popular, and great progress has been made in product research and knowledge projects.

Deep convolutional neural networks (CNN) are used to improve precision, speed, and accuracy in most image recognition and object detection research. CNN can learn features from big data with prior preparation. It does not need manual processing and contains a lot of general information. Recent developments in object identification algorithms, including SSD, Fast R-CNN, Faster R-CNN, R-FCN, and YOLO, have used CNN. We used "You Look Only Once" (YOLO), a single-shot detection network with minimal propagation latency and strong detection capabilities. A lot of the neural networks that are available today are accurate, but they don't work in real-time, therefore they require a lot of GPUs for training.

YOLO solves these problems. YOLOv8 ('You only look once'), an object detection algorithm that splits photos into a grid structure, is a cutting-edge piece of technology. Every grid cell oversees for items inside of itself. YOLOv8, a well-known object detection method, is renowned for its accuracy and quickness. This system will analyse photos captured in real-time by a car's front-facing camera and help the driver by alerting him or her to potential problems via auditory output or a navigation display in the car.

II. Literature

Wang C Y et.al [1] To identify and recognise traffic signs, this research investigates the use of deep learning techniques, most likely using convolutional neural networks (CNNs) or architectures similar to them.

Bochkovsky et,al [2] YOLOv4, a cutting-edge object identification model renowned for striking the ideal balance between speed and accuracy, is presented in this research. Popular object detection

framework YOLO (You Only Look Once) completes detection in a single network pass, which makes it faster than other methods.
Sheikh M A A et.al [3] This research aims to increase accuracy and resilience by proposing a system for traffic sign identification and classification that combines neural networks with colour characteristics.
G. Wang et.al [4] This research provides a coarse-to-fine approach based robust traffic sign identification algorithm. The approach achieves accurate and dependable findings by first identifying possible candidates at a coarse level and then fine-tuning the detection process. Neural networks or related methods are probably used in this approach to manage the intricacy and unpredictability of traffic sign detection tasks.
Yuan Y et.al [5] In order to improve the model's focus on pertinent areas of the input image and increase the accuracy and efficiency of traffic sign detection tasks, the network integrates vertical spatial sequence attention methods. This method probably solves issues like the different scales and orientations of traffic signs in pictures.

2.1 DEEP LEARNING MODEL.

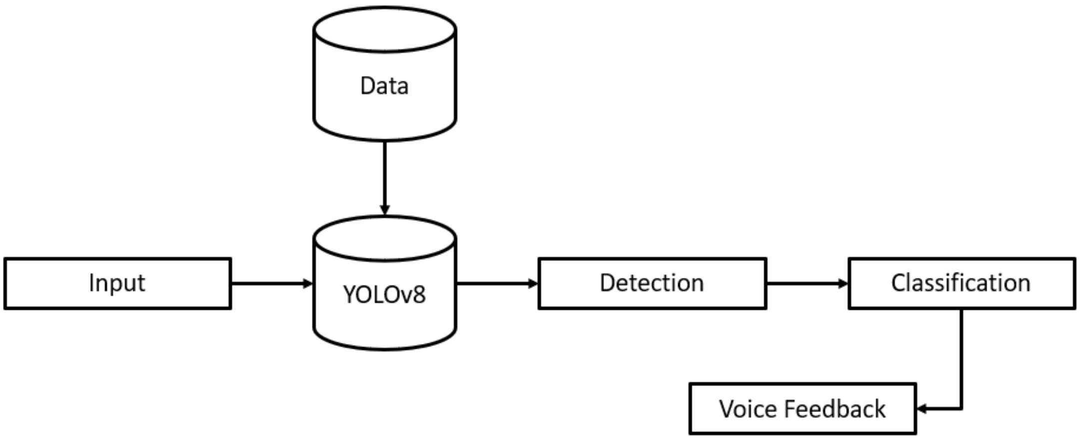

Fig 1: System Block Diagram

A YOLO-based deep learning model trained on the Traffic Sign Dataset to detect traffic signs. The aim is to build a system capable of continuously detecting and classifying traffic signs, and integrating it with a Python text-to-speech module to alert the driver with a voice when a traffic sign is detected. The key components of this system are object localization, which involves drawing bounding boxes around objects in images, and classification, where objects are assigned class names.
Data set: A custom traffic sign data set consisting of 9948 images with 8 classes is used for implementing this project. The dataset is divided into three parts, one is for training, one is for testing and the other is for validating. The collection includes natural traffic sign images that were taken on several types of roadways (rural, metropolitan, and highway) at different times of day, at sunset, and with varying weather patterns. These variations in lighting and weather cause traffic signs to get obstructed or have varied orientations.
PYTTSX3: PYTTSX3 is a text-to-speech conversion library which is developed using Python programming language. This library is supported by many types of operating systems and it can be operated without an internet connection. It can be used with the various versions of the Python programming language. To install this library, we have to use a pip package manager. This Python library offers two voices: male and female.
YOLOv8 Algorithm: YOLO algorithms are deep learning models that are widely used for object detection and classification because of their speed and accuracy. The YOLOv8 is the latest version of

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.1, May : 2024

YOLO, which differs in speed and accuracy from other versions. The YOLO is still developing algorithm, it has to go through several iterations to improve and achieve a greater performance.
The YOLOv8 model consists of three main components, the Backbone, the Neck and the Head. Each component performs separate operations that are responsible for the final detection of the object. The Backbone and the Head are connected by the Neck and helps in the process of feature extraction for the detection of the objects.
Backbone: Yolov8 is built on a feature-rich backbone network. By extracting features from the input image, the network offers a thorough representation of the visual data. YOLOv8's Backbone is a modified version of the Darknet architecture called CSPDarknet53. By incorporating Cross Stage Partial networks, this change improves the effectiveness and capacity of learning.
Neck: The neck structure in the architecture allows features to blend. This is essential for merging multi-scale data and enhancing the model's capacity to identify objects of different sizes. PANet (Path Aggregation Network), a feature pyramid network that enables information flow across various scales, is introduced by YOLOv8. The model can handle objects with different scales more effectively.
Head: Based on the features that the neck architecture and backbone network have retrieved, this component makes predictions. For every anchor box connected to a grid cell, the YOLO head forecasts bounding box coordinates, object scores, and class probabilities. Anchor boxes are used by the architecture to forecast items of various sizes and shapes with efficiency.

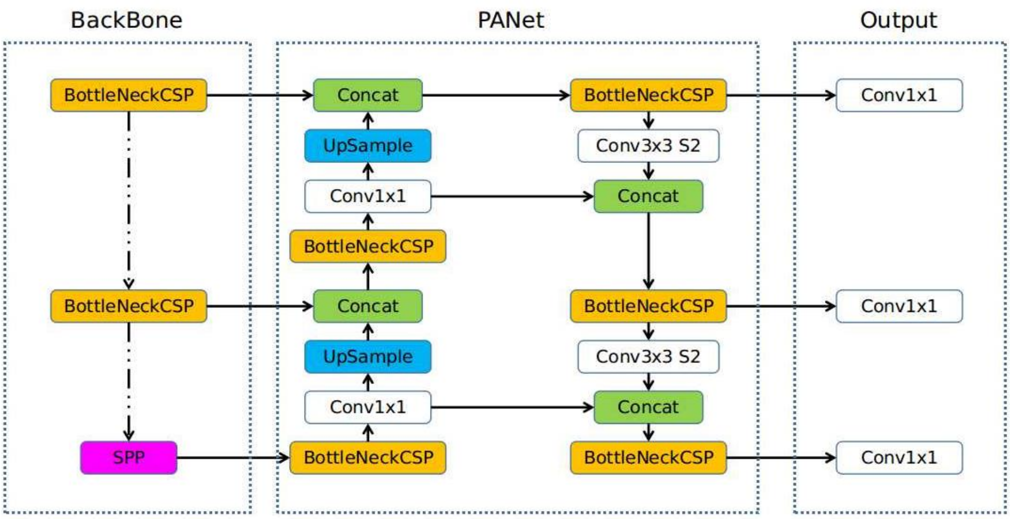

Fig 2: YOLO Architecture

2.2 IMPLEMENTATION

Training the model: A custom dataset was used to train the model in Google Colaboratory. The Google Colaboratory is a web-based tool which is used to run and execute codes. It offers free access to GPUs and computational resources and doesn't require any installation. To access and train the YOLO model in the Google Colaboratory using a custom dataset, we need to execute a code. A "best.pt" file is produced once the model has been trained. Later on, this file is inserted into the Python code which is executed in Visual Studio code.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.1, May : 2024

2.3 IoT and AI for crop selection

Ishita Dasgupta et al [43] combined the use of IOT devices, wireless sensor network (WSN), and AI techniques in order to provide farmers with a faster and more accurate recommendation of suitable crops. These recommendations are based on a number of factors, including temperature, annual precipitation, total available land size, previous crop grown history, and other resources. In addition, the identification of undesired plants on crops, also known as weed detection, is carried o

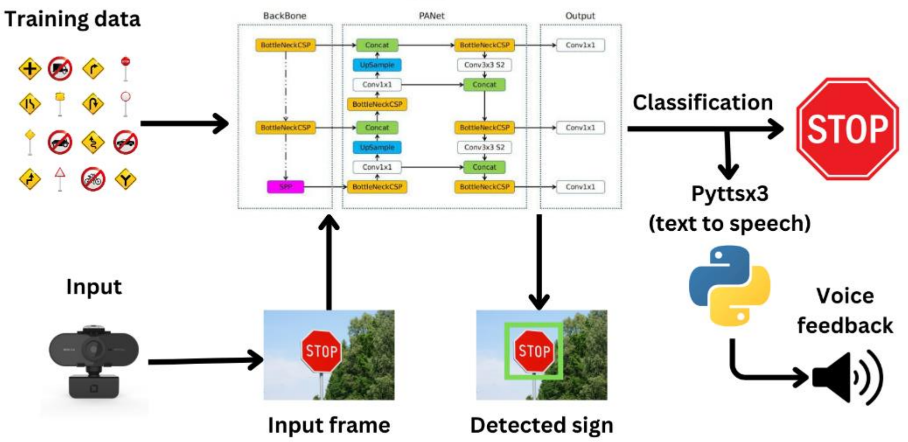

Fig 3: System Design

Input data: After successful training and execution of the code in Visual Studio code, We will use a camera as an input device to feed video frames to the trained model. We can also use an image or a video as input for the trained model.

Detection and Classification: The Trained Traffic sign detection model then processes the input frames to detect and classify the object creates bounding boxes around the object and labels them with the class name.

Voice Alert: When a traffic sign is detected, it is labelled with the class name and this class name is stored as a string. By using PYTTSX3, these strings are converted to speech and the system alerts the driver through a speaker.

RESULTS AND ANALYSIS:

An image or a video can be given as an input. We have used a image as input and the results are shown in the below figure.

Fig 4: Detection of sign in an image

In real time, A Camera that is fixed to the vehicle acts as an input device to feed video frames to the trained model. We have used web camera as an input device to feed video frames to the model. The results are shown in the below figures.

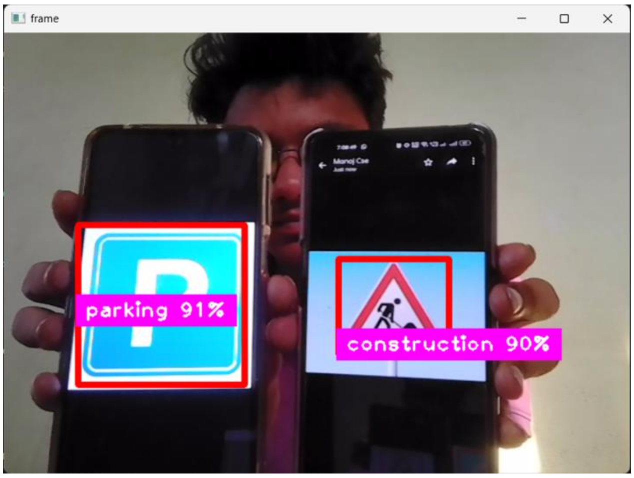

Fig 5: Detecting multiple signs using webcam

The trained model is capable of detecting multiple signs with the accuracy more than 80%.We have achieved the highest accuracy of 97% in detecting the traffic sign by using this model. The detection of objects through camera is also depends on the quality of the camera. The highest compatible camera can be helpful to detect traffic sign quickly from more distance.

2.4 CONCLUSION

In this article, a traffic sign detection method using a deep learning model is presented. We explored the detection of traffic signs using YOLOv8 with the goal of automation traffic sign detection and alerting the driver through voice alert. We suggested applying the YOLOv8 algorithm, which is a method used for traffic sign detection. This method yields faster identification results than other object detection techniques like Fast R-CNN, Faster R-CNN, and other R-CNN algorithms. The system detects a large number of categories rapidly and effectively by using a deep network to learn them. A camera will be used to take the video frames as input during the image detection. The system sounds an alert when it detects a traffic sign at a point. When precise and safe navigation is required, this approach works well.

III. Conclusion

In order to solve the problems of diminishing arable land and the rising demand for food brought on by an expanding global population, improved and more effective methods of crop production are required. Everyone should make it a priority to educate themselves on the importance of food security in relation to environmentally responsible agriculture. The proliferation of new technology that may boost agricultural yields and encourage inventive young people to take up farming as a respectable vocation are two positive outcomes of this trend. This article stressed the role that many of the technologies now employed in farming, notably IoT and AI, play in making agriculture smarter and more successful so that it can meet the demands of the future. Scholars and engineers might benefit from taking notice of the present issues confronted by the sector as well as the future potential. Because of this, every acre of farmland should be used to its full potential in order to maximize agricultural output. This may be accomplished by using environmentally friendly sensors and communication systems that are powered by artificial intelligence and the internet of things.

References

[1] Wang C Y, "Exploring and utilising deep learning for traffic sign identification and detection", International Conference on Robots & Intelligent Systems (ICRIS),2018.

[2] Bochkovsky, Alexey, H. Liao, and Chien-Yao Wang. "YOLOv4: Optimal Velocity and Precision for Object Identification." n. pag., ArXiv abs/2004.10934 (2020).
[3] Sheikh M A A, Kole A., Maity T, "Traffic sign detection and classification using the colour feature and neural network," International Conference on Intelligent Control Power and Instrumentation, 2016.
[4] G. Wang, G. Ren, Z. Wu, Y. Zhao, and L. Jiang, "A robust, coarse-to-fine traffic sign detection method ", Proceedings of the 2013 International Joint Conference on Neural Networks, Dallas, TX, USA, 2013, pp. 754–758.
[5] "VSSA-NET: Vertical Spatial Sequence Attention Network for Traffic Sign Detection," by Yuan Y, Xiong Z, and Wang Q, IEEE Transactions on Image Processing, 2019, (99):1-1.
[6] "Linear and circular extraction method based on Hough transformation", by Chen W. W., Wu W. Electron mass, 383(02):25–27 (2019).
[7] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi."You Only Look Once: Unified, Real-Time Object Detection" arXiv:1506.02640v5.
[8] Yuan Y, Xiong Z, Wang Q, "VSSA-NET: Vertical Spatial Sequence Attention Network for Traffic Sign Detection," IEEE Transactions on Image Processing, 2019, (99):1-1.
[9] Chen W W, Wu W, "Linear and circular extraction method based on Hough transformation. Electron mass, "2019, 383(02):25-27.
[10] Zhu Z, Liang D, Zhang S, et al, "Traffic-Sign Detection and Classification in the Wild," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[11] Sheikh M A A, Mukhopadhyay S, "Noise tolerant classification of aerial images into manmade structures and natural-scene images based on statistical dispersion measures," IEEE India Conference. 2012, 653– 658.
[12] Gomez-Moreno H , Maldonado-Bascon S , Gil-Jimenez P, et al, "Goal Evaluation of Segmentation Algorithms for Traffic Sign Recognition," IEEE Transactions on Intelligent Transportation Systems, 2010, 11(4):917-930.
[13] Zhu Y, Zhang C, Zhou D, et al, "Traffic sign detection and recognition using fully convolutional network guided proposals," Neurocomputing, 2016:S092523121630741X.
[14] Zhao G F, " Traffic sign recognition Based on Machine Learning " Hangzhou University of Electronic Science and Technology, 2015.
[15] G. Wang, G. Ren, and T. Quan, "A traffic sign detection method with high accuracy and efficiency," in Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE), Hangzhou, China, 2013, pp. 22–23.
[16] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA,,2014, pp. 580–587.
[17] Redmon, J., Farhadi, "A.: YOLO9000: better, faster, stronger". In: IEEE CVPR, pp. 7263–7271 (2017).
[18] Adrian Rosebrock, Intersection over Union (IoU) for object detection on November 7, 2016.https://pyimagesearch.com.
[19] Xu, R.; Lin, H.; Lu, K.; Cao, L.; Liu, Y. A Forest Fire Detection System Based on Ensemble Learning. Forests 2021, 12, 217.
[20] Chen T, Lu S, "Accurate and Efficient Traffic Sign Detection Using Discriminative Adaboost and Support Vector Regression," IEEE Transactions on Vehicular Technology, 2016, 65(6):4006-4015.