IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14101

Optimizing YOLOv10 for Real-Time Traffic Sign Detection and Recognition: A Bangladeshi Perspective

Mrittika Mahbub[1], Md. Habib Ehsanul Hoque[2]

Lecturer, Dept. of CSE, Pundra University of Science & Technology, Bangladesh[1]

Assistant Professor, Dept. of CSE, Pundra University of Science &Technology, Bangladesh[2]

Abstract: Traffic sign detection and identification not only support driver assistance technologies but also play a vital role in enhancing traffic management. These capabilities are essential for ensuring safe transportation and effective operation of self driving vehicles, aiding in real-time decision making for both human and autonomous drivers. Recognizing traffic signs in Bangladesh is especially difficult due to the country's unique driving environment. This includes non-standard signage, diverse road conditions, and the frequent presence of pedestrians and livestock on roadways. By utilizing technologies like machine learning and computer vision, these systems can be tailored to local contexts, ultimately enhancing roadway safety in Bangladesh. This research proposes a smart assistant system that utilizes a dataset of 6,000 diverse traffic sign images from Bangladeshi road environments to improve road safety, especially in regions with potential driver compliance issues. The dataset encompasses 41 traffic sign categories and presents various real-world challenges, including faded color, weather conditions, blurring and vibration, occlusion, variable size of traffic sign and low light conditions. To enhance the dataset's robustness, data augmentation techniques such as random rotations, shearing and zooming were applied. We trained the YOLOv10 deep learning model, renowned for its real-time object detection capabilities, on this dataset. The model achieved significant results, with a mean Average Precision (mAP) of 0.80, a recall of 0.87, a precision of 0.92, and an F1 score of 0.89, demonstrating its effectiveness in real-world traffic sign detection and classification.

Keywords: TSR, Smart Traffic, Feature Extraction,Traffic Sign,YOLOv10,Computer Vision

I. INTRODUCTION

Bangladesh is grappling with a severe road safety crisis, reflected in high rates of accidents and fatalities. According to a recent study published in national newspapers, core factors contributing to this crisis include unskilled drivers, disregard for traffic signs, poor road conditions, and risky driving behaviors. In 2023 alone, there were 7,902 fatalities from 6,261 collisions, with 10,372 injuries. This represents 32.43% of all incidences, frequently as a result of young, in-experienced riders driving recklessly [1].The rapid increase in vehicles on the road has made traffic management increasingly challenging, emphasizing the need for accurate traffic sign detection and identification to ensure effective and safe control. For the adoption of smart driving and accident prevention, real time detection and classification of traffic signs are essential to accurately identify and categorize sign types. This process is critical in computer vision applications, contributing significantly to traffic monitoring, autonomous driving, safety, and environmental management. The government is implementing an extensive upgrade of the country's highways, aiming to expand all major highways to four to six-lane standards in phases. Key routes such as the Dhaka-Sylhet, Dhaka-Rangpur, and Dhaka-Chittagong Highways are part of this project, with plans to convert approximately 4,000 kilometers of road to wider lanes, beginning with a 13.5-kilometer expressway extension into a six-lane route [2]. This initiative addresses the rising demand for traffic capacity and aims to improve road safety by ensuring strict adherence to traffic signs and regulations. However, many local drivers and their assistants lack awareness of traffic laws or face significant literacy barriers, making compliance difficult. This knowledge gap leads to frequent and severe accidents, not only on highways but also on rural and urban roads, underscoring a critical public safety issue with a high toll of traffic-related deaths and injuries. Additionally, limited technological resources hinder effective law enforcement in identifying and penalizing violators. Our research aims to address these issues, offering a valuable solution for both the government and society. Looking ahead, adopting smart assistant technology to detect and identify traffic signs along roadsides can significantly enhance driver awareness and improve travel safety. Our main objective is to achieve accurate real-time detection and classification of traffic signs. While various methods have shown promising results [3], real-time detection remains challenging due to complex factors like crowded roads, environmental conditions, and issues such as blurriness, occlusions, small sign sizes, low lighting, color blending, and poster obstructions. To address these, we assembled a comprehensive dataset of diverse and challenging real-time traffic sign images and videos. Using the

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025
DOI: 10.17148/IJARCCE.2025.14101

advanced YOLOv10 deep learning model [4], we evaluate detection performance, optimizing it for flexible deployment across platforms.

II. LITERATURE REVIEW

Computer vision has produced several methods for recognizing and detecting traffic signals, but many of them have drawbacks that restrict their applicability for intelligent traffic automation and real-time detection.The YOLO(You Only Look Once) series (from v1 to v10) [5] created in recent years stands out among the several real-time object identification systems. R-CNN [6] was one of the first deep learning methods to achieve significant success in object detection, using a two-stage approach. Recently, Zhang, Guirong, et al. [7] introduced the RTS R-CNN Instance Segmentation Network, a refined model designed to enhance real-time traffic sign detection accuracy. Commonly applied in image processing, this model uses CNN primarily for feature extraction, focusing on efficient and precise traffic sign identification under various conditions. Various techniques have been developed in computer vision for detecting and identifying traffic signs, but many face challenges that limit their suitability for real-time detection and intelligent traffic automation. In this section, we review several key studies on traffic sign detection. To increase the accuracy of traffic sign identification, Y Sun et al. [8] presented a CNN and twin support vector machine (SVM) hybrid model that combines deep learning and machine learning approaches. While the twin SVM classifier improves classification accuracy, the CNN is used for feature extraction, capturing key sign properties. Using deep learning technology, specifically CNNs, Lim, Xin Roy, et al. [9] were able to match traffic signs in real time with greater accuracy. In order to increase the dependability of traffic sign recognition (TSR) systems, their work highlighted the significance of large-scale datasets. Fast R-CNN and Faster R-CNN [10] further advanced object detection by introducing more efficient feature extraction and transitioning toward an end-to-end approach. Fast R-CNN used SPPNet [11] to accelerate feature extraction, while Faster R-CNN introduced Region Proposal Networks, enabling a more streamlined process for identifying objects within images. In 2015, Joseph Radmon proposed YOLO [12]. It completes object detection in a single phase by using per gird prediction. This innovative method raises the bar for real-time object detection to a completely new level. The efficiency of the YOLO method was emphasized by Flores-Calero et al. [13], specifically its high accuracy and real-time competence in detecting difficult and occluded traffic signs. Real-time object identification has benefited greatly from one-stage detectors like RetinaNet [14] and [15], and scaled-YOLOv4 [16] has further improved model efficiency through sophisticated scaling techniques. While X. Chen et al. presented an enhanced YOLOv5 [17] for traffic signal recognition, optimizing quality and efficiency through fine-tuned attribute combinations, T. H. Tran et al. [18] assessed YOLOv4's performance against previous YOLO versions and other models. The difficulty of identifying tiny traffic signs in complicated weather situations is addressed by Qu, Shenming, et al.'s enhanced YOLOv5 model [19]. By improving feature extraction and increasing detection accuracy under challenging circumstances, this improved model adapts YOLOv5 to more accurately detect tiny, far-off indications. However, the YOLOv7-TS [20] model enhances traffic sign detection using sub-pixel convolution to identify better small signs and feature fusion to integrate multi-scale information. It demonstrates superior performance over the standard YOLOv7, making it effective for accurate detection in complex traffic scenarios. A YOLOv8-CE-based approach for autonomous vehicle-specific real-time traffic sign recognition and identification is presented by Luo, Yuechen, et al. [21]. The method improves the operational safety of autonomous driving systems by increasing the speed and accuracy of recognizing different traffic signs by utilizing the sophisticated capabilities of the YOLOv8 architecture. In their study of road signs in Bangladesh, Ahsan, Sk Md Masudul, et al. [22] emphasized the development of a national dataset, enhancing road safety, adjusting to local road circumstances, and addressing both global and regional significance. The newest model in the YOLO family of object detection models, YOLOv10 [4], was created to enhance real-time detection capabilities while resolving some of the issues with earlier iterations. To achieve good accuracy, we developed a dataset specific to Bangladeshi real-time traffic signs and employed the YOLOv10 model for detection and classification, ensuring that the system can effectively recognize and interpret various traffic signs under diverse conditions, ultimately contributing to improved road safety and traffic management in the region.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14101

III. PROPOSED METHODOLOGY

We divided our workflow into several key steps to implement the proposed strategy, as shown in Fig. 1. The following subsections provide a detailed explanation of each step.

A. Data Collection

For this work, data was gathered by capturing images of road signs from diverse locations, including highways, urban, rural, and overpass areas, using both stationary and moving viewpoints. To ensure variety, images were taken from multiple angles at different times of day and across varied environments. In our dataset, certain traffic signs presented unique challenges, including signs that were small, blurry, obstructed, or poorly lit, and those on posters. Clear, well-lit images were more identifiable, while fog, low light, or obstructions from trees made detection harder. We gathered 41 unique types of traffic signs, along with real-time roadside data, which presented challenges like traffic barriers and various objects obstructing clear traffic sign detection. To address this, we gathered data both while stationary and while in motion, using multiple vehicles to accurately measure distances and assess the appropriate scenarios for road signs and the detection devices employed which are shown in Fig. 2. This approach ensured a comprehensive understanding of the environmental conditions affecting sign visibility and detection accuracy.

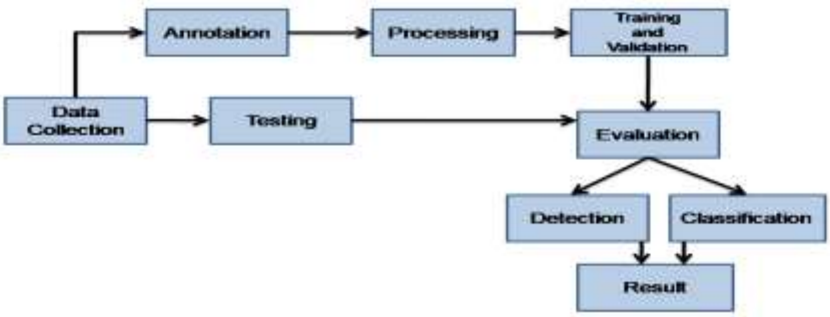

Fig. 1. An overview of the planned methodology.

Fig. 2. Several Challenges Faced in Detection.

B. YOLO Algorithm Concept

YOLO (You Only Look Once) is one of the most widely used techniques in computer vision, which uses a variety of models to train datasets in object detection. The YOLOv1 model, first presented by Joseph Redmon in Darknet, was the first object detection network to integrate class identification with bounding box prediction in a single pass. From YOLOv2 to YOLOv8 [5], iterations have significantly increased speed and accuracy, establishing YOLO as a top computer vision framework for object detection. The most recent iteration of the YOLO series, YOLOv10, was created to improve real-time object detecting capabilities. This model adds new features that increase accuracy and efficiency while building on the advantages of its predecessors. This is a thorough rundown of YOLOv10 [4] that emphasize its applications, performance, and architecture. The architecture of YOLOv10, shown in Fig. 4, is divided into three main sections:

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14101

• **Backbone:** An enhanced version of CSPNet (Cross Stage Partial Network), which improves gradient flow and reduces computational redundancy to produce more efficient feature extraction from input photos, forms the foundation of YOLOv10.

• **Neck:** This component aggregates features across various scales using Path Aggregation Network (PAN) layers, allowing for efficient multiscale feature fusion. This ensures that features from different levels of the backbone are combined effectively to boost detection performance.

• **Head:** YOLOv10 presents a dual-head design:
  – One-to-Many Head: During training, this head produces several predictions for every item, offering helpful supervisory cues that raise learning precision.
  – One-to-One Head: This head reduces processing time considerably and does away with the need for Non-Maximum Suppression (NMS) by generating a single optimal prediction for each item during inference.
  YOLOv10 introduces several unique architectural innovations that enhance its performance and efficiency:

- **Dual Label Assignment**: YOLOv10 uses a label assignment approach similar to DATE, incorporating a stop-
- gradient operation in the one-to-one branch to improve training efficiency.
- **NMS-free Object Detection:** By leveraging a one-to-one matching mechanism, YOLOv10 eliminates the need for Non-Maximum Suppression (NMS) in the post-processing phase, thereby streamlining the prediction process.
- **Rank-guided Block Design:** YOLOv10 introduces a rank-based approach that decides where to apply conventional convolutions versus depth-wise convolutions, optimizing the model's architecture for better performance.
- **Partial Self-Attention:** YOLOv10 combines CSP-Net with Transformer-inspired elements to introduce a self-attention module, enhancing feature extraction and overall model performance.

These architectural features collectively improve YOLOv10's accuracy and efficiency for real-time applications.

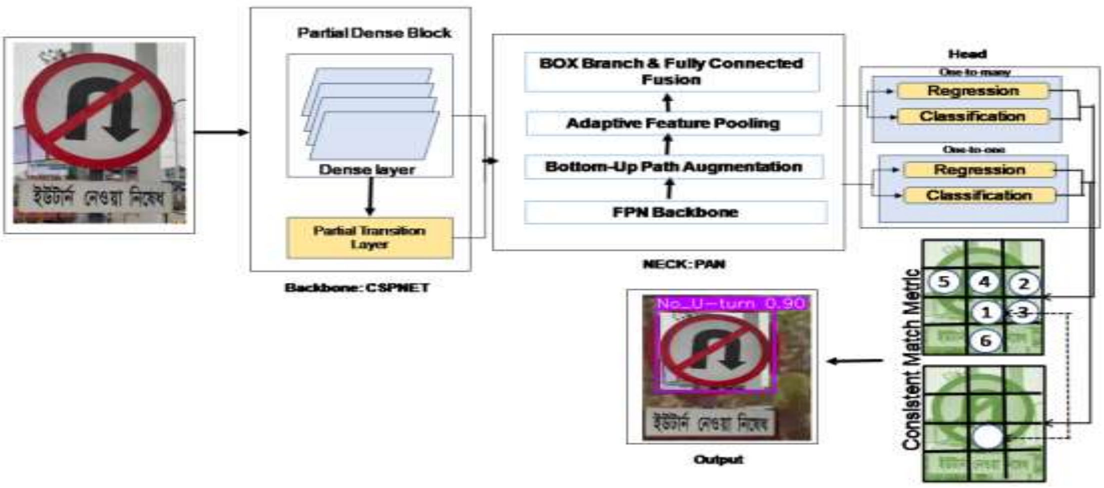

Fig. 3. The YOLOv10 Architecture Used for Traffic Sign Detection.

C. Traffic Sign Detection Using YOLOv10

In order to handle the various features of traffic signs which can differ greatly in size, shape, colour, and appearance due to elements like shifting lighting conditions, occlusions, and different distances from the camera the YOLOv10 architecture makes use of sophisticated deep learning techniques. YOLOv10 arranges incoming photos into a grid that predicts bounding boxes, confidence scores, and class probabilities for particular regions and its internal mechanism in order to perform efficient object recognition.

Backbone Feature Extraction: Convolutional layers are used at the core of YOLOv10 to extract features. The output of a convolutional layer, referred to as Backbone Feature
Extraction can be expressed as follows. :

$$O_{i,j}(k)=\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} I_{m,n}^{(c)}\cdot K^{(k)}_{i-m,j-n} \qquad (1)$$

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025
DOI: 10.17148/IJARCCE.2025.14101

Where:

$O^{(k)}_{i,j}$ is the output at position (i, j) for the k-th output.
$I^{(c)}_{m,n}$ is the input position (m, n) in the c-th channel.
$K^{(k)}_{i-m,j-n}$ is the kernel used produce the output.

Path Aggregation Network (PAN): The feature aggregation process can be described as

$$F_{out} = F_{low} + \text{Upsample}(F_{high}) \tag{2}$$

Where:

- F_{out} is the output feature map after aggregation.
- F_{low} is a feature map from a lower layer in the backbone.
- F_{high} is a feature map from a higher layer.
- Upsample refers to an operation that increases the spatial dimensions of the feature map, often achieved via bilinear interpolation or nearest neighbor methods.

Loss Function:

$$\text{Loss} = \lambda_{loc} \cdot L_{loc} + \lambda_{conf} \cdot L_{conf} + \lambda_{cls} \tag{3}$$

Where:

- L_{loc} is the localization loss (typically MSE for bounding box coordinates).
- L_{conf} is the confidence loss for the predicted bounding boxes.
- L_{cls} is the classification loss.
- λ_{loc}, λ_{conf}, λ_{cls} are hyperparameters that control the importance of each component.

Bounding Box Prediction: YOLOv10 predicts bounding boxes for detected traffic signs by calculating the center coordinates (b_x, b_y), width (b_w), and height (b_h) of each bounding box.

$$b_x = \sigma(t_x) + c_x, \tag{4}$$

$$b_y = \sigma(t_y) + c_y, \tag{5}$$

$$b_w = p_w \cdot e^{t_w}, \tag{6}$$

$$b_h = p_h \cdot e^{t_h} \tag{7}$$

Object Confidence Score: The object confidence score indicates the likelihood that a bounding box contains an object (traffic sign). It is defined as:

$$\text{Confidence} = P(\text{object}) \times IOU_{pred,\ truth} \tag{8}$$

NMS-Free Mechanism (YOLOv10 Innovation): In YOLOv10, the dual-head architecture enables NMS-free processing during inference, allowing a One-to-One Matching Mechanism:

$$\text{Prediction} = \text{argmax}(\text{One-to-One Head}) \tag{9}$$

This mechanism allows YOLOv10 to produce a single prediction per object, reducing latency and improving efficiency

IV. RESULT AND DISCUSSION

To conduct our studies, we used a virtual machine on Kaggle that has a P100 GPU. The system's performance was also assessed using a device that had an Intel Core i5 9th Gen CPU and 8 GB of RAM. To take pictures, a 1080p camera with a live feed of 30 frames per second was used. We built up the datasets, ran tests with the YOLOv10 model, and evaluated the correctness of the model in our experimental evaluation.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14101

A. Dataset Preparation

The total number of photos in our collection was 6,000. About 30% of the images were taken from the video clips, and the remaining 70% were taken as still photos,4,800 training and validation photos, or 80% of the total dataset, were obtained by using 118 images for each class as training and validation data, Figure 3 shows the sample of our dataset's several kinds of Bangladeshi Real-time traffic signs. For testing, the remaining 1,200 photos (20% of the entire dataset) were employed. The second column of Table I displays the quantity of testing photos for each class. We utilized Make Sense AI, a free, open-source, and user-friendly web program, to annotate the dataset.

Fig. 4. A sample of our dataset's several kinds of Bangladeshi Real-time traffic signs.

B. YOLOv10 Model Training

With pre-trained weights from YOLOv10, the training model is YOLOv10s. Every image has a 640 x 640 pixel resolution. To obtain the best fit and avoid both overfitting and underfitting, the model was trained over 100 epochs.

C. Model Accuracy Measurement

The following subsections give an overview of the most commonly used metrics to assess model accuracy, where TP, FP, and FN represent True Positives, False Positives, and False Negatives, respectively.

1) Precision and Recall: Precision reflects the ratio of correctly identified positives among all detected instances, indicating the model's accuracy in recognizing traffic signals:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall measures the proportion of true positives among actual instances in the dataset, representing the model's ability to capture all real signs:

$$\text{Recall} = \frac{TP}{TP + FN}$$

2) F1 Score: The F1 Score, a balance between precision and recall, is the harmonic mean of these metrics:

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{Precision + Recall}$$

3) Mean Average Precision (mAP): The mean of all Aver-age Precision (AP) scores across classes, mAP gives an overall measure of model performance across the dataset:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} \text{APi}$$

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14101

Fig. 5. Our test dataset's sample detection outcomes across a range of test settings.

TABLE I
ASSESSMENT METHODS FOR EVERY CLASS

Information				mAP	
S/N	**Class**	**P**	**R**	**mAP @0.5**	**mAP @(50-0.95)**
1	Double Bend (Left)	0.883	0.602	0.757	0.457
2	Double Bend (Right)	0.883	0.582	0.782	0.479
3	Bazar Ahead	0.959	0.952	0.992	0.416
4	Crossroad Ahead	0.967	0.432	0.681	0.291
5	Keep Left	1.000	0.815	0.911	0.526
6	Keep Right	1.000	0.833	0.917	0.673
7	Left Side Road	0.623	0.310	0.434	0.165
8	Right Side Road	0.537	0.560	0.532	0.345
9	No Parking	0.966	1.000	0.995	0.556
10	No U-turn	0.869	0.903	0.907	0.573
11	Sharp Bend (Left)	0.800	0.154	0.493	0.339
12	Sharp Bend (Right)	0.569	0.767	0.603	0.259
13	School	0.832	0.939	0.927	0.480
14	U-turn	1.000	0.339	0.673	0.505
15	Bus Stop	1.000	0.611	0.812	0.618
16	Hospital	0.955	0.952	0.975	0.667
17	Mosque	0.947	0.819	0.926	0.648
18	Toilet	1.000	1.000	0.995	0.736
19	Road Hump	0.895	0.636	0.773	0.522
20	No Overtaking	0.863	0.197	0.531	0.362
21	No Use of Horn	0.824	0.812	0.868	0.623
22	Narrow Bridge	0.926	0.802	0.889	0.473
23	No Rickshaws	0.833	0.950	0.964	0.774
24	Pedestrian Crossing	0.958	0.372	0.673	0.424
25	Railway Crossing with	0.788	0.788	0.735	0.342
26	Church	1.000	1.000	0.995	0.686
27	Give Way	1.000	0.826	0.913	0.766
28	Left Side Road Narrow	0.996	1.000	0.995	0.682
29	One Way	1.000	1.000	0.995	0.664
30	Pass Either Side	1.000	1.000	0.995	0.687
31	Speed Limit 5	1.000	1.000	0.995	0.633
32	Speed Limit 15	1.000	1.000	0.995	0.793
33	Speed Limit 20	0.972	0.644	0.823	0.712
34	Speed Limit 30	0.952	0.909	0.952	0.825
35	Speed Limit 40	0.849	0.164	0.478	0.380

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14101

36	Speed Limit 50	0.952	1.000	0.990	0.806
37	Speed Limit 60	0.921	0.759	0.868	0.626
38	Speed Limit 80	0.962	0.870	0.931	0.745
39	No Stopping	0.900	0.903	0.965	0.818
40	Sharp Bend Left	1.000	0.968	0.995	0.735
41	No Trucks	1.000	0.975	0.989	0.767
Total	**1235**	**0.912**	**0.766**	**0.844**	**0.575**

4) The Detection Performance: An overview of our method's performance on test data is given in Table I. The category of test samples is shown in the second column, while the third and fourth columns provide specifics on Precision(P) and Recall (R). The mean Average Precision (mAP) at 50% and 50%-90% confidence levels are displayed in the fifth and sixth columns, respectively. Across 41 categories, the model's average recall was 0.76, its precision was 0.912,and its F1 score was 0.80. The YOLOv10 model shows strong performance across various metrics, as evidenced by the detection samples. Figure 6 illustrates the Precision-Recall curves, Figure 7 presents the F1-Confidence curves, and Figure 8 displays the confusion matrix.

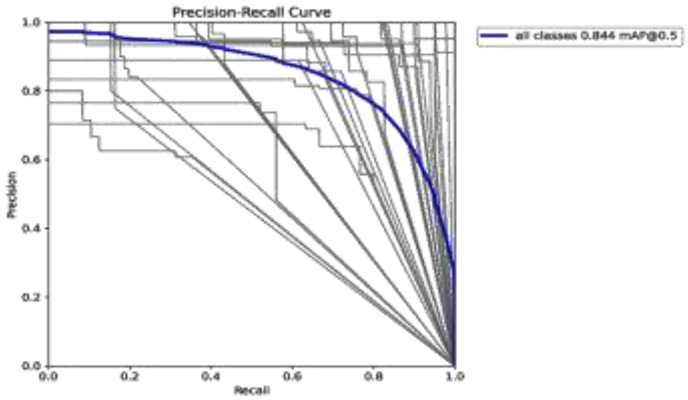

Fig. 6. Over the performance of YOLOv10, the precision-recall curve

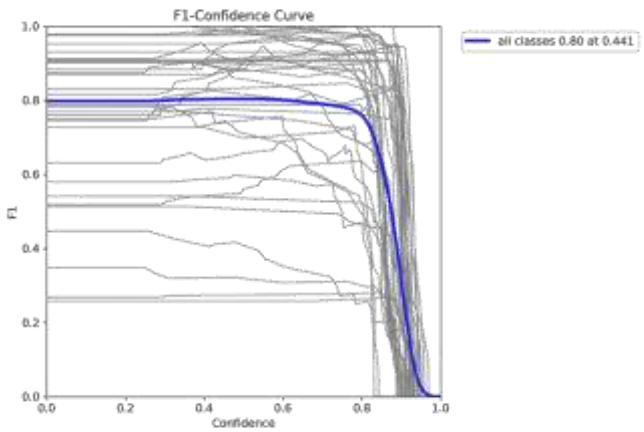

Fig. 7. F1-score curve

The model demonstrates high precision across many categories, particularly for signs such as "Keep Left," "Keep Right," "No Parking," "One Way," and "Toilet," each achieving 100% precision. However, recall varies, with certain categories like "Sharp Bend (Left)" and "Crossroad Ahead" showing lower recall values, indicating potential under detection. Figure 5 shows the sample detection outcomes across a range of test settings. Overall, the model shows balanced performance across precision and recall, with mean Average Precision (mAP) metrics at different confidence thresholds underscoring its effectiveness. Signs with high mAP@0.5 and mAP@0.5-0.95 scores confirm the robustness of detection, though specific classes could benefit from further training to improve recall.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⋈ Peer-reviewed & Refereed journal ⋈ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14101

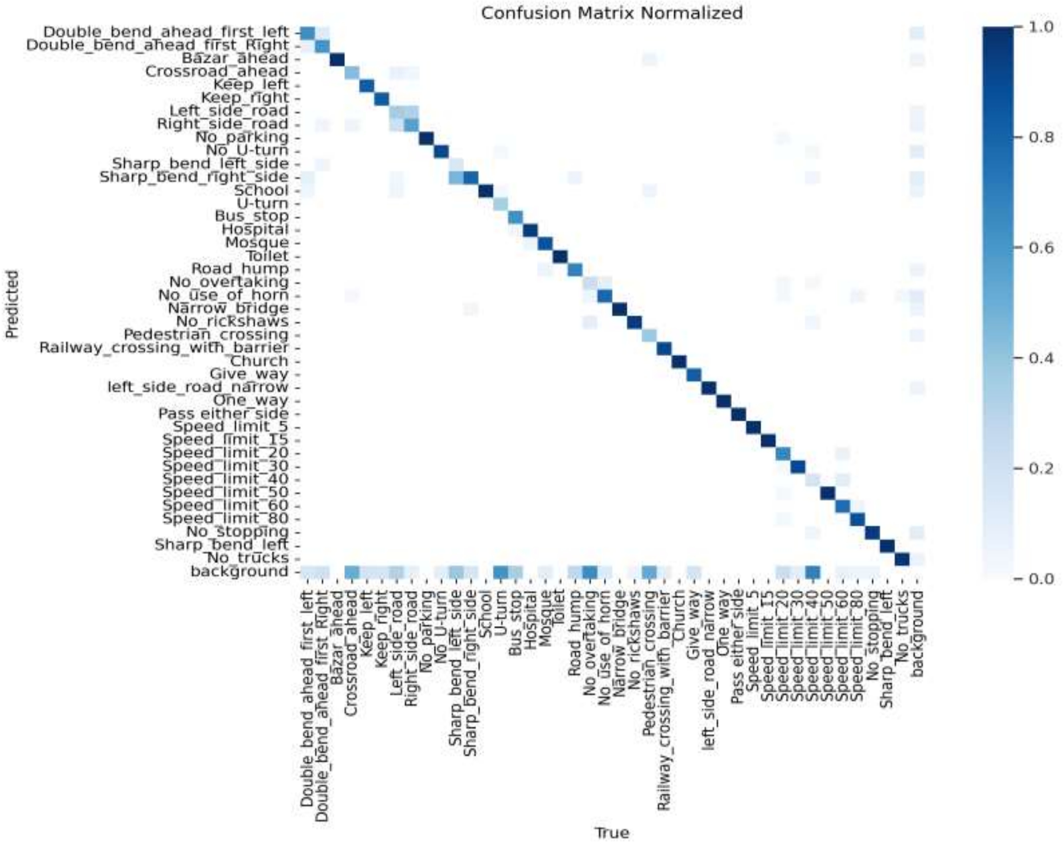

Fig. 8. Confusion Matri

V. CONCLUSION

This paper explores the development and evaluation of an advanced traffic sign recognition (TSR) system utilizing advanced machine learning techniques to improve accuracy and real-time performance. Our findings highlight the effectiveness of deep learning models, specifically YOLOv10, in identifying and classifying traffic signs in challenging conditions and complex environments. Extensive experiments on our real-world dataset yielded competitive accuracy rates and robust performance. Future research will focus on the efficient deployment of these systems on edge devices, addressing challenges with hybrid models. Our study underscores the potential of advanced intelligent traffic systems to enhance road safety and support autonomous driving and ADAS. Future work should prioritize improving the adaptability, efficiency, and reliability of these systems in diverse real-world situations.

REFERENCES

[1]. bdnews24.com. (2024) Bangladesh lost 7,902 lives to road accidents in 2023: study. Accessed: 2024-11-01. [Online]. Available: https://bdnews24.com/bangladesh/pwh9c85xxr

[2]. The Business Standard. (2024) Country's 4,000km highways to be 4-6 lanes: Govt. Accessed: 2024-11-02. [Online]. Avail-able: https://www.tbsnews.net/bangladesh/countrys-4000km- highways-be-4-6-lanes-govt-355141

[3]. M. Manawadu and U. Wijenayake, "Voice-assisted real-time traffic sign recognition system using convolutional neural network," arXiv preprint arXiv:2404.07807, 2024.

[4]. A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, and G. Ding,"Yolov10: Real-time end-to-end object detection. arxiv 2024," arXiv preprint arXiv:2405.14458.

[5]. C.-Y. Wang and H.-Y. M. Liao, "Yolov1 to yolov10: The fastest and most accurate real-time object detection systems," arXiv preprint arXiv:2408.09332, 2024.

[6]. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation,"in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.

[7]. G. Zhang, Y. Peng, and H. Wang, "Road traffic sign detection method based on rts r-cnn instance segmentation network," Sensors, vol. 23, no. 14, p. 6543, 2023.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14101

[8]. Y. Sun and L. Chen, "Traffic sign recognition based on cnn and twin support vector machine hybrid model," Journal of Applied Mathematics and Physics, vol. 9, no. 12, pp. 3122–3142, 2021.
[9]. X. R. Lim, C. P. Lee, K. M. Lim, T. S. Ong, A. Alqahtani, and M. Ali,"Recent advances in traffic sign recognition: approaches and datasets,"Sensors, vol. 23, no. 10, p. 4674, 2023.
[10]. X. Gao, L. Chen, K. Wang, X. Xiong, H. Wang, and Y. Li, "Improved traffic sign detection algorithm based on faster r-cnn," Applied Sciences, vol. 12, no. 18, p. 8948, 2022.
[11]. K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," IEEE transactions on pattern analysis and machine intelligence, vol. 37, no. 9, pp. 1904–1916, 2015.
[12]. J. Redmon, "You only look once: Unified, real-time object detection,"in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
[13]. M. Flores-Calero, C. A. Astudillo, D. Guevara, J. Maza, B. S. Lita, B. Defaz, J. S. Ante, D. Zabala-Blanco, and J. M. Armingol Moreno,"Traffic sign detection and recognition using yolo object detection algorithm: A systematic review," Mathematics, vol. 12, no. 2, p. 297, 2024.
[14]. T.-Y. Ross and G. Doll´ar, "Focal loss for dense object detection," in proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2980–2988.
[15]. J. Redmon,"Yolo: Real-time object detection,"https://pjreddie.com/darknet/yolo/, accessed: 2024-10-31.
[16]. Roboflow, "Scaled-yolov4: Scaling cross stage partial network object detection," https://blog.roboflow.com/scaled-yolov4-tops-for efficientdet/, accessed: 2024-10-31.
[17]. J. Wang, Y. Chen, Z. Dong, and M. Gao, "Improved yolov5 network for real-time multi-scale traffic sign detection," Neural Computing and Applications, vol. 35, no. 10, pp. 7853–7865, 2023.
[18]. S.-S. Park, V.-T. Tran, and D.-E. Lee, "Application of various yolo models for computer vision-based real-time pothole detection," Applied Sciences, vol. 11, no. 23, p. 11229, 2021.
[19]. S. Qu, X. Yang, H. Zhou, and Y. Xie, "Improved yolov5-based for small traffic sign detection under complex weather," Scientific reports, vol. 13, no. 1, p. 16219, 2023.
[20]. S. Zhao, Y. Yuan, X. Wu, Y. Wang, and F. Zhang, "Yolov7-ts: A traffic sign detection model based on sub-pixel convolution and feature fusion,"Sensors, vol. 24, no. 3, p. 989, 2024.
[21]. Y. Luo, Y. Ci, H. Zhang, and L. Wu, "A yolov8-ce-based real-time traffic sign detection and identification method for autonomous vehicles,"Digital Transportation and Safety, vol. 3, no. 3, pp. 82–91, 2024.
[22]. S. M. M. Ahsan, S. Das, S. Kumar, and Z. La Tasriba, "A detailed study on bangladeshi road sign detection and recognition," in 2019 4th International Conference on Electrical Information and Communication Technology (EICT). IEEE, 2019, pp. 1–6.