Improved Detection of Bad Traffic Signs using Non-Maximum Suppression Ensemble

Rung-Ching Chen[1], Shao-Kuo Tai[2], Jun-Ci Chen[3], and Ibrahim Yahaya Garta*
Department of Information Management, Chaoyang University of Technology, Taichung, 41349, Taiwan, R.O.C
*s11114907@cyut.edu.tw +886 975847661

Abstract— **Algorithms for traffic sign detection and recognition affected by light and weather conditions have successfully been implemented. We propose an improved method for detecting traffic signs affected by varying lighting, adverse weather, and physical conditions. A transformer block is introduced into the backbone of Yolov5s and Yolov5m models to improve feature extractions. Predictions of the models are used to implement a Non-Maximum Suppression ensemble where the most confident bounding boxes are selected and the redundant ones are suppressed. The process has improved the overall mean average precision over the union of 50 (mAP@50) of the proposed models across all classes. We test our model on some GTSRD and TT100K datasets.**

I. Introduction

Traffic signs regulate traffic flow and provide useful information to drivers thereby mitigate the risk of accident. They are exposed to various environments, weather, lighting, human, and other factors that affect their visibility and effectiveness. Combinations of one or more of these factors can render the signs as 'bad' which may result in miss-detection and misclassification of those signs. When these signs become difficult to detect or classify correctly, it can lead to dangerous situations on the road.

Traffic Recognition Systems (TRS) are technologies designed to address the above challenges [1]. However, the systems may struggle to detect bad signs [2]. TRS is built on algorithms that are constantly improved to enhance its performance. Many such algorithms have been proposed and successfully implemented. Lai et al. [3] used data augmentation to generate occlusion, snow, fog, blur, and noise samples of the TT100K dataset. The authors train and evaluate an enhanced STC-YOLO model based on the Yolov5 model to detect small traffic signs in complex environments. Ahmed et al [4] proposed a dual-CNN approach to detect and classify traffic signs under challenging weather conditions. Zhang et al [5] proposed an enhanced YOLOv7 detection subnet with a multiscale feature fusion attention (MSFFA) structure to improve the detection of traffic signs in foggy scenes

This research aims to improve the detection performance of bad traffic signs. Our contributions are: (1) improve Yolov5s and Yolov5m by incorporating transformer block-C3TR module into their backbones to enhance feature extraction. (2) Implement Non--Maximum Suppression ensemble using predictions of the enhanced Yolov5s and Yolov5m models. (3) Collect a novel dataset of traffic signs affected by physical, lighting, and weather conditions.

II. Methodology

A. Dataset

We use 2400 bad traffic signs with eight classes namely good, occluded, faded, displaced, lighting affected, rusted, perforated, and defaced. Figure 1 shows samples of the dataset. Each of the class has 300 images. We use 1680 for training, 510 for validation, and 210 for tests.

Fig. 1 Samples of the dataset

B. Improved YOLO models and Non-Maximum Suppression Ensemble

We improve Yolov5s and Yolov5m models by introducing C3TR into their backbones, which enhances feature extraction. YOLOv5 is a state-of-the-art object detection model in the "You Only Look Once" (YOLO) family. Developed by Ultralytics, it is built on PyTorch and is known for its high speed and accuracy.

The C3TR module combines CSP (Cross Stage Partial), 3 Convolutions, and a Transformer Block to enhance object detection and classification by efficiently extracting features and capturing long-range dependencies.

Non-maximum suppression (NMS) Ensemble [6] is implemented by combining predictions of the improved Yolov5s and Yolov5m models. NMS selects the most confident bounding boxes and reduces false positives, thereby improving the accuracy of the proposed model. Figure 2 shows flowchart of the proposed model.

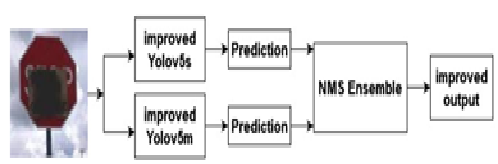

Fig. 2 Flowchart of the proposed model

C. Environment for experiments

We used Python 3.10.12 torch-2.2.1+cu121. The model was trained with 100 epochs, and 16 batch size, learning rate of 0.01..

III. RESULTS

The following equations evaluate the performance of the model.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

In equations 1, 2, and 3, True Positive (TP) is the outcome where the model correctly predicts the positive class. False Positive (FP) is an outcome where the model incorrectly predicts the positive class while False Negative (FN) is an outcome where the model incorrectly predicts the negative class.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \quad (3)$$

where N is the number of classes and AP_i is the Average Precision for class i. AP is a metric used to evaluate the accuracy of object detection models for a specific class.

After improvement, mAP@50 of Yolov5s and Yolov5m increase as shown in Table 1.

TABLE I

RESULTS OF YOLO MODELS AFTER IMPROVEMENT

Model	P	m AP@50	GFLOP
Yolov5s	0.786	0.818	16.3
Yolov5m	0.848	0.826	48.4

The ensemble of Yolov5s and Yolov5m has resulted in an increase of mAP@50 0.833 across all classes. Figure 3 shows the Precision-Recall curve of the NMS ensemble showing the performance of each class. The Perforated traffic sign and rusted traffic sign classes recorded high precision which indicates that the model the classes have distinct features that the model learnt and recognize.

We test our model on some GTRSD and TT100K images as shown in Figure 4.

On comparison, our model outperforms other state-of-the-art models as shown in Table 2.

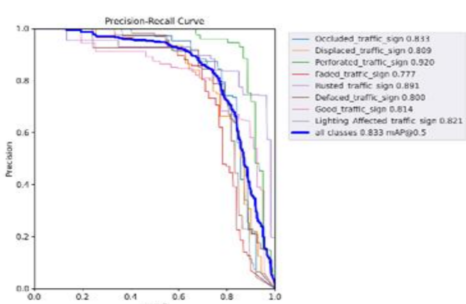

Fig. 3 Precision-Recall curve of NMS ensemble.

Fig. 4 Detection results of GTRS and TT100K datasets respectively.

TABLE III

COMPARISON OF THE PROPOSED MODEL WITH STATE-OF-THE-ART

Model	Dataset	mAP@50
YOLO-TSF [7]	Foggy-TT100K	0.831
Yolov5 [8]	TT100K	0.6514
Ours	Bad traffic signs	0.833

IV. CONCLUSIONS

In this research, we improve Yolov5s, and Yolov5m models by introducing C3TR to the backbones of their models. Predictions of the improved models are used to implement the Non-Maximum Suppression ensemble which improved the proposed model.

ACKNOWLEDGMENT

This paper is supported by NSTC, Taiwan. Project No. NSTC-112-2221-E-324-003 MY3, NSTC-111-2622-E-324-002 and NSTC-112-2221-E-324-011-MY2.

REFERENCES

[1] V. M. Shivanna, and J.-I. Guo, "Object Detection, Recognition, and Tracking Algorithms for ADASs-A Study on Recent Trends," *Sensors*, 24, 249. 2024.

[2] J. Gua, J. Lu, Y. Qu, C. Li. "Traffic-Sign Spotting in the Wild via Deep Features," In Proceedings of the IEEE Intelligence Vehicles Symposium (IV), Changshu, China, 2018, pp.120-125.

[3] H. Lai, L. Chen, W. Liu, Z. Yan, and S. Ye. "STC-YOLO: Small Object Detection Network for Traffic Signs in Complex Environments," *Sensors* 2023, 23,5307.

[4] S. Ahmed, U. Kamal, and M. K. Hasan, "DFR-TSD: A Deep Learning Based Framework for Robust Traffic Sign Detection Under Challenging Weather Conditions," in *IEEE Transactions on Intelligent Transportation* Systems, vol. 23, no. 6, pp. 5150-5162, June 2022.

[5] Zhang and X. Hu, "MSFFA-YOLO Network: Multiclass Object Detection for Traffic Investigations in Foggy Weather," in *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-12, 2023.

[6] R. Rahman, Z. B. Azad, and M. B. Hasan, "Densely-Populated Traffic Detection using YOLOv5 and Non-Maximum Suppression Ensembling," arXiv:2108.12118v1 [cs.CV] 27 Aug 2021.

[7] R. Li, Y. Chen, Y. Wang, and C. Sun, "YOLO-TSF: A Small Traffic Sign Detection Algorithm for Foggy Road Scenes," *Electronics*, 13, 3744, 2024.

[8] J. Wang, Y. Chen, Z. Dong and G. Mingyu," Improved YOLOv5 network for real-time multi-scale traffic sign detection," *Neural Comput & Applic* 35, pp. 7853–7865 2023.