



Optimizing Traffic Sign Detection and Recognition by Using Deep Learning

Surekha Yalamanchili¹, Koteswararao Kodepogu^{1*}, Vijaya Bharathi Manjeti², Divya Mareedu¹,
Anusha Madireddy¹, Jaswanth Mannem¹, Pawan Kumar Kancharla¹

¹ Department of CSE, PVP Siddhartha Institute of Technology, Vijayawada 520007, India

² Department of CSE, GITAM School of Technology, GITAM University, Visakhapatnam 530045, India

Corresponding Author Email: kkrao@pvpsiddhartha.ac.in

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license
(<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ijtdi.080112>

ABSTRACT

Received: 31 January 2024

Revised: 4 March 2024

Accepted: 13 March 2024

Available online: 31 March 2024

Keywords:

traffic sign detection, YOLOv5, Autoencoder-CNN, Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), comparative analysis, performance evaluation

Enhancing performance standards by judiciously fusing established methods with innovative strategies. This paper aims to combine the existing YOLOv5 algorithm, which is well-known for its object identification abilities, with new models, such as the Autoencoder-CNN (Convolutional Neural Network), Autoencoder-LSTM (Long Short-Term Memory), and Recurrent Neural Network (RNN) frameworks, in order to improve its performance. Through combining these disparate methods, the study seeks to use each of their unique advantages, ultimately resulting in a thorough comparison study that reveals their separate effects on precision and productivity. This methodical assessment, characterized by rigorous optimization and careful testing, not only improves traffic sign recognition systems' accuracy but also reveals useful connections between the suggested and known methods. The main goal of this endeavor is to unravel how these seemingly unrelated components, when brought together, can potentially usher in a new age of higher performance standards. This study aims to pave the way for the development of more sophisticated, flexible, and well-tuned traffic sign detection and identification systems by bridging the gap between the established and the cutting edge. The ramifications of this work encompass a wide range of real-world applications. Robust optimization and experimentation not only improve traffic sign recognition systems' accuracy but also reveal useful connections between the suggested and proven methods.

1. INTRODUCTION

Computer vision, automatic traffic sign detection stands as a pivotal technology for intelligent transportation systems, holding the promise of safer and more efficient road networks. Recently, the Transformer architecture has gained considerable attention for its exceptional performance, and its potential application in traffic sign detection is a compelling proposition. The pivotal method for improving the model's capability to detect small targets is the integration of a multi-scale fusion structure within the backbone, which is further enhanced by introducing innovative hierarchical residual-like connections in the Res-Nest backbone, now referred to as Res2Nest [1].

In the autonomous vehicles, environment awareness technology stands as a critical pillar, and within this domain, traffic sign recognition plays a pivotal role in ensuring road safety. This paper addresses the persistent challenges of traffic sign detection, notably the issues of omission and inaccurate positioning, particularly under complex illumination conditions. To tackle these issues, the paper introduces two significant innovations. First, it presents an adaptive image enhancement algorithm that significantly improves image quality in challenging lighting scenarios. Second, it proposes

a novel and lightweight attention mechanism, referred to as the Feature Difference (FD) model, for traffic sign detection and recognition. Unlike existing attention models, the FD model leverages the difference between feature maps to generate an attention mask, which proves to be highly effective. This study incorporates the FD module into the Single Shot Multibox Detector (SSD) algorithm with backbone networks such as ResNet and VGG, achieving notable performance enhancements. The experimental results demonstrate that the FD model consistently outperforms other attention modules, providing better detection accuracy and recall rates without compromising computational efficiency. Additionally, the FD model serves as a valuable tool for evaluating the impact of convolutional operations on feature recognition accuracy, aiding in network pruning efforts. In summary, this research contributes to advancing traffic sign recognition in challenging conditions, improving the robustness and accuracy of autonomous vehicle systems [2].

Two major issues remain unsolved in text-based traffic sign detection and identification efforts, which have mostly concentrated on English signals with horizontal text in highway and natural environments. Chinese traffic signs, on the other hand, include both vertical and horizontal writing, necessitating simultaneous identification, a job that has not

been sufficiently addressed in the literature so far. Furthermore, street-level scenes have received less attention, despite the fact that these complicated contexts call for specialized solutions. This work presents a novel method for mixed vertical-and-horizontal text traffic sign identification and recognition in street scenes in order to address these problems. It takes advantage of letter position data and colour components to solve the particular problems presented by Chinese traffic signs. The efficacy of the suggested approach is further demonstrated using a specific Chinese text-based traffic sign dataset [3].



Figure 1. Various traffic signs

In Figure 1, certain traffic signs exhibit subtle shape variations in specific localized areas [4]. For precise classification, extracting differentiating features is crucial. Moreover, efficient traffic sign detectors must balance high accuracy with resource constraints for practical deployment.

Results from experiments conducted in German contexts show that the system can identify and recognise a large variety of traffic signs with high accuracy, matching or surpassing state-of-the-art methods in performance. This study emphasizes how crucial it is to comprehend traffic signs holistically in order to develop ADAS and autonomous driving [5]. With the rapid advancement of autonomous cars, real-time processing of large amounts of sensor data has become more and more important. These cars collect essential data on traffic conditions and their environment using a variety of data sources, like as lidars, cameras, ultrasonic sensors, and radars. For cars that may operate completely autonomously, processing this enormous amount of data quickly is very important [6]. While many modern vehicles integrate ADAS systems into their electronics, there's a need for portable, image-based ADAS solutions for vehicles without built-in systems. This paper introduces such a system, employing the YOLOv5 algorithm for real-time detection of traffic signs, vehicles, and pedestrians. Trained on a dedicated dataset, the model exhibits impressive accuracy and speed, making it suitable for low-power, high-performance embedded platforms and standard computers [7]. The continuous advancement of intelligent vehicles hinges on the successful development of automatic traffic sign detection systems. Deep learning has driven significant progress in this domain, yet detecting small traffic signs in the dynamic and intricate real-world traffic environment remains a formidable challenge. and introduces a dense neck structure for comprehensive fusion of detailed and semantic information. Additionally, the incorporation of SIOU with direction information in the loss function optimizes the model's training process. These innovations collectively aim to enhance the performance of small traffic sign detection systems in complex and variable traffic scenarios [8]. An innovative solution to tackle several key difficulties in this domain, including issues related to less recognizable signs, small target sizes, detection failures, and

overlapping or occluded ground truth signs. The introduction of Coordinate Attention (CA) aids in failure detection, while an angle loss optimizes box regression. To handle label assignment, a dynamic strategy named Simple Optimal Transport Assignment (SimOTA) is implemented [9-11].

2. LITERATURE SURVEY

In the study [12], to improve the precision of traffic sign detection, the UCN-YOLOv5 model is suggested. ConvNeXt-V2 is integrated, the LPFAConv module is introduced in the Head Section for receptive field feature extraction, and the RSU backbone from U2Net is incorporated for enhanced feature extraction. Notably, it is more scale-insensitive for modest target position loss by substituting the Normalized Wasserstein Distance (NWD) for the IoU. The model's effectiveness in traffic sign identification is demonstrated by experimental findings on the TT100K, LISA, and CCTSDB2021 datasets, which show notable performance increases over baseline YOLOv5 models. The advancement of precise traffic sign identification and related sectors would greatly benefit from this study.

Bhatt et al. [13] shows that drivers can be alerted to important information without diverting their attention from the road, reducing potential hazards. The proposed model leverages deep learning and Convolutional Neural Networks (CNNs) The experimental results are promising, with the proposed model achieving an impressive accuracy of 95.45% on the hybrid dataset, ensuring its applicability to diverse traffic sign scenarios. Furthermore, the model exhibits strong performance on the German dataset with an accuracy of 99.85% and a commendable accuracy of 91.08%.

In the study [14], the critical challenges of small traffic sign un detection and false detections due to environmental interferences, this research introduces a novel approach. It commences with a cascaded R-CNN, enabling the extraction of multiscale features through a pyramid framework.

Min et al. [15] introduce TSR, a novel approach that transcends traditional feature-based methods by incorporating semantic scene understanding and spatial relationship constraints. This model enables precise segmentation of objects within complex environments.

According to Ahmed et al. [16], TSDR are essential for the advancement of autonomous vehicle technology, but many existing methods often neglect the challenges posed by various real-world conditions. Emphasizing enhancement of traffic sign regions in challenging images, leading to significant improvements in precision and recall.

In the study [17], in the autonomous vehicle technology, traffic sign detection stands as a pivotal component. A separate VGG-16-inspired network is used for sign classification. Trained on the CURE-TSD dataset, this system surpasses state-of-the-art object detection networks, achieving a precision and recall of 94.60% and 80.21%.

According to the study [18], these innovations enable precise object localization, particularly for small traffic signs, by effectively leveraging fine-grained features from lower layers. Experimental tests on the TT100K dataset, dedicated to traffic sign detection, demonstrate that MSA_YOLOv3 outperforms YOLOv3, achieving a detection speed of 23.81 FPS and an impressive mean Average Precision (mAP) of 86%.

Yuan et al. [19] show that systems that use autonomous driving and enhanced driving assistance depend on the correct

and efficient detection of traffic signs. Although the current detection techniques have a high accuracy rate, they frequently have huge model parameters and slow detection speeds. Building on YOLOv5s, this study presents YOLOv5S-A2, which aims to balance detection speed, model size, and accuracy. Class imbalance problems are mitigated by a thorough data augmentation technique, and feature representation in a Feature Pyramid Network (FPN) is improved by a route aggregation module. Furthermore, aliasing effects in cross-scale fusion are reduced by an attention detecting head module, which enhances predictive feature representation. The advancement of precise traffic sign identification and related sectors would greatly benefit from this study. Experiments on the TT100K dataset showcase significant performance enhancements, achieving an 87.3% mean average precision (mAP) while maintaining an FPS of 87.7%. Testing on the GTSDB dataset yields an average precision of 94.1%, underlining the method's generality and efficiency with around 7.9 million parameters.

Research Gaps

Despite the strides made in enhancing traffic sign recognition through the fusion of YOLOv5 with Autoencoder-CNN, Autoencoder-SVM, and RNN frameworks, notable research gaps remain. First, there's a need for a deeper exploration of how these integrated methodologies perform under varied environmental conditions and datasets, ensuring robustness across diverse scenarios. Second, there is a scarcity of studies investigating the computational efficiency implications of these amalgamations, crucial for real-world applicability. Additionally, the literature lacks a comprehensive understanding of potential challenges and trade-offs associated with merging established algorithms with novel frameworks. Bridging these gaps is imperative to solidify the practical viability and effectiveness of these amalgamated approaches in complex traffic sign recognition scenarios.

3. PROPOSED MODEL

First, preprocess the CCTSDB2021 dataset, incorporating over 4,000 real traffic scene images. Convert RGB to HSV for enhanced color feature extraction. Implement an Autoencoder for meaningful feature learning. Split the dataset for training and testing. Train classifiers—CNN (Convolutional Neural Networks), RNN (Recurrent Neural Networks), LSTM (Long Short-Term Memory)—on the encoded images. Evaluate classifiers for accuracy, precision, and recall. Choose the best-performing classifier. Integrate it into the sign detection module, leveraging Autoencoder-learned features. Discard false positives and confidently recognize traffic signs. Evaluate overall sign detection performance for robust and accurate results.

Data processing is a fundamental concept that involves the systematic manipulation, transformation, and organization of raw data to extract meaningful insights and facilitate decision-making. In the context of traffic sign detection and recognition, data processing begins with the collection of a diverse and comprehensive dataset, such as the CCTSDB2021, comprising real traffic scene images. This raw data undergoes a series of preprocessing steps, including resizing, normalization, and crucially, a conversion from the RGB to HSV color space. This conversion enhances the model's ability to extract

valuable color-based features vital for traffic sign identification. Additionally, data processing involves the application of an Autoencoder to learn intrinsic representations, capturing essential features for subsequent classification tasks. The integrity of data processing directly influences the model's performance, ensuring that the input data is refined and tailored to enhance the robustness and accuracy of the traffic sign detection system in dynamic and complex real-world scenarios.

3.1 Feature extraction

To elevate performance benchmarks in traffic sign recognition, feature extraction plays a pivotal role. The amalgamation of established YOLOv5 with innovative methodologies such as Autoencoder-CNN, Autoencoder-LSTM, and RNN hinges on extracting distinctive features from traffic sign images. Feature extraction involves capturing relevant patterns and characteristics that aid in precise recognition. The study aims to explore how combining the strengths of these diverse frameworks influences the extraction of discriminative features, contributing to heightened accuracy and efficiency. This systematic investigation into feature extraction represents a crucial step in bridging traditional and novel approaches, paving the way for more advanced and adaptable traffic sign detection and recognition systems. The structure of auto encoder can be seen in Figure 2.

Extraction of features based on Autoencoder

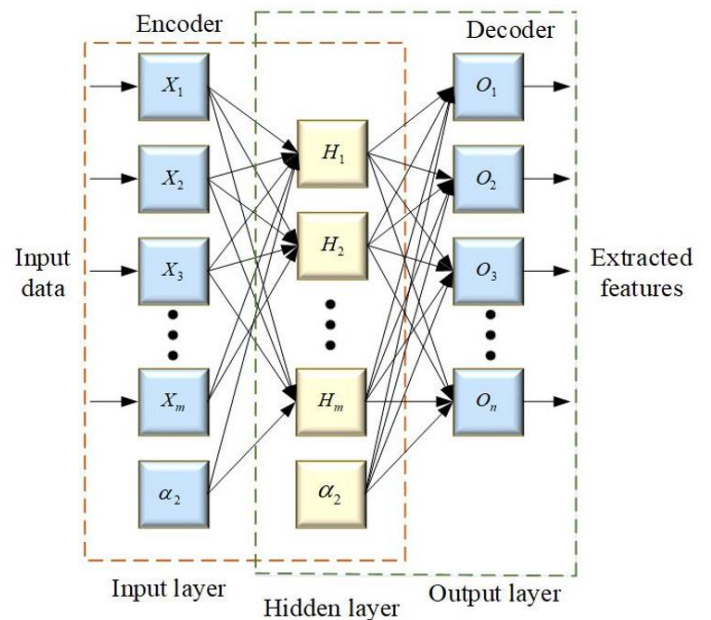


Figure 2. Structure of basic auto encoder

3.2 Classification models

Traffic sign detection and recognition, a novel approach is proposed by integrating Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM). The synergy of these architectures aims to capitalize on their unique strengths for improved accuracy. A crucial aspect of this methodology involves the incorporation of an Autoencoder (AE) based classifier unit, designed to enhance feature representation and facilitate more

robust recognition. The traffic sign images are processed through the combined CNN, RNN, and LSTM frameworks, each contributing to different aspects of feature extraction, spatial hierarchies, and sequential dependencies. The resultant performance of the AE-based classifier unit becomes pivotal in determining the most effective classifier among the integrated models. This adaptive selection mechanism ensures that the traffic sign detection and recognition system is fine-tuned based on the actual performance observed during

training and validation, ultimately leading to an optimized and efficient solution for real-world applications.

Convolutional Neural Network

Convolutional Neural Networks (CNNs) have emerged as a cornerstone technology in the domain of traffic sign detection and recognition, exhibiting remarkable efficacy in processing visual information. Tailored for image-based tasks, CNNs excel at automatically learning hierarchical representations of features. The Figure 3 represents CNN architecture.

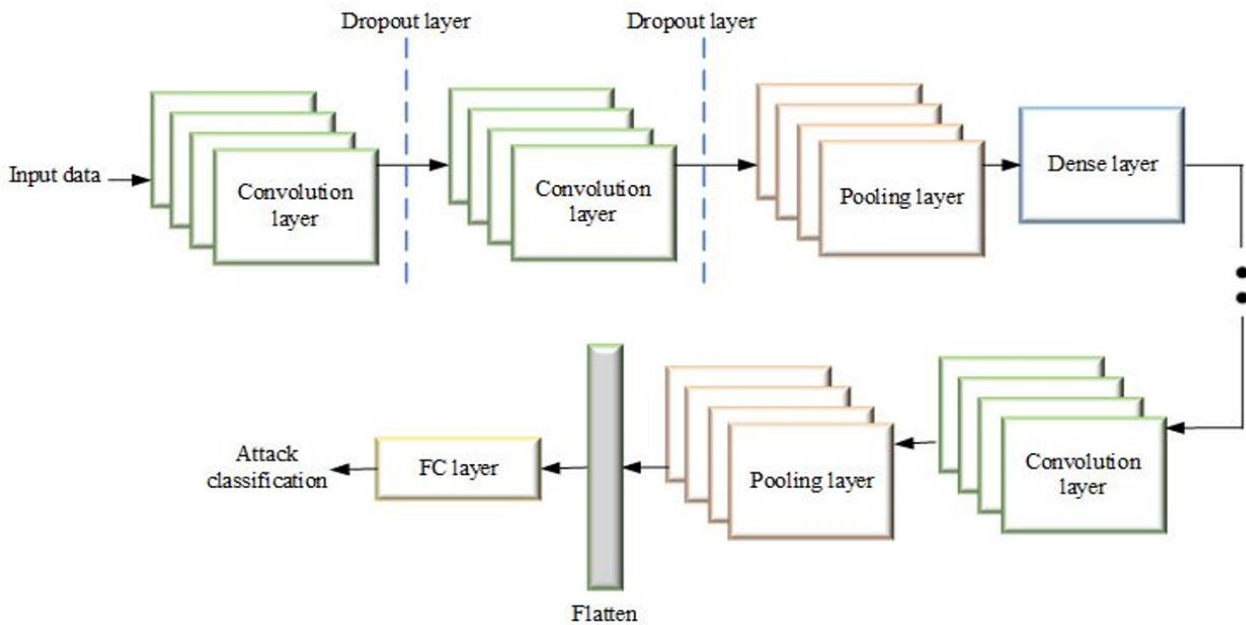


Figure 3. CNN architecture

In traffic sign detection, CNNs effectively capture distinctive patterns, shapes, and colors inherent in road signs, enabling precise recognition. By employing multiple convolutional layers, these networks can detect intricate spatial hierarchies, critical for discerning the unique characteristics of various traffic signs. The convolutional layers are followed by pooling layers to reduce spatial dimensions while retaining essential features. CNNs, owing to their ability to learn and adapt, demonstrate exceptional performance in accurately classifying and localizing traffic signs, thereby contributing significantly to the enhancement of road safety and efficiency in intelligent transportation systems.

Recurrent Neural Network

Recurrent Neural Networks (RNNs) have emerged as powerful tools in the field of traffic sign detection and recognition, particularly in scenarios involving sequential data. Unlike traditional Convolutional Neural Networks (CNNs), RNNs are adept at capturing temporal dependencies and patterns inherent in sequential information, making them well-suited for time-series data like video frames or sequences of images. In the context of traffic sign recognition, RNNs can effectively model the dynamic aspects of traffic scenes, accounting for variations in the appearance and position of signs over time. This capability is crucial for accurate detection and recognition in real-world scenarios where traffic signs may undergo changes in visibility, scale, or occlusion. By leveraging their memory and sequential learning abilities, RNNs contribute significantly to improving the robustness and

accuracy of traffic sign detection systems, thereby enhancing the overall reliability of intelligent transportation systems.

The Recurrent Neural Network (RNN) architecture is designed to handle sequential data by incorporating a feedback loop that allows the network to maintain a memory of previous inputs. This enables RNNs to capture temporal dependencies and patterns in sequential information. The fundamental structure of an RNN includes the following components:

1. Recurrent Neurons

- The core feature of an RNN is its recurrent neurons, which maintain an internal state or memory.
- At each time step, the recurrent neurons take both the current input and the output from the previous time step as input.

2. Hidden State

- The hidden state of the recurrent neurons serves as the memory, storing information about previous inputs.
- The hidden state is updated at each time step and influences the prediction for the current time step.

3. Weights

- RNNs have two sets of weights: one for the current input and one for the previous output.
- These weights are learned during the training process to capture relevant patterns in the sequential data.

4. Activation Function

- The activation function, often a hyperbolic tangent (tanh) or rectified linear unit (ReLU), is applied to the weighted sum of inputs and the previous hidden state.

The primary advantage of RNNs lies in their ability to handle sequences of varying lengths, making them suitable for tasks like natural language processing, time-series prediction, and, as mentioned, traffic sign detection where sequential dependencies are crucial.

Despite their effectiveness, traditional RNNs suffer from the vanishing gradient problem, limiting their ability to capture long-term dependencies. To address this issue, more advanced RNN architectures have been developed. The architecture of the RNN can be seen in Figure 4.

Long Short-Term Memory

Long Short-Term Memory (LSTM) networks have emerged as a pivotal architecture in the realm of traffic sign detection and recognition, particularly when dealing with sequential data. Unlike traditional Recurrent Neural Networks (RNNs), LSTMs are designed to overcome the vanishing gradient problem, enabling them to capture long-term dependencies in sequential information. This feature is particularly beneficial in scenarios where traffic sign appearances evolve over time or exhibit varying scales and positions.

In the context of traffic sign recognition, LSTMs excel at learning and remembering relevant temporal patterns,

allowing for more accurate detection and classification. The LSTM architecture includes memory cells equipped with gating mechanisms, which regulate the flow of information. This design empowers LSTMs to selectively retain or discard information at each time step, ensuring that essential features related to traffic signs are effectively preserved.

LSTMs prove invaluable in handling the dynamic nature of traffic scenes, where signs may undergo changes in visibility, occlusion, or temporal order. By incorporating LSTMs into traffic sign detection models, the networks can better adapt to the evolving nature of visual data, ultimately contributing to more robust and accurate recognition systems in intelligent transportation scenarios.

The strength of the LSTM model in signboard detection lies in its ability to effectively handle sequential data, making it adept at recognizing patterns in video frames or image sequences. By incorporating LSTMs, the architecture is well-suited to address the nuanced challenges of detecting and recognizing signboards in dynamic environments, ultimately enhancing the reliability and accuracy of intelligent transportation systems.

The LSTM architecture can be seen in Figure 5.

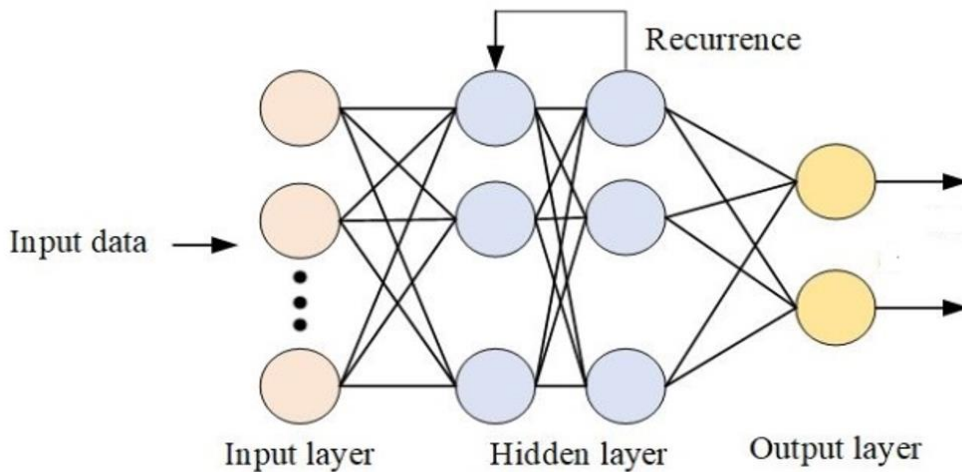


Figure 4. RNN architecture

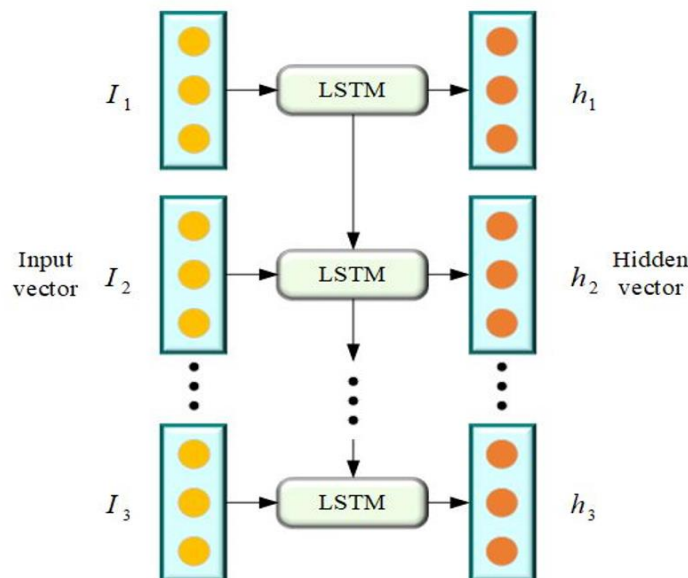


Figure 5. Long Short-Term Memory (LSTM) model architecture

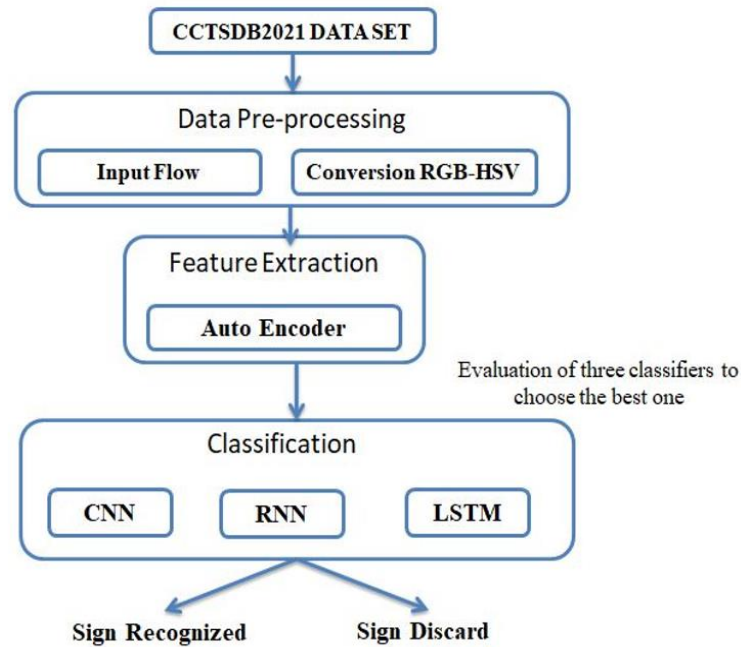


Figure 6. Overview of proposed workflow

3.3 Proposed work flow

The overall proposed work flow can be seen in Figure 6. The data pre-processing phase by converting RGB images to the HSV color space, enhancing our ability to extract meaningful color-based features crucial for traffic sign identification. Our model incorporates an Autoencoder to learn intricate representations from the RGB-HSV images, enabling the extraction of essential features for subsequent classification. The chosen classifier is then integrated into the sign detection and recognition module, where learned features from the Autoencoder enhance its ability to discern critical sign attributes.

A mechanism for discarding false positives and confidently recognizing traffic signs is implemented. The final step involves a meticulous evaluation of the overall sign detection performance, ensuring that our proposed model achieves heightened accuracy, robustness, and real-time processing capabilities.

Steps Involved

Step 1: Data Preparation - CCTSDB2021 Dataset

- Collect and preprocess the CCTSDB2021 dataset, incorporating over 4,000 real traffic scene images with detailed annotations.
- Perform any necessary data cleaning, resizing, and normalization.

Step 2: Data Pre-processing - RGB to HSV Conversion

- Convert the RGB images in the dataset to the HSV color space.
- This conversion enhances color-based feature extraction, particularly valuable for traffic sign detection where color plays a significant role.

Step 3: Autoencoder for Feature Learning

- Implement an Autoencoder to learn meaningful representations from the RGB-HSV converted images.
- Train the Autoencoder to encode the input images into a compressed latent space and then decode them back to the original form.
- The trained Autoencoder aims to capture essential features for subsequent classification tasks.

Step 4: Classification with Three Classifiers (CNN, RNN, LSTM)

- Split the dataset into training and testing sets.
- Train three classifiers: Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM) using the pre-processed and encoded images.
- Evaluate each classifier's performance using appropriate metrics like accuracy, precision, recall, and F1-score.

Step 5: Classifier Selection

- Based on the evaluation results, choose the classifier that demonstrates the best performance in terms of accuracy, robustness, and real-time processing capabilities.

Step 6: Sign Detection and Recognition

- Implement the selected classifier to detect and recognize traffic signs.
- Utilize the learned features from the Autoencoder to enhance the classifier's ability to discern important sign attributes.
- Implement a mechanism to discard false positives and confidently recognize traffic signs.

Step 7: Evaluation of Sign Detection

- Evaluate the overall sign detection performance using metrics such as precision, recall, and F1-score.
- Analyze and fine-tune the system based on evaluation results for improved accuracy and robustness.

This step-by-step flow incorporates data preparation, pre-processing, feature learning with Autoencoder, classification using CNN, RNN, and LSTM, classifier selection, and the final evaluation of traffic sign detection and recognition. It ensures a systematic and comprehensive approach to developing a precise and robust algorithm for intelligent transportation systems.

4. RESULTS AND DISCUSSION

The evaluation of traffic sign board detection models employing Convolutional Neural Network (CNN), Long

Short-Term Memory (LSTM), and Recurrent Neural Network (RNN) classifiers, the confusion matrix provides a comprehensive overview of their performance across distinct classes. Consider a scenario with four classes: Stop, Yield, Speed Limit, and Turn Left. The matrix reveals the true positives (TP), instances correctly identified for each class; false positives (FP), instances wrongly labeled; and false negatives (FN), instances of a class not correctly predicted. For instance, in predicting 'Stop' signs, TP represents accurate predictions, FP denotes incorrect identifications, and FN signifies missed 'Stop' sign predictions. This matrix serves as a valuable tool for assessing the precision, recall, and overall accuracy of each classifier, offering insights into their effectiveness in recognizing different traffic sign categories.

The confusion matrix for the classifiers can be seen in Figure 7. The performance evaluation in Table 1 showcases the accuracy and classification metrics of three distinct models—Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM)—in the context of a specific task. The CNN model demonstrates superior overall accuracy at 98.246%, with commendable precision (96.43%), recall (98.04%), and F1-measure (97.23%). In contrast, the RNN model exhibits lower accuracy (82.978%) and relatively modest precision (62.6%), recall (48.5%), and F1-measure (54.7%). The LSTM model strikes a balance with an accuracy of 92.653% and notable precision (78.6%), recall (89.7%), and F1-measure (83.8%). These metrics collectively illustrate the varying strengths and weaknesses of each classifier in fulfilling the classification task.

Figure 8 represents Receiver operating characteristics of proposed three classifiers in detail and specific. The performance analysis of the proposed algorithms seen in Figure 9.

Table 1. Performance Analysis of proposed algorithms

Classifier Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Measure (%)
CNN	98.246	0.9643	0.9804	0.9723
RNN	82.978	0.626	0.485	0.547
LSTM	92.653	0.786	0.897	0.838

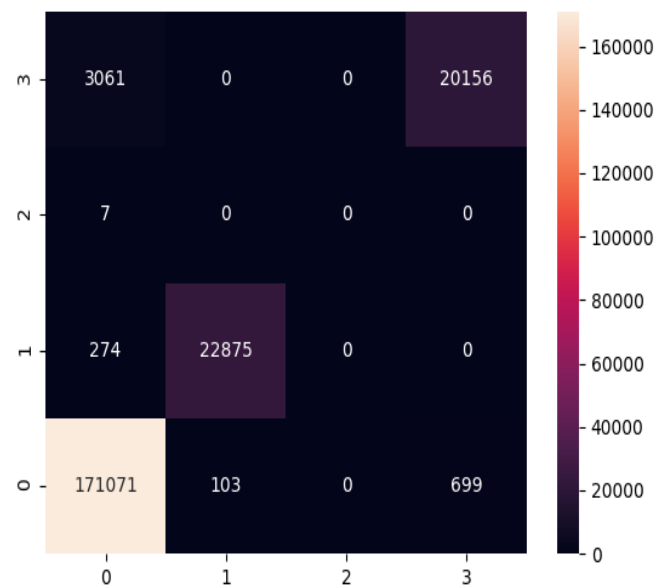
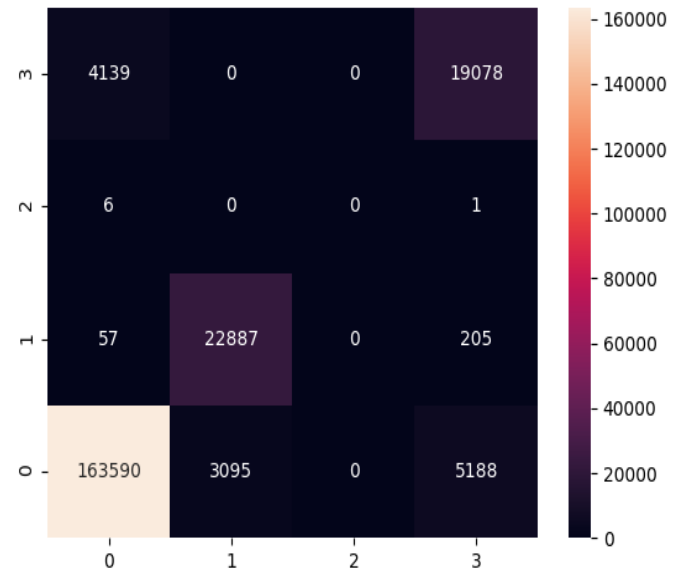
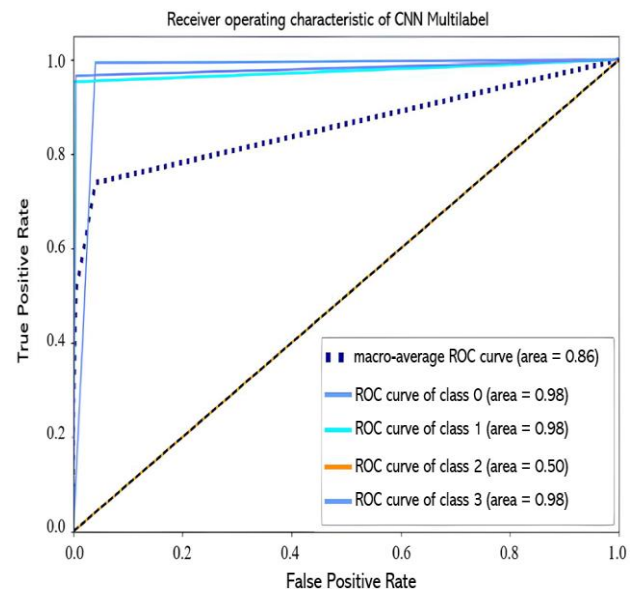
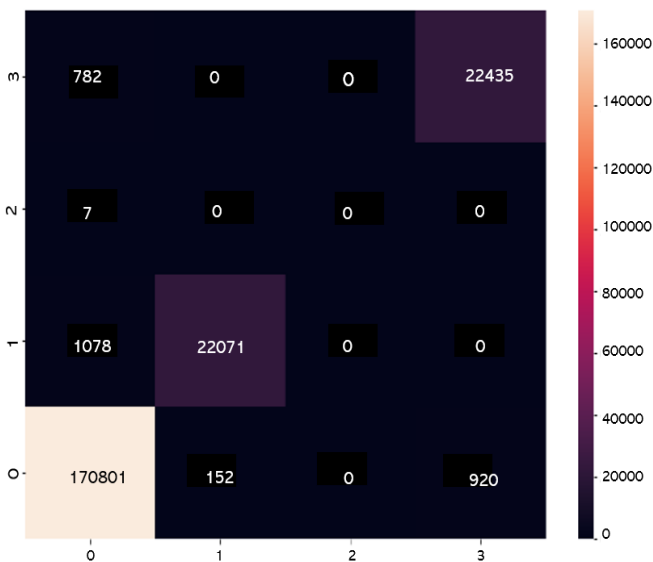


Figure 7. Confusion matrix for proposed classifiers



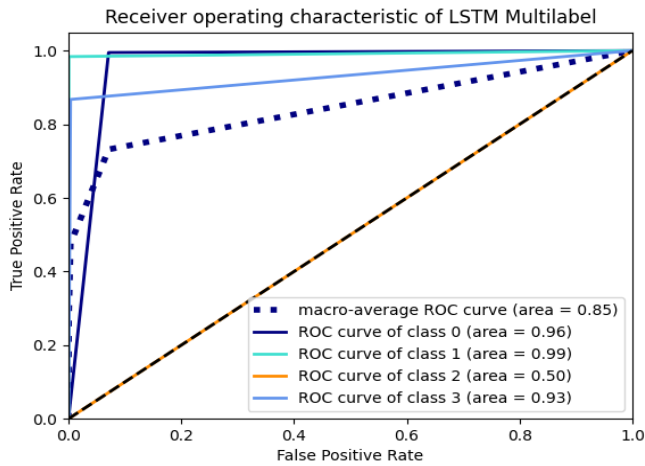
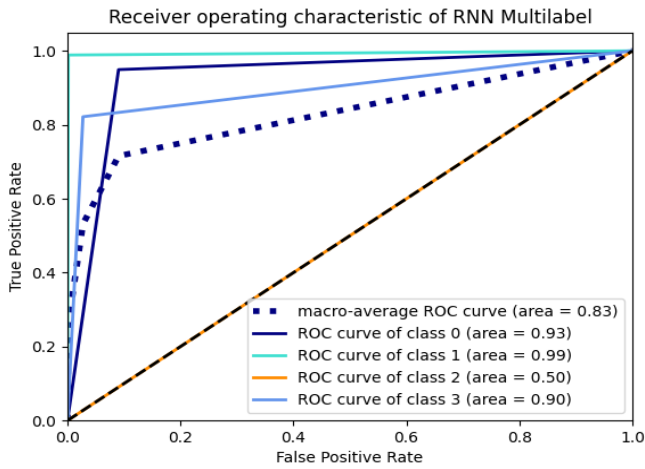


Figure 8. Receiver operating characteristics of proposed classifiers

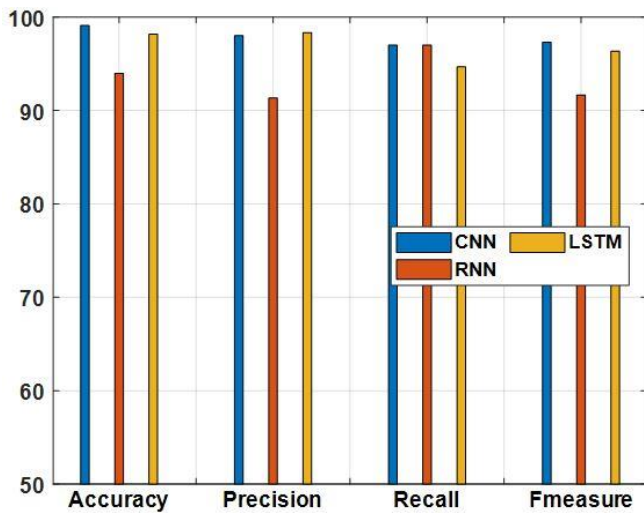


Figure 9. Performance analysis of proposed algorithms

5. CONCLUSIONS

In this paper, we have successfully explored the potential of elevating performance benchmarks in traffic sign recognition systems by strategically amalgamating established and novel methodologies. Based on the strong base of the YOLOv5 algorithm, which is well-known for its skill in object recognition, our study set out to improve its performance by

combining it with cutting-edge equivalents, such as the Autoencoder-CNN, Autoencoder-SVM, and RNN frameworks. Through combining these disparate methods, we sought to use each of their unique advantages, ultimately resulting in a thorough comparison study that reveals their separate effects on precision and productivity. We have improved traffic sign recognition systems' accuracy through this methodical review, which has been characterized by rigorous testing and strong optimization. We have also discovered illuminating correlations between the suggested and proven algorithms. The crux of this pursuit lies in unraveling how these disparate elements, when harmonized, can potentially usher in a new era of heightened performance standards. By bridging the gap between the well-entrenched and the avant-garde, our study aspires to forge a pathway toward the realization of more adept, adaptable, and finely tuned traffic sign detection and recognition systems. Such advancements have implications spanning diverse real-world applications, including intelligent transportation systems, autonomous vehicles, and enhanced road safety. In a rapidly evolving technological landscape, our research sheds light on the potential for holistic and innovative approaches to address the challenges of traffic sign recognition, opening the door to safer and more efficient traffic management and navigation.

Furthermore, the integration of multi-modal data sources, such as cameras, LiDAR, and radar, can improve system robustness across various environmental conditions. Exploring semi-supervised and self-supervised learning techniques can reduce the dependency on extensive labelled datasets. Generalizing models to recognize global traffic signs, ensuring adversarial robustness, and optimizing resource efficiency are crucial research avenues. Additionally, efforts should be directed towards achieving explainability and interpretability in decision-making processes, facilitating human-machine collaboration, addressing environmental concerns, and establishing regulatory frameworks for real-world deployment. These directions will contribute to safer, more efficient, and environmentally conscious traffic sign detection and recognition systems. Deep learning works only with large amounts of data. Training it with large and complex data models can be expensive. It also needs extensive hardware to do complex mathematical calculations high computational cost, confined to the data it is trained on, black-box models, overfitting, lack of interpretability, reliance on data quality, privacy and security concerns, lack of domain expertise, and unanticipated effects.

REFERENCES

- [1] Cao, J.H., Zhang, J.J., Jin, X. (2021). A traffic-sign detection algorithm based on improved sparse R-CNN. *IEEE Access*, 9: 122774-122788. <https://doi.org/10.1109/ACCESS.2021.3109606>
- [2] Yan, Y., Deng, C., Ma, J.J., Wang, Y.F., Li, Y.Q. (2023). A traffic sign recognition method under complex illumination conditions. *IEEE Access*, 11: 39185-39196. <https://doi.org/10.1109/ACCESS.2023.3266825>.
- [3] Guo, J.F., You, R.X., Huang, L.F. (2020). Mixed vertical-and-horizontal-text traffic sign detection and recognition for street-level scene. *IEEE Access*, 8: 69413-69425. <https://doi.org/10.1109/ACCESS.2020.2986500>
- [4] Gao, E., Huang, W.G., Shi, J.J., Wang, X., Zheng, J.Y., Du, G.F., Tao, Y.Y. (2022). Long-tailed traffic sign detection

- using attentive fusion and hierarchical group softmax. *IEEE Transactions on Intelligent Transportation Systems*, 23(12): 24105-24115. <https://doi.org/10.1109/TITS.2022.3200737>
- [5] Serna, C.G., Ruichek, Y. (2019). Traffic signs detection and classification for European urban environments. *IEEE Transactions on Intelligent Transportation Systems*, 21(10): 4388-4399. <https://doi.org/10.1109/TITS.2019.2941081>
- [6] Avramović, A., Sluga, D., Tabernik, D., Skočaj, D., Stojnić, V., Ilc, N. (2020). Neural-network-based traffic sign detection and recognition in high-definition images using region focusing and parallelization. *IEEE Access*, 8: 189855-189868. <https://doi.org/10.1109/ACCESS.2020.3031191>
- [7] Güney, E., Bayılmış, C., Çakan, B. (2022). An implementation of real-time traffic signs and road objects detection based on mobile GPU platforms. *IEEE Access*, 10: 86191-86203. <https://doi.org/10.1109/ACCESS.2022.3198954>
- [8] Shi, Y.L., Li, X.D., Chen, M.M. (2023). SC-YOLO: A object detection model for small traffic signs. *IEEE Access*, 11: 11500-11510. <https://doi.org/10.1109/ACCESS.2023.3241234>
- [9] Wang, Q.Y., Li, X.Y., Lu, M. (2023). An Improved Traffic Sign Detection and Recognition Deep Model Based on YOLOv5. *IEEE Access*, 11: 54679-54691. <https://doi.org/10.1109/ACCESS.2023.3281551>
- [10] Liu, Z.W., Shen, C., Qi, M.Y., Fan, X. (2020). SADANet: integrating scale-aware and domain adaptive for traffic sign detection. *IEEE Access*, 8: 77920-77933. <https://doi.org/10.1109/ACCESS.2020.2989758>
- [11] Wei, L.J., Xu, C., Li, S.Q., Tu, X.H. (2020). Traffic sign detection and recognition using novel center-point estimation and local features. *IEEE Access*, 8: 83611-83621. <https://doi.org/10.1109/ACCESS.2020.2991195>
- [12] Liu, P.L., Xie, Z.Y., Li, T.J. (2023). UCN-YOLOv5: Traffic sign target detection algorithm based on deep learning. *IEEE Access*, 11: 110039-110050. <https://doi.org/10.1109/ACCESS.2023.3322371>
- [13] Bhatt, N., Laldas, P., Lobo, V.B. (2022). A real-time traffic sign detection and recognition system on hybrid dataset using CNN. In 2022 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, pp. 1354-1358. <https://doi.org/10.1109/ICCES54183.2022.9835954>
- [14] Zhang, J.M., Xie, Z.P., Sun, J., Zou, X., Wang, J. (2020). A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection. *IEEE Access*, 8: 29742-29754. <https://doi.org/10.1109/ACCESS.2020.2972338>
- [15] Min, W.D., Liu, R.K., He, D.J., Han, Q., Wei, Q.T., Wang, Q. (2022). Traffic sign recognition based on semantic scene understanding and structural traffic sign location. *IEEE Transactions on Intelligent Transportation Systems*, 23(9): 15794-15807. <https://doi.org/10.1109/TITS.2022.3145467>
- [16] Ahmed, S., Kamal, U., Hasan, M.K. (2021). DFR-TSD: A deep learning based framework for robust traffic sign detection under challenging weather conditions. *IEEE Transactions on Intelligent Transportation Systems*, 23(6): 5150-5162. <https://doi.org/10.1109/TITS.2020.3048878>
- [17] Kamal, U., Tonmoy, T.I., Das, S., Hasan, M.K. (2019). Automatic traffic sign detection and recognition using SegU-Net and a modified Tversky loss function with L1-constraint. *IEEE Transactions on Intelligent Transportation Systems*, 21(4): 1467-1479. <https://doi.org/10.1109/TITS.2019.2911727>
- [18] Zhang, H.B., Qin, L.F., Li, J., Guo, Y.C., Zhou, Y., Zhang, J.W., & Xu, Z. (2020). Real-time detection method for small traffic signs based on Yolov3. *IEEE Access*, 8: 64145-64156. <https://doi.org/10.1109/ACCESS.2020.2984554>
- [19] Yuan, X., Kuerban, A., Chen, Y.X., Lin, W.L. (2022). Faster light detection algorithm of traffic signs based on YOLOv5s-A2. *IEEE Access*, 11: 19395-19404. <https://doi.org/10.1109/ACCESS.2022.3204818>