2024 International Conference on Recent Progresses in Science, Engineering and Technology (ICRPSET)
Faculty of Engineering, University of Rajshahi, Rajshahi-6205, Bangladesh, 07-08 December, 2024

Vehicle Classification and Detection Using YOLOv8: A Study on Highway Traffic Analysis

Naima Islam
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
naimaislam204100@gmail.com

Sajeeb Kumar Ray
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
sajeebbro2002@gmail.com

Md. Anwar Hossain
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
manwar.ice@pust.ac.bd

Mirza AFM Rashidul Hasan
Department of Information and Communication Engineering
University of Rajshahi
Rajshahi, Bangladesh
mirzahasanice@gmail.com

Alamin
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
alamin.200610@s.pust.ac.bd

Md. Basim Al Zabir Shammo
Department of Information and Communication Engineering
Pabna University of Science and Technology
Pabna, Bangladesh
basim.210628@pust.ac.bd

***Abstract*—The accelerated advancement of urban infrastructure has underscored the significance of smart city technologies, especially within traffic management, wherein vehicle detection and classification assume a crucial role. This manuscript presents a vehicle identification framework specifically designed for the roadways of Bangladesh, employing the YOLOv8n model to optimize traffic flow, alleviate congestion, and augment road safety. The system utilizes a custom dataset from real-time traffic footage in Bangladesh, classifying vehicles into eight categories. As part of this study, we evaluated our model utilizing Bangladeshi real-time traffic video data. The study showcases the model's capability to navigate intricate urban settings, although accuracy enhancement is necessary, especially in congested traffic. The findings underscore the model's potential for improving traffic surveillance and supporting smart city developments.**

Keywords—YOLOv8n, Vehicle detection, Traffic management, Bangladeshi roads, Real-time video analysis, Intelligent transportation systems.

I. Introduction

The idea of a smart city has gained a lot of attention in the quickly changing urban development landscape. A smart city uses technology to improve quality of life, increase the efficiency of urban services, and encourage sustainable development. One of the most important aspects of creating a smart city is intelligent traffic management, where real-time vehicle detection and classification are vital for optimizing traffic flow, minimizing congestion, and guaranteeing road safety all essential components of contemporary urban infrastructure.

Vehicle detection and categorization have been extensively studied in the modern era. As a result of the field's many developments, a wide range of methods and models have been created to increase precision and effectiveness in diverse applications. Mahi et al. [1, 2] employed the automatic wrong-way vehicle identification system that is intended to improve traffic monitoring and lower accident rates with the use of direction identification, object detection, and DeepSORT tracking. Samuel et al. [3, 4] proposed a YOLOv4-based system for classifying and counting vehicles. Unlike previous methods that merely count vehicles without categorization, their methodology improves traffic monitoring by combining both vehicle classification and counting into a single system. Bakirci, M. [5, 6] proposed the YOLOv8n model for vehicle detection using the images captured from UAV missions. Terada et al. [7, 8] suggested a Yolov2-tiny technique for vehicle identification in the article. The initial step in training the network is to fine-tune its parameters and cluster data using the K-means method. The detection component is trained on the darknet. Utilizing the enhanced CNN network, the kind of vehicle is classified. After modifying the network using the AlexNet method, they include SPP to address the issue of low classification accuracy caused by picture resizing and rescaling. Yun et al. [9] employed a real-time highway vehicle object dataset for an end-to-end highway vehicle identification model in this research publication, where they suggested the YOLO-v3 method for object detection. Deep learning systems are leveraged for various classification and detection functionalities [10]. An improved Convolutional Neural Network architecture based on deep learning methods was proposed by Chia-Chi Tsai et al. [11] for vehicle identification and classification systems utilized in intelligent transportation applications. Siraj et al. [12] suggested a YOLOv8-inspired framework for identifying and classifying vehicles unique to Bangladesh, incorporating transfer learning from aerial video sources. A semi-supervised convolutional neural network strategy was suggested by Dong et al. [13] for the categorization of cars based on their front views. However, the CNN-trained features are too skewed to function well in raster pictures. A traffic video analysis system based on computer vision techniques was introduced by Ahmad Arinaldi et al. [14]. The primary function of the system is to recognize and classify automobiles. To this end, two models have been developed: the first is an MoG + SVM system, while the second is based on Faster RCNN, a newly popular deep learning architecture for object recognition in pictures. A YOLO algorithm-based object recognition technique for identifying moving automobiles versus traffic was proposed by Rahman et al. [15, 16]. Bounding boxes for vehicles are made. After that, the direction is determined using the centroid-tracker approach. Sabina et al. [17] intended to evaluate two object recognition systems that used CNN to recognize objects within frames. They assessed how well the MobileNet SSD model and the YOLOv5s object identification model performed in various scenarios. While MobileNet SSD showed a better detection time, YOLOv5s produced more accurate findings. All they did, though, was compare these two detection models. Tejal Palwankar and Kushal Kothari suggested a deep learning-based object identification system that makes use of SSD and MobileNet to provide an effective method for both tracking and detection tasks [18].

In this study, our main goal is to create a vehicle identification system specifically designed for Bangladeshi roads, which will enhance traffic control and safety and forward the goal of smart cities. Bangladesh's highways are seeing an increase in the number of vehicles, hence it is imperative to create effective technology for traffic monitoring and analysis. Our work is relevant to different smart city efforts since it uses an advanced YOLOv8n model and a custom dataset to recognize and categorize vehicles from real-time video footage. We think that this technology will contribute to safer and more effective transportation systems as well as traffic monitoring, ultimately boosting Bangladesh's development of smart cities.

The main contribution of this study is the following:

- Using the YOLOv8n model, a vehicle detection system with the ability to recognize and categorize various vehicle kinds from real-time video footage was developed specifically for Bangladeshi roadways.
- By increasing traffic flow, safety, and the effectiveness of the road infrastructure in metropolitan areas, the system is made to improve traffic monitoring and management, which supports the larger goal of developing smart cities.
- As part of this study, we evaluated our model utilizing Bangladeshi real-time traffic video data.

II. Materials and Methods

A. System Utilized

The investigation engaged the Kaggle platform, which supplies superior computational resources ideal for machine learning processes. It features an interactive Jupyter notebook interface with access to NVIDIA Tesla P100 GPUs, crucial for efficiently training the YOLOv8n model. The system's specifications included 13 GB of RAM and 20 GB of storage, accommodating large datasets required for vehicle detection. Kaggle operates on a Linux-based environment, ensuring compatibility with various Python libraries such as Ultralytics YOLOv8n, OpenCV, and NumPy. The platform's integration of these libraries enabled real-time training, validation, and testing of the model. Video files were directly uploaded to Kaggle, and outputs were saved for subsequent analysis. This cloud infrastructure supported effective workflow management for vehicle detection tasks.

B. Dataset Preparation

This study constructed a dataset utilizing a subset from the “Poribohon-BD” dataset, established by Tabassum et al. [19]. A total of 240 images were allocated for training and 90 for validation. These images represented 8 vehicle categories: Bicycle, Bike, Bus, Car, CNG, Leguna, Rickshaw, and Truck. The selection aimed for a balanced class distribution to enhance model efficacy. Each image was labeled with the corresponding vehicle designation, and the dataset conformed to YOLOv8n specifications. The selected images were unlabeled initially, necessitating the use of CVAT for annotation and bounding box creation on the designated images. This organized preparation enabled effective training and validation, improving vehicle detection and classification accuracy in practical applications.

C. Model description

In the framework of this research, the YOLOv8n system was deployed for the recognition and organization of vehicles. The YOLOv8n variant, celebrated for its lightweight nature, is integrated into the YOLO object detection lineage. The "n" denotes "nano," reflecting its design for speed and efficiency in resource-constrained settings, particularly in real-time video analysis.

The YOLOv8n model was specifically trained to identify and categorize eight vehicle types: Bicycle, Bike, Bus, Car, CNG, Leguna, Rickshaw, and Truck. It processes images by dividing them into a grid, where each cell predicts bounding boxes and the likelihood of object presence. This integrated approach enables YOLOv8n to detect and classify objects concurrently at high velocity.

The diagram which is displayed in Fig. 1 represents a vehicle detection system employing a YOLOv8n model. Input images are resized to 1024x1024 pixels. These images are subsequently processed by the YOLOv8n model. After training the YOLOv8n model, it is employed for detecting objects in videos in real-time and recognizing vehicles of various types. The ultimate output is a real-time video displaying bounding boxes around detected vehicles along with their confidence levels.

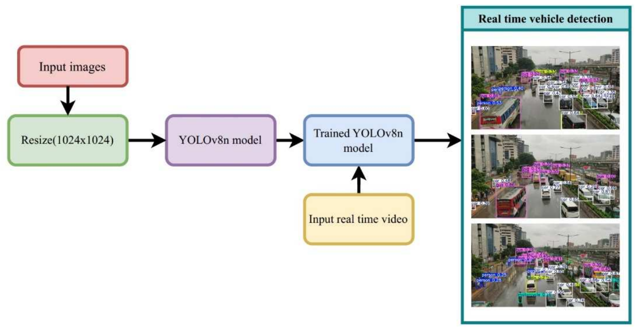

Fig. 1. YOLOv8n-Based Vehicle Detection System Flow Diagram.

In this investigation, the model was refined using a custom dataset based on the “Poribohon-BD” dataset. The training focused on optimizing model weights to enhance detection and classification accuracy for the targeted vehicle categories. YOLOv8n's capacity to generalize on unseen data was assessed through vehicle video footage, confirming its applicability in real-world scenarios such as traffic surveillance and vehicle enumeration. Despite its compactness, YOLOv8n effectively balanced speed and accuracy, although further enhancements are suggested to improve precision and recall.

D. Training

This work used a customized vehicle identification dataset with eight different classes to train the YOLOv8n model. Many vehicle types, including bicycles, bikes, buses, cars, CNG, legunas, rickshaws, and trucks, are included in the dataset. With a batch size of 32, the training procedure was carried out across 100 epochs to guarantee effective utilization of the GPU resources. An increased pixel count of 1024 × 1024 was employed to improve the model's precision in identifying tiny objects. To maximize model performance, 0.01 was chosen as the learning rate. On this particular dataset, pre-trained weights were used to refine the model, enabling it to make use of prior training experience on bigger datasets. To further enhance the model's generalization capacity, data augmentation techniques were also used to add changes to the training pictures. With the help of this method, the model was able to acquire reliable features for vehicle detection in a variety of classes. Fig. 2 displays a group of images from a batch of the train loader.

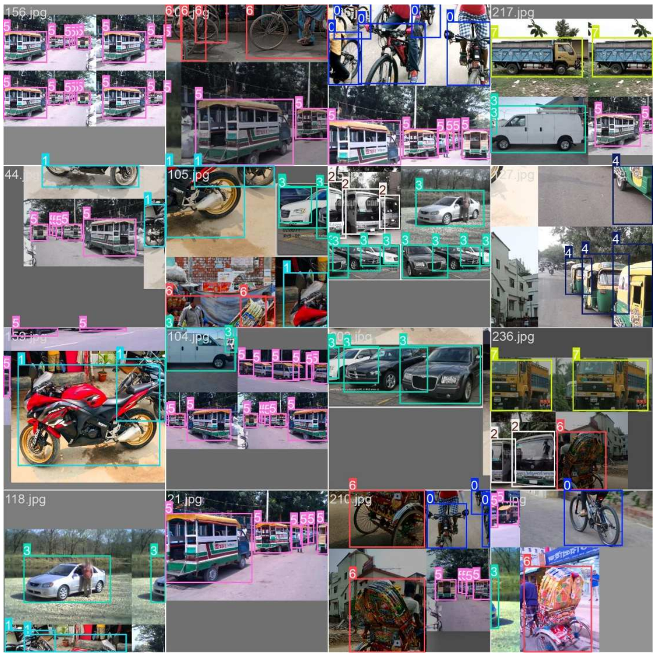

Fig. 2. Images from a batch of a train loader.

III. RESULTS AND DISCUSSION

Our study's findings show that although the suggested model was able to effectively identify and classify vehicles using real-time video data, it was not always very accurate. The model had difficulties, especially in intricate settings like Bangladesh's congested roadways.

A unique dataset was employed in this study for both training and validation. For validation, the dataset had 90 images in total. A batch of these images was displayed in Fig. 3 as part of the validation procedure once the model had been trained to evaluate the model's performance. We were able to visually examine the model's detection and classification performance using this method, which gave us valuable information about regions that needed more tuning.

The confusion matrix elucidates the YOLOv8n model's efficacy in vehicle type classification. It delineates correct and incorrect predictions across categories such as Bicycle, Bike, Bus, Car, CNG, Leguna, Rickshaw, Truck, and background. The diagonal entries signify accurately classified instances, with the model demonstrating commendable precision in identifying vehicles like Bicycle (65%) and Rickshaw (66%). Notably, confusion arises with the background class, where many vehicles, including Bike (74%) and Bus (61%), were erroneously classified as background. Furthermore, misclassifications among similar vehicle types, such as Car and CNG, reveal the model's challenges in distinguishing certain categories.

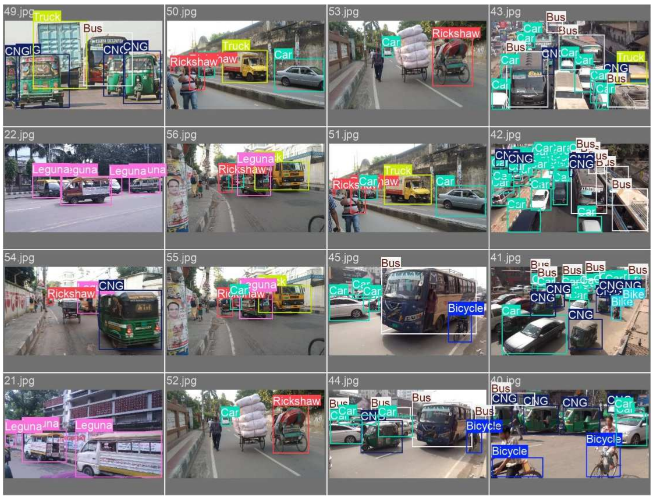

Fig. 3. Images from a batch of validation phase.

This confusion underscores the necessity for enhancements, potentially via data augmentation or model fine-tuning, to improve classification accuracy across all vehicle categories. The confusion matrix offers critical insights into the model's capabilities and limitations, especially in differentiating similar classes and minimizing background misclassifications. Fig. 4 shows the confusion matrix analysis of this study.

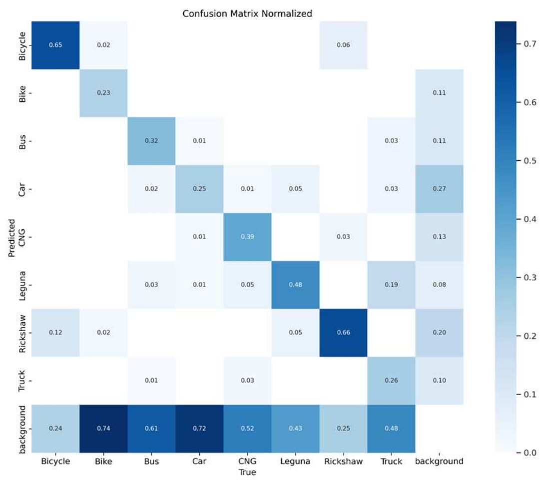

Fig. 4. Confusion matrix using YOLOv8n model.

We utilized a real-time video that was taken from a Bangladeshi road that the model had not seen during training to evaluate our model. After the trained system analyzed the footage, the model was able to identify and categorize different types of automobiles in actual driving situations. This illustrates the model's resilience in recognizing various vehicle kinds in hypothetical, real-world situations. The precise detection of automobiles as they appear in the video frames is demonstrated in Fig. 5.

Our study faced several limitations, primarily due to the imbalance in our custom dataset, where the number of images for different vehicle types was not uniform. This data imbalance contributed to lower accuracy during both the validation and testing phases, resulting in certain vehicle types not being properly recognized in the test video. In the future, we aim to address this by expanding and balancing our custom dataset, which will help improve the accuracy of our model. Additionally, we plan to enhance our work by incorporating lane-by-lane vehicle counting using real-time video from Bangladesh roads.

Fig. 5. The visualization of real-time vehicle detection.

In future endeavors, we plan to expand and normalize our tailored dataset to boost model accuracy. Additionally, we propose to improve our methodology by integrating lane-specific vehicle counting through real-time video data from Bangladeshi roadways, along with license plate text recognition. This integration will facilitate more accurate vehicle monitoring and categorization, thus bolstering sophisticated applications in traffic regulation, law enforcement, and vehicular safety.

IV. Conclusion

This comprehensive research endeavor successfully conceptualized, developed, and executed an advanced vehicle detection system that utilized the sophisticated YOLOv8n model, which was meticulously tailored to address the inherently dynamic and challenging traffic conditions that are characteristic of the roadways in Bangladesh. By employing cutting-edge real-time video analysis techniques, the model exhibited a remarkable ability to not only identify but also classify a diverse array of vehicle types, thereby significantly contributing to the ongoing efforts aimed at enhancing traffic control measures and ensuring the safety of road users. Despite these notable accomplishments, the system faced considerable challenges regarding its accuracy, primarily stemming from the pronounced imbalance present in the custom dataset utilized for training, as well as the intricate complexities associated with urban traffic environments that further compounded the issue. Future research initiatives will concentrate on a multifaceted approach that includes the augmentation of the dataset to achieve a more representative sample, the refinement of model precision to bolster its reliability, and the integration of advanced features such as lane-by-lane vehicle counting, all of which are intended to substantially elevate the system's practical applicability in the context of real-world smart city scenarios.

Acknowledgment

We acknowledge the support in conducting this research from the Department of Information and Communication Engineering (ICE) at Pabna University of Science and Technology (PUST).

References

[1] B. Siddique Mahi, F. S. Eshita and T. Helaly, "An Automated System for Wrong-Way Vehicle Detection using YOLO and DeepSORT," in 2023 5th International Conference on Sustainable Technologies for Industry 5.0 (STI), Dhaka, Bangladesh, pp. 1-6, 2023.

[2] Z. Rahman, A. M. Ami and M. A. Ullah, "A Real-Time Wrong-Way Vehicle Detection Based on YOLO and Centroid Tracking," *2020 IEEE Region 10 Symposium (TENSYMP)*, Dhaka, Bangladesh, pp. 916-920, 2020.

[3] S. I. C. Imperial, A. L. L. Lucas and M. V. Caya, "Vehicle Type Classification and Counting Using YOLOv4 Algorithm," in 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, pp. 1-6, 2022.

[4] Wang, M., Zhao, X., Chen, Z., & Chen, Q., “Vehicles Detection and Counting Method from Overhead View based on Embedded Vision,” *2023 IEEE 6th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, pp. 164-169, 2023.

[5] Murat Bakirci, "Enhancing vehicle detection in intelligent transportation systems via autonomous UAV platform and YOLOv8 integration, " Applied Soft Computing, vol. 164, 2024.

[6] Jingyu Liu, Jing Wu, Mingyu Liu, “UAV monitoring and forecasting model in intelligent traffic oriented applications,” Computer Communications, vol. 153, pp. 499-506, 2020.

[7] Karungaru S, Dongyang L, Terada K. "Vehicle detection and type classification based on CNN-SVM, " International Journal of Machine Learning and Computing, pp. 304-10, 2021.

[8] Dong, Z., Wu, Y., Pei, M., & Jia, Y., “Vehicle Type Classification Using a Semisupervised Convolutional Neural Network,” IEEE Transactions on Intelligent Transportation Systems, vol. 16, pp. 2247-2256, 2015.

[9] Song H, Liang H, Li H, Dai Z, Yun X. "Vision-based vehicle detection and counting system using deep learning in highway scenes, " European Transport Research Review, pp. 1-6, December 2019.

[10] M. A. Hossain, S. K. Ray, N. Islam, A. Alamin, and M. A. F. M. R. Hasan, "Enhanced human activity recognition through deep multi-layer perceptron on the UCI-HAR dataset," International Journal of Advances in Applied Sciences, vol. 13, pp. 429-438, 2024.

[11] Tsai, C., Tseng, C., Tang, H., & Guo, J. "Vehicle Detection and Classification based on Deep Neural Network for Intelligent Transportation Applications, " 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 1605-1608, 2018.

[12] S. U. Salekin, M. H. Ullah, A. A. A. Khan, M. S. Jalal, H. H. Nguyen, and D. M. Farid, "Bangladeshi Native Vehicle Classification Employing YOLOv8," in Intelligent Systems and Data Science. ISDS 2023, N. Thai-Nghe, T. N. Do, and P. Haddawy, Eds., Communications in Computer and Information Science, vol. 1949, Singapore: Springer, 2024.

[13] Z. Dong, Y. Wu, M. Pei and Y. Jia, "Vehicle Type Classification Using a Semisupervised Convolutional Neural Network," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 4, pp. 2247-2256, Aug. 2015.

[14] Ahmad Arinaldi, Jaka Arya Pradana, Arlan Arventa Gurusinga, “Detection and classification of vehicles for traffic video analytics,” Procedia Computer Science, vol. 144, pp. 259-268, 2018.

[15] Rahman, Z., Ami, A.M., & Ullah, M.A. “A Real-Time Wrong-Way Vehicle Detection Based on YOLO and Centroid Tracking,” 2020 IEEE Region 10 Symposium (TENSYMP), pp. 916-920, 2020.

[16] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 779-788, 2016.

[17] Sabina N, Aneesa M. P, Haseena P. V, “Object Detection using YOLO And Mobilenet SSD: A Comparative Study,” International Journal of Engineering Research & Technology (IJERT), vol. 11, June 2022.

[18] Palwankar, T., & Kothari, K. “Real Time Object Detection using SSD and MobileNet,” International Journal for Research in Applied Science and Engineering Technology, vol. 10, 2022.

[19] Tabassum, Shaira, Ullah, Md. Sabbir, Al-nur, Nakib Hossain, Shatabda, Swakkhar, “Poribohon-BD”, Mendeley Data, V2, 2020.