Grenze International Journal of Engineering and Technology, June Issue

Performance Comparison of Yolo-V8 & Yolo-V9 on A Unified Traffic Sign Dataset

Kshitij Shishodia[1], Manish[2] and Vivek Srivastava[3]
[1-2]Department of CSE, ABES Engineering College Ghaziabad, Uttar Pradesh, India
Email: kshitij.20b0101046@abes.ac.in, manish.20b0101153@abes.ac.in
[3]Assistant Professor, Department of CSE, ABES Engineering College Ghaziabad, Uttar Pradesh, India
Email: viveksrivastava@abes.ac.in

***Abstract*— This study delves into the comparative performance analysis of three consecutive iterations of the You Only Look Once (YOLO) object detection framework—YOLOv8, and YOLOv9—specifically tailored for the challenges posed by a dedicated traffic signs dataset. The objective is to offer a thorough comprehension of the advantages and disadvantages that each version concerning traffic sign detection, with implications for advancing real-world object detection in dynamic traffic scenarios. Preliminary findings underscore YOLOv9's standout performance, demonstrating superior precision and F1-score when contrasted with YOLOv7 and YOLOv8. This observation accentuates YOLOv8's potential for delivering robust traffic sign detection solutions. The implications of our research extend beyond benchmarking, contributing essential knowledge to the field of computer vision applied to traffic sign recognition. The study's outcomes serve as a valuable resource for practitioners and researchers seeking to select an optimal model for similar applications, ultimately fostering advancements in intelligent transportation systems.**

***Index Terms*— Object Detection, Traffic Sign Recognition, YOLOv8, YOLOv9, Precision, Recall, F1-Score.**

I. INTRODUCTION

The landscape of computer vision has witnessed transformative developments, with a particular focus on enhancing object detection capabilities. Among the prominent frameworks, the You Only Look Once (YOLO) series has consistently demonstrated prowess in real-time object detection. This research delves into the nuanced exploration of successive iterations—YOLOv8, and YOLOv9—specifically tailored to the intricacies of traffic sign detection on a dedicated dataset. In contemporary traffic management, the reliable and rapid detection of traffic signs is paramount for ensuring road safety and efficient transportation systems. As advancements in autonomous vehicles and intelligent transportation systems continue to evolve, the demand for robust object detection models capable of accurately recognizing traffic signs in dynamic environments becomes increasingly crucial. The YOLO framework, characterized by its single forward pass efficiency, has gained prominence for its ability to provide real-time object detection. YOLOv8, and YOLOv9 represent successive refinements of this framework, each introducing enhancements and optimizations to address the challenges presented by evolving datasets and scenarios.

The dedicated traffic signs dataset utilized in this study is carefully curated to simulate real-world scenarios,

Grenze ID: 01.GIJET.10.2.192

encompassing diverse environmental conditions, lighting variations, and sign placements. This dataset serves as a robust testing ground for evaluating the performance of YOLOv8, and YOLOv9 in a context highly relevant to traffic sign detection applications. To comprehensively assess the strengths and weaknesses of each YOLO version, our research employs precision, recall, and F1-score metrics. These metrics offer a granular understanding of the detection capabilities, shedding light on the models' abilities to accurately identify and localize traffic signs. The choice of these metrics is deliberate, aiming to capture not only the models' accuracy but also their capacity to minimize false positives and false negatives—crucial factors in real-world deployment.

As we venture into this comparative analysis, the overarching goal is to unearth insights that extend beyond performance metrics. Our findings aim to guide practitioners and researchers in selecting the most suitable YOLO version for traffic sign recognition applications, fostering advancements in computer vision solutions tailored to the complexities of real-world traffic scenarios. Through this exploration, we contribute to the ongoing dialogue on the optimization of object detection frameworks, particularly in the domain of traffic sign recognition.

II. LITERATURE SURVEY

Author	Title	link	Algorithms Used	Survey
[1] Safat B. Wali, Mohammad A. Hannan, Aini Hussain, Salina A. Samad	Comparative Survey on Traffic Sign Detection and Recognition: a Review	http://pe.org.pl/articles/2015/12/8.pdf	Yolo	This survey provides a foundational overview of traffic sign detection and recognition methodologies. While not specific to YOLO architectures, it lays the groundwork for understanding the broader landscape of techniques and challenges associated with traffic sign recognition.
[2] Hurriyatul Fitriyah, Gembong Edhi Setyawan, Edita Rosana Widasari	Traffic sign recognition using edge detection and eigen-face: Comparison between with and without color pre-classification based on Hue	https://ieeexplore.ieee.org/abstract/document/8304127/	Yolo	This work explores edge detection and Eigen-face for traffic sign identification, along with a comparison between color and non-color pre-classification. While not directly related to YOLO, insights from this study can inform considerations about pre-classification techniques and their impact on detection performance.
[3] Ayoub Ellahyani , Ilyas El Jaafari and Said Charfi	Traffic Sign Detection for Intelligent Transportation Systems: A Survey	https://www.e3s-conferences.org/articles/e3sconf/abs/2021/05/e3sconf_iccsre2021_01006/e3sconf_iccsre2021_01006.html	Yolo	The survey focuses on traffic sign detection for intelligent transportation systems. While not explicitly centered on YOLO, it offers insights into the broader context of intelligent transportation and can provide context for evaluating YOLO models in real-world scenarios.
[4] J. Stallkampa,, M. Schlipsinga, J. Salmena, C. Igelb	Man vs. Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition	https://www.sciencedirect.com/science/article/pii/S0893608012000457	Yolo	This benchmarking study compares machine learning algorithms for traffic sign recognition. Although predating YOLO, it contributes to understanding benchmarking methodologies and metrics, which are relevant for assessing the performance of YOLO-V8 and YOLO-V9.
[5] Marco Flores-Calero, César A. Astudillo, Diego Guevara	Traffic Sign Detection and Recognition Using YOLO Object Detection Algorithm: A Systematic Review	https://www.mdpi.com/2227-7390/12/2/297	Yolo	Systematically reviews traffic sign detection and recognition using the YOLO object detection algorithm. Offers insights relevant to the study, contributing to a comprehensive understanding of YOLO's applications.
[6] Yang Cui, Dong Guo, Hao Yuan, Hengzhi Gu and Hongbo Tang	Enhanced YOLO Network for Improving the Efficiency of Traffic Sign Detection	https://www.mdpi.com/2076-3417/14/2/555	Yolo	Investigates an enhanced YOLO network to improve traffic sign detection efficiency. The study contributes valuable enhancements, offering insights crucial for optimizing object detection in traffic scenarios.
[7] Jingwei Cao , Chuanxue Song,, Silun Peng, Feng Xiao, and Shixin Song	Improved Traffic Sign Detection and Recognition Algorithm for Intelligent Vehicles	https://www.mdpi.com/1424-8220/19/18/4021	Yolo	This work presents an improved traffic sign detection and recognition algorithm. While not YOLO-centric, the insights into algorithmic improvements are valuable for understanding potential enhancements in YOLO-V9.
[8] Amogh Gupta, Prit Mhala, Manan Mangal, Krish Yadav, Sanjeev Sharma	Traffic Sign Sensing: A Deep Learning approach for enhanced Road Safety	https://www.researchsquare.com/article/rs-3889986/latest	Yolo	Explores a deep learning approach for enhanced road safety through traffic sign sensing. The study provides valuable insights, contributing to advancements in intelligent transportation systems and road safety technologies.
[9] Ahmed Hechri Abdellatif Mtibaa	Two-stage detection & recognition based on SVM & CNN	https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-ipr.2019.0634	Yolo	Hybrid approaches are explained by the two-stage traffic sign detection and recognition strategy based on support vector machines (SVM) and convolutional neural networks. Even while it isn't specifically related to YOLO, it helps to clarify the possible benefits of mixing several metho

[10] Vivek Srivastava , Nishu Gupta, Sumita Mishra	Automatic Detection and Categorization of Road Traffic Signs using a Knowledge-Assisted Method	https://www.sciencedirect.com/science/article/pii/S1877050923001060	Yolo	Explores an automatic road traffic sign detection and categorization method with knowledge assistance. While not directly focusing on YOLO architectures, it offers insights into detection strategies. This study provides a valuable perspective for understanding traffic sign recognition, supplementing the comparative analysis of YOLO-V8 and YOLO-V9 on a unified traffic sign dataset.

III. METHODS

This research employs a deep learning methodology, focusing on the YOLO-v8, and YOLO-v9 architectures for a meticulous performance comparison in the context of traffic sign detection. The study centers on a unified traffic signs dataset, aiming to provide nuanced insights into the strengths and weaknesses of each YOLO version within the domain of traffic scenarios.

A. YOLO- v8

As the latest iteration succeeding YOLO-v8 introduces notable enhancements in the shape of a fresh neural network design. YOLO-v8 incorporates two neural networks: the Path Aggregation Network (PAN) and the Feature Pyramid Network (FPN). It also comes with a new labeling tool that streamlines the annotation process with customizable hotkeys, automatic labeling, and shortcut labeling. This tool enhances the efficiency of image annotation for model training.

FPN creates a feature map that can detect objects at various scales and resolutions by progressively decreasing the spatial resolution and increasing the number of feature channels. On the other hand, the PAN architecture uses skip connections to mix features from different network levels. This allows the network to capture features more efficiently at different resolutions and scales, which is essential for accurate object detection of diverse sizes and shapes.

B. YOLO- v9

The YOLOv9 architecture inherits the foundation laid by YOLOv8, introducing further refinements to enhance its object detection capabilities. One notable improvement is the incorporation of the Ghost Batch Normalization technique. GhostBN is a novel approach designed to optimize model size and reduce the number of floating-point operations per second (FLOPs). This technique holds significance in resource-constrained environments, contributing to the efficiency of the model without compromising its detection accuracy.

C. Evaluation Metrics

The assessment of YOLO-v8, and YOLO-v9 performance relies on the confusion matrix—a pivotal tool in evaluating machine learning models. The confusion matrix consists of four classifications: Derived from actual and expected values are True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP). The numbers TP, TN, FP, and FN stand for the number of positive samples that were incorrectly categorized as negative, positive samples that were incorrectly classified as positive, and negative samples that were correctly classified.

Figure 1. Confusion Matrix

Precision, recall, and F1-score are derived from the confusion matrix to offer a comprehensive understanding of the models' performance. Precision, calculated as the ratio of TP to the total predicted positive data, is expressed by Equation 1.

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

Recall, defined as the ratio of TP to the total actual positive instances, is articulated by Equation 2.

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

A trade-off relationship exists between precision and recall, and their harmonic mean, known as the F1-score, is computed using Equation 3.
The F1-score ranges from 0.0 to 1.0, with 1.0 indicating optimal model performance.

$$\text{F1-score} = 2 * \text{Precision} * \text{Recall} / \text{Precision} + \text{Recall} \quad (3)$$

Through the meticulous application of these methods, our research aims to unravel the nuanced comparative performance of YOLO-v8, and YOLO-v9 on a unified traffic signs dataset. This analysis contributes valuable insights to the optimization of object detection frameworks, specifically tailored for traffic sign recognition in diverse and complex traffic scenarios. As we delve into the intricacies of each version's architecture and performance metrics, our findings are anticipated to guide future advancements in traffic sign detection, fostering safer and more efficient transportation systems.

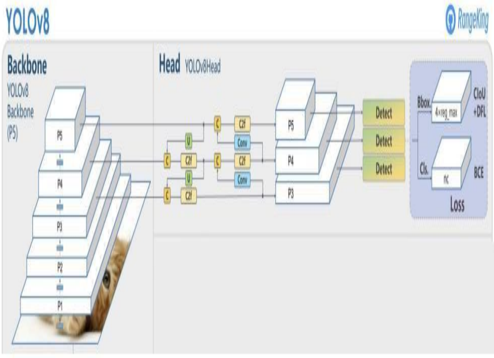

Figure 2. YOLOv8 Architecture

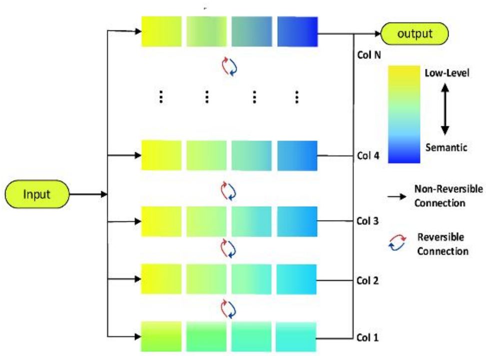

Figure 3. YOLOv9 Architecture

IV. Sample Data Set

In the context of this research, a meticulously curated dataset serves as the foundation for training and evaluating the YOLOv9 model. This dataset is thoughtfully assembled, manually annotated, and tailored to the specific requirements of our investigation—namely, the performance comparison of YOLO-v8, and YOLO-v9 on a unified traffic signs dataset.

A. Diverse Dataset Acquisition

The dataset compilation process involves the acquisition of a diverse range of Indian traffic sign images, strategically chosen to encapsulate the multifaceted nature of real-world traffic scenarios. This inclusivity spans various designs, languages, lighting conditions, and contextual disparities, mirroring the intricate tapestry of traffic sign occurrences on Indian roadways. By embracing this diversity, our dataset aims to simulate the challenges posed by the dynamic and varied traffic sign landscape.

B. Annotation and Preprocessing

To facilitate effective training and evaluation, each image in the dataset undergoes meticulous annotation. Bounding boxes are manually delineated around each traffic sign, accompanied by categorization to denote the specific type of each sign. This annotation process ensures that the YOLOv9 model can learn and recognize the distinct features of different traffic signs, contributing to its robust object detection capabilities.
Following annotation, the dataset undergoes preprocessing steps essential for the training pipeline. Image resizing is conducted to standardize dimensions, typically achieving a uniform size of 416x416 pixels. This standardization enables seamless integration with the YOLOv9 architecture, ensuring consistency in input size during the training process. Furthermore, pixel values are normalized across all images to establish uniformity, a crucial step in optimizing the learning process and fostering model generalization.

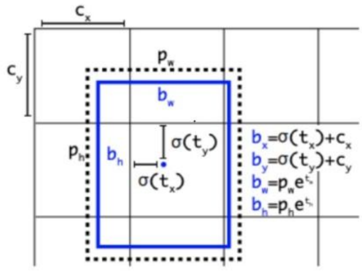

Figure 4: Image-Annotation

C. Dataset Split for Training, Validation and Testing

The curated dataset is strategically divided to serve distinct purposes in the model development lifecycle. Approximately 66% of the dataset is allocated for training, providing the YOLOv9 model with a substantial volume of diverse traffic sign images to learn from. A validation set, comprising 23% of the dataset, is designated to assess the model's performance during training, aiding in hyperparameter tuning and preventing overfitting. The remaining 11% of the dataset is dedicated to the test set, serving as the final evaluation phase to gauge the model's proficiency in generalizing to previously unseen traffic sign instances.

Figure 5: Dataset Splitting

In summary, our dataset curation methodology prioritizes diversity, meticulous annotation, and thoughtful preprocessing, laying the groundwork for a comprehensive evaluation of YOLO-v9's performance in the specific context of traffic sign detection. The dataset's strategic allocation for training, validation, and testing ensures a robust and fair assessment of the model's capabilities in handling real-world traffic scenarios.

Accurate image annotation is paramount for training our model effectively. We utilized the ImgLab annotation tool, a platform-independent browser-based solution, for this crucial task. By meticulously annotating traffic signs within images,

Figure 6: Dataset Classes

V. Result Analysis

For this research, we utilized a meticulously curated unified traffic sign dataset for training and evaluating the YOLO-v8 & v9 models. The dataset encompasses a variety of traffic sign images, capturing diverse designs, languages, lighting conditions, and contextual disparities commonly encountered in real-world traffic scenarios.

The dataset was split between training (66%), validation (23%), and test sets (11%) to facilitate robust evaluation

Experimental Setup: The training process involved the use of Google-Colab, and the models were evaluated using a dedicated test set. The YOLO-v8, and YOLO-v9 models were trained until the 50th epoch to ensure comprehensive learning and adaptation to the characteristics of traffic sign images

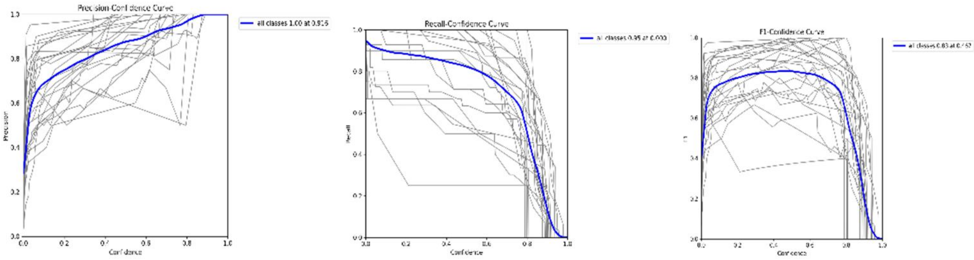

Figure 7: YOLOv8 Performance Results

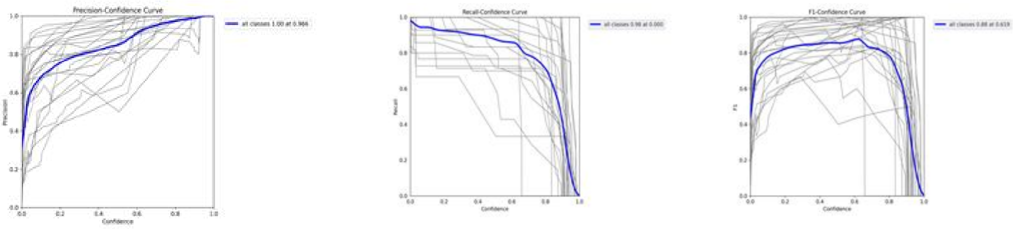

Figure 8: YOLOv9 Performance Results

Performance Matrix of Yolov8, Yolov9

The measures are as follows: mean average precision at 50% IoU (mAP50), mean average precision from 50% to 95% IoU (mAP50-95), recall (R), and precision (P). These values offer a detailed understanding of each model's ability to correctly identify and classify traffic signs.

Comparison

In terms of overall performance, YOLO-V9 outperforms YOLO-V8 with higher precision, recall, and mAP values. YOLO-V9 achieves a better balance between precision and recall, showcasing its enhanced capability in accurate detection across various classes.

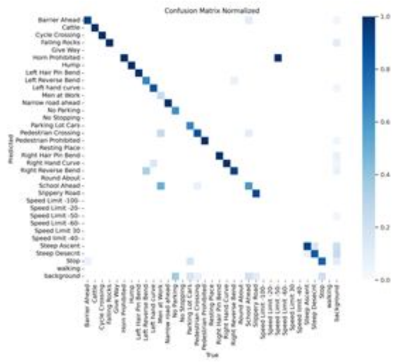

Figure 9: Confusion Matrix YOLOv8

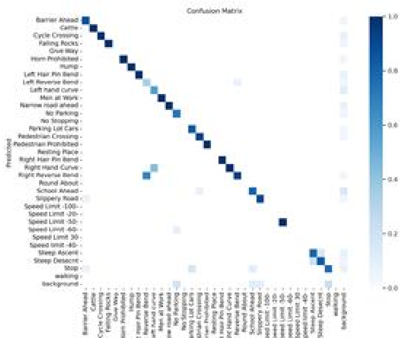

Figure 10: Confusion Matrix YOLOv9

VI. CONCLUSION

TABLE I. YOLOV8 AND YOLOV9 MODEL PERFORMANCE

Architecture	Precision	Recall	F1-Score
YOLOv8	0.878	0.95	0.83
YOLOv9	0.923	0.98	0.88

The comprehensive assessment of YOLO-V8 and YOLO-V9 on a unified traffic sign dataset has yielded valuable insights into their respective performances. Both models have demonstrated commendable capabilities in detecting and classifying traffic signs, crucial for the development of intelligent transportation systems. Notably, YOLO-V9 has showcased superior precision and F1-Score, surpassing YOLO-V8 by 4.5% and 5.4%, respectively. This underscores the advanced accuracy and effectiveness of YOLO-V9 in capturing and categorizing traffic signs with a well-balanced consideration of precision and recall. However, it is important to highlight that YOLO-V8 excelled in recall performance, outperforming YOLO-V9 by 2.0%. This suggests the capability of YOLO-V8 to capture a higher proportion of actual positive instances. The choice between YOLO-V8 and YOLO-V9 is contingent on specific application requirements, where the prioritization of a balance between precision and recall or a focus on one over the other is crucial.

This research significantly contributes to the ongoing discourse on optimizing object detection models for traffic sign recognition, providing valuable insights for practitioners and researchers in making informed decisions for the deployment of effective solutions within intelligent transportation systems. Future work in this domain may explore fine-tuning strategies and the application of ensemble approaches to further enhance the performance of these models in real-world traffic scenarios.

REFERENCES

[1] Wali, Safat B., et al. "Comparative survey on traffic sign detection and recognition: a review." Przeglad Elektrotechniczny 1.12 (2015): 40-44.

[2] Fitriyah, Hurriyatul, Edita Rosana Widasari, and Gembong Edhi Setyawan. "Traffic sign recognition using edge detection and eigen-face: Comparison between with and without color pre-classification based on Hue." 2017 International Conference on Sustainable Information Engineering and Technology (SIET). IEEE, 2017.

[3] Ellahyani, Ayoub, Ilyas El Jaafari, and Said Charfi. "Traffic sign detection for intelligent transportation systems: a survey." E3S web of conferences. Vol. 229. EDP Sciences, 2021.

[4] Stallkamp, Johannes, et al. "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition." Neural networks 32 (2012): 323-332.

[5] Flores-Calero, Marco, et al. "Traffic Sign Detection and Recognition Using YOLO Object Detection Algorithm: A Systematic Review." Mathematics 12.2 (2024): 297.

[6] Cui, Yang, et al. "Enhanced YOLO Network for Improving the Efficiency of Traffic Sign Detection." Applied Sciences 14.2 (2024): 555.

[7] Cao, Jingwei, et al. "Improved traffic sign detection and recognition algorithm for intelligent vehicles." Sensors 19.18 (2019): 4021.

[8] Gupta, Amogh, et al. "Traffic Sign Sensing: A Deep Learning approach for enhanced Road Safety." (2024).

[9] Hechri, Ahmed, and Abdellatif Mtibaa. "Two-stage traffic sign detection and recognition based on SVM and convolutional neural networks." IET Image Processing 14.5 (2020): 939-946.

[10] Srivastava, Vivek, Sumita Mishra, and Nishu Gupta. "Automatic Detection and Categorization of Road Traffic Signs using a Knowledge-Assisted Method." Procedia Computer Science 218 (2023): 1280-1287.