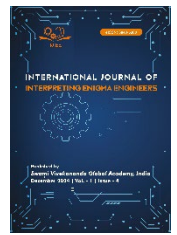




HELMET AND NUMBER PLATE DETECTION USING YOLOV5 AND YOLOV8



Tanusha Gorak¹, Dr.C.Kishor Kumar Reddy¹

Original Article

¹Department of Computer Science and Engineering, Stanley College of Engineering and Technology for Women.

*Corresponding Author's Email: tanushaa15@gmail.com, ckishorkumar@stanley.edu.in

Abstract

The purpose of this project is to create a reliable and effective system for detecting vehicle number plates using two well-known versions of the object detection algorithm YOLO (You Only Look Once), specifically YOLOv5 and YOLOv8. This system aims to automatically identify and extract number plates from images of vehicles, which is essential for applications such as automated tolling, traffic monitoring, and law enforcement. The project involves training both YOLOv5 and YOLOv8 models on a specially created dataset of labelled vehicle images, and then evaluating and comparing the models based on detection accuracy, inference speed, and their robustness in various environmental scenarios. Additionally, the system's performance is improved by incorporating Optical Character Recognition (OCR) to retrieve text from the identified number plates. The results of this comparison indicate that YOLOv8, thanks to its enhanced architecture, offers greater accuracy and faster inference rates compared to YOLOv5, making it a better choice for real-time applications. This work emphasizes the efficacy of deep learning methods in addressing number plate recognition challenges, providing a scalable and practical solution for intelligent transportation systems.

Keywords: *YOLO (You Look Only Once), YOLOv5, YOLOv8, Deep Learning.*

INTRODUCTION

Vehicle number plate detection is critical in modern transportation systems, with applications in areas such as automated tolling, parking management, traffic monitoring, and law enforcement [1]. Traditional methods of number plate recognition have been limited in accuracy and scalability, especially when dealing with various vehicle types, environmental conditions, and real-time processing demands [2]. However, with the improvements in computer vision and Deep learning, these challenges can now be addressed more effectively. In recent years, the object detection algorithm YOLO (You Only Look Once) has emerged as one of the most popular and efficient solutions for tasks like real-time object detection. YOLO is capable of simultaneously predicting both the class and bounding box coordinates for multiple objects within an image, making it particularly well-suited for tasks like vehicle number plate detection. Among various versions of YOLO, YOLOv5 and YOLOv8 have gained significant attention due to their impressive performance and ease of implementation. YOLOv5, although already highly efficient, has been surpassed by the newer YOLOv8 in terms of accuracy, speed, and model optimization [3]. YOLOv8 introduces several architectural improvements and optimizations that enhance detection performance and inference speed, by an ideal candidate for real-time applications where processing speed is important.

This project aims to implement and comparison of the performance of YOLOv5 and YOLOv8 in detecting vehicle number plates from images. By training both models on a custom dataset of vehicle images with labelled number plates, we will evaluate their accuracy, inference speed, and robustness in various environmental conditions [4]. Additionally, Optical Character Recognition (OCR) will be integrated to extract the textual information from detected number plates, further improving the system's functionality.

The primary goal of this research is to provide a comprehensive comparison between YOLOv5 and YOLOv8 in the context of number plate detection, thereby contributing valuable insights into the effectiveness of different YOLO architectures for real-world applications in intelligent transportation systems.

Recognizing vehicle number plates is a vital part of today's intelligent transportation systems (ITS) and is essential for a variety of purposes, such as automatic toll collection, parking management, traffic monitoring, and surveillance. Accurately identifying number plates in real time can really make traffic management smoother, boost law enforcement efforts, and enhance safety for everyone on the road [5]. Still, this task can be quite tricky because of the many different vehicle types, varying lighting conditions, environmental elements like weather and traffic congestion, and the unique formats of number plates found around the globe [6].

Historically, conventional techniques for recognizing vehicle number plates depended on image processing methods such as edge detection, contour analysis, and template matching. Although these methods were successful in controlled settings, they often faced challenges due to variations in images (for instance, plate orientation, occlusion, and distortion) and required considerable manual feature extraction and parameter tuning [7],[14]. Consequently, the necessity for more sophisticated, automated, and scalable solutions became apparent.

In recent years, object detection models based on deep learning, especially the YOLO (You Only Look Once) series of algorithms, have transformed the landscape of image analysis and recognition [8],[13]. YOLO is a robust object detection model that can simultaneously detect and locate several objects in a single forward pass, providing a notable advantage over traditional methods in terms of both accuracy and speed [9]. The architecture of YOLO facilitates real-time processing, which is essential for applications like vehicle number plate recognition, where images or video frames need to be analyzed rapidly to enable prompt decision-making.

The initial YOLO algorithm has evolved through numerous versions, each bringing enhancements in both accuracy of detection and efficiency [10],[12]. YOLOv5, a well-regarded and commonly utilized version, has established itself as a benchmark in the object detection field due to its strength, flexibility, and simplicity of implementation [11],[15]. YOLOv5 has been applied to an array of scenarios, such as identifying faces, pedestrians, vehicles, and license plates. Nonetheless, YOLOv8, the most recent version in the YOLO series, further refines the network for improved detection accuracy, enhanced generalization, and quicker inference times. These advancements render YOLOv8 especially attractive for real-time number plate detection in vehicles, where the capacity to swiftly process frames is crucial for achieving success.

LITERATURE SURVEY

Yixiao Zhang et al.2024. Detecting and locating vehicles is a crucial task in automatic driving systems. Conventional detection techniques are often vulnerable to variations in lighting, obstructions, and changes in scale within complex environments, which hinders detection precision and reliability. To address these challenges, this paper introduces a vehicle detection and location approach utilizing YOLOv5 (You Only Look Once version 5) and binocular vision. Binocular vision employs two cameras to capture images from different perspectives simultaneously. By analyzing the discrepancies between the two images, it is possible to obtain more accurate depth information [1].

Sunil Kumar et al, 2023.In recent times, progress in sustainable intelligent transportation has highlighted the importance of vehicle detection and tracking for managing real-time traffic flow on highways. Nonetheless, the performance of current deep learning-based methods continues to face significant challenges due to variations in

vehicle sizes, occlusions, and other real-time traffic situations. To tackle the issues of vehicle detection and tracking, a smart and efficient approach is introduced, utilizing You Only Look Once (YOLOv5) for vehicle detection at a speed of 140 FPS, and subsequently integrating Deep Simple Online and Realtime Tracking (Deep SORT) into the detection outcomes to monitor and forecast vehicle positions [2].

Wojciech Lindenheim-Locher et al, 2023. This study concentrates on the initial phase of tracking 3D drone challenge, specifically accurate identification of the drones in the images captured through a multi-camera system. The YOLOv5 DL model, utilizing various input resolutions, is trained and evaluated using real, multimodal data that consists of synchronized sequences of videos alongside motion capture data as a reliable truth reference. The boxes are established based on 3D location and orientation of the drone an asymmetric cross affixed to top of drone, with a distance known to the center of drone. The markers registered by motion capture acquisition are used to recognize the arms of the cross. In addition to the traditional (mAP) mean average precision, a suitable measure for assessing detection effectiveness in tracking of 3D is introduced, which is the distance between the centroids of detected drones and corresponding references, factoring in rates of both false positives and false negatives. Furthermore, the videos created in the Air Sim simulation environment were utilized in both the training and testing phases [3].

Qing An et al, 2023. Safety helmets are crucial in various work environments, both indoors and outdoors, such as in high-temperature metallurgical operations and in the construction of tall buildings, to prevent injuries and promote safe production practices. However, relying on manual oversight can be expensive and often lacks consistent enforcement, influenced by various human factors. Additionally, detecting small objects can be challenging and often lacks accuracy. Enhancing safety helmets through an improved detection algorithm for helmet could solve the problems and represents a favourable strategy. In the research, it is introduced a version of the YOLOv5s network which is a modified one, the lightweight model based on Deep Learning for recognition of object. The suggested model builds upon the YOLOv5s framework, enhancing its effectiveness by recalibrating the predicted frames, employing metric for clustering the Intersection over Union (IoU), and adjusting the frames using of Kmeans++ algorithm. To boost the backbone and neck networks of the YOLOv5s architecture, the convolutional block attention module (CBAM) and the global attention mechanism (GAM) have been incorporated. These attention modules improve feature extraction capabilities in deep learning neural networks by decreasing the loss of information features and augmenting the global interactions representations. Additionally, the CBAM model is involved into the CSP module to refine target feature extraction while decreasing computational demands for model execution [4].

Chenyang Wei et al, 2023. The integrated fast detection technology for electric motorcycles, riders, helmets and numberplates plays a crucial role in enhancing safety in the traffic.

YOLOv5 represents the most sophisticated single-state object detection algorithms available.

However, deploying it on the embedded systems like Unmanned Aerial Vehicles (UAV) proves challenging due to the high demands and memory required. This paper introduces a version of YOLOv5 designed for the rapid detection of electric bike helmets and license plates, by implementing two strategies to enhance the original YOLOv5 [5].

Ju Han et al,2023. The detection of helmets for safety in the construction sector necessitates high resolution image transmission, which poses challenges for current image detecting techniques to gain rapid detection. Addressing this issue, a new Super-Resolution (SR) reconstruction module has been developed for enhancing the image resolution prior to the detection phase. Within this reconstruction module, a multichannel attention mechanism is incorporated to broaden the feature extraction capabilities. Additionally, a novel Cross Stage Partial module based on You Only Look Once v5 has been introduced for minimizing the reduce gradient confusion and information loss. Experiment have been conducted for assessing effectiveness of the proposed algorithm[6].

Yiduo Zhang et al, 2023. In the smart monitoring of construction sites, detecting safety helmets is critically important. Nevertheless, due to the small size of helmets and the significant noise levels typical in construction settings, current detection techniques frequently face challenges with inadequate accuracy and robustness. To tackle this issue, this

paper presents a new algorithm for safety helmet detection, named FEFD-YOLOV5. The FEFD-YOLOV5 algorithm improves detection effectiveness by integrating a shallow detection head tailored for identifying small targets and utilizing an SENet channel attention module to condense global spatial data, thereby enhancing the model's mean average precision (mAP) in relevant contexts. Furthermore, this algorithm introduces an innovative denoise module, guaranteeing that the model preserves high accuracy and robustness amid various noise conditions, thus boosting its generalization ability to satisfy the requirements of real-world applications. Experimental findings indicate that the proposed enhanced algorithm achieves a detection accuracy of 94.89% in environments without noise and still maintains 91.55% accuracy in highly noisy conditions, showcasing superior detection performance compared to the original algorithm [7].

Weipeng Tai et al, 2023. Helmet recognition algorithms that utilize deep learning are designed to facilitate continuous detection and documentation of violations, such as not wearing a helmet. Nevertheless, real-world situations can present challenges due to weather and human influences, complicating the process of detecting safety helmets. Issues like camera shake and head occlusion frequently arise, resulting in inaccurate outcomes and reduced availability. To tackle these real-world challenges, this study introduces a new algorithm named DAAMYolov5. The DAAM-Yolov5 algorithm enhances dataset under various conditions to boost the mean Average Precision (mAP) in relevant scenarios through the implementation of Mosaic9 data augmentation[8].

Shuai Chen et al, 2022. A new sensing detection method that combines model super resolution reconstruction, a transformer-spatial attention mechanism, and the YOLOv5 image classifier is presented in response to the current difficulty of detecting helmet use among riders through aerial photography with unmanned aerial vehicles. The small size of targets, high size fluctuation, motion blur in aerial photos is some of the reasons why the detection model has low accuracy and poor generalization. A ladder-type multi-attention network is created for target recognition in order to overcome these difficulties. It decreases information loss, fully captures visual features, and enables information integration and sharing at all levels [9].

Jianfeng Han et al, 2024. Ensuring that workers wear helmets at construction sites is crucial for preventing accidents, making supervision necessary. This demand leads to a strong requirement for run-time performance. We optimized the architecture based on Yolov7. For boosting real-time capabilities, we incorporated Ghost Module after assessing different modules, resulting in a more efficient design that produces additional feature mapping with few linear operations. After evaluating various attention mechanisms, we added SE blocks to emphasize information in the image. The improved run-time performance while preserving the high accuracy fulfilling the detection requirements [10].

Ahatsham Hayat et al, 2022. For the construction business, worker safety on building sites is becoming a more significant concern. Although there are a number of factors that contribute to incorrect use, properly worn safety helmets can help reduce worker injuries at these sites. Thus, it's critical to put in place a computer vision-based automatic safety helmet detecting system. The detection of helmets in construction contexts has received little attention, despite the fact that numerous researchers have developed Machine Learning and Deep Learning helmet recognition algorithms. This research presents a You Only Look Once (YOLO) architecture-based real-time automatic safety helmet identification system for construction sites. For real-time helmet detection, YOLO provides a high-speed solution with a processing rate of 45 frames per second. [11].

Maged Shoman et al, 2024. In this study, we propose a new method for detecting helmet usage violations among motorcycle riders by integrating Deep Convolutional Generative Adversarial Networks (DCGANs) with the YOLOv8 object detection algorithm. This study aims to increase the accuracy of existing helmet violation detection techniques, which are prone to errors and frequently rely on manual checks. The suggested method involves using a large dataset of both synthetic and actual photos to train the model, which has excellent accuracy in identifying helmet infractions, even when there are several riders. In order to improve model performance in real-world scenarios, we use data augmentation in conjunction with artificial images produced by DCGANs to improve the training dataset. We pay special attention to resolving imbalanced classes. [12].

PROPOSED METHODOLOGY

Helmet and Numberplate Detection:

In video footage, helmets and license plates are detected. Using a method known as object detection, it can first identify objects like helmets and license plates inside video frames. Once the helmet has been identified, the algorithm proceeds to determine whether it is being worn—a process called helmet classification. To improve visual comprehension, it uses boxes to highlight the objects that have been identified and shows whether or not a helmet is being worn. Importantly, it works in real time, processing video frames as they are received and optimizing them for dynamic scenarios. By displaying the processed video and pausing when the user presses the "Esc" key, the application facilitates user involvement. Additionally, it ensures effective performance during video processing by deftly managing computational resources. In summary, the algorithm can identify helmets and license plates in films by detecting, classifying, and displaying real-time information thanks to these traits taken together.

This study advances using a methodical approach to develop a trustworthy system for identifying license plates and helmets. After a thorough data collection process, a diverse dataset with annotated photographs that depict actual circumstances is put together. An effective deep learning model for object detection is selected after data preprocessing, which includes normalization, resizing and augmentation. This procedure prioritizes a balance between accuracy and computational efficiency.

In order to fine-tune the chosen pre-trained model using the labeled dataset and adjust its capabilities to the particular task needs, transfer learning is essential. To avoid overfitting, the training method involves dividing the dataset carefully, adjusting the hyperparameters, and continuously validating the results. The model is improved using post-processing techniques to improve detection accuracy as well as assessment measures including precision, recall, and F1 score. The seamless integration of the trained model into real-time systems, with an emphasis on speed and efficiency optimization, is at the heart of the study. The model is thoroughly tested and validated to guarantee its adaptability to a range of settings and circumstances. Road safety is improved and intelligent transportation systems are advanced through the process of iterative fine-tuning and comprehensive documentation, which guarantees the established system's reproducibility and scalability. Use the labelled dataset to optimize the whole network, including the detecting head, Region Proposal Network (RPN), and backbone. To benefit from the pre-trained weights created from the picture classification challenge, use transfer learning. Use a network that is completely connected as the object detecting head. This network fine-tunes the bounding box coordinates and classifies the suggested regions into predefined types (helmet, number plate).

YOLOv5 and YOLOv8:

YOLOv5 and YOLOv8 represent two versions of the You Only Look Once (YOLO) object detection framework, both providing high-performance capabilities for real-time object detection, although they differ in their architecture, performance, and adaptability. YOLOv5, created by Ultralytics and built using PyTorch, has achieved significant popularity thanks to its user-friendly interface, rapid inference speed, and superb accuracy. It allows for a straightforward configuration for detecting objects in images, videos, or live webcam feeds. YOLOv5 comes in various model sizes, including YOLOv5small, YOLOv5medium, YOLOv5large, and YOLOv5extra-large, enabling users to select an appropriate model based on their computing resources and the balance between speed and accuracy they wish to achieve.

The YOLOv5 model provides users with the capability to train tailored models using their own datasets, making it applicable to a broad range of scenarios beyond the conventional pretrained models. Its customization flexibility and access to pretrained weights (like those trained on COCO or specialized datasets) have established YOLOv5 as a preferred choice for numerous real-time detection applications, particularly when optimizing computational efficiency is essential. The PyTorch implementation of YOLOv5 allows developers to swiftly adjust and expand the model according to specific requirements, rendering it a versatile option for both novices and experienced users.

Conversely, YOLOv8 represents the latest advancement in the YOLO series, coming with a range of enhancements compared to its earlier versions, including YOLOv5. While YOLOv5 has established a benchmark for real-time detection, YOLOv8 features an improved architecture, advanced optimization strategies, and a more refined framework that greatly boosts the model’s efficiency in both speed and precision. YOLOv8 is included in the ultralytics Python package, which not only provides detection functionality but also offers tools for model training, assessment, and exportation. YOLOv8 has been designed to be more modular than YOLOv5, giving users greater adaptability for customizing their training and exporting the model in various formats like ONNX and TensorFlow, facilitating deployment across different environments.

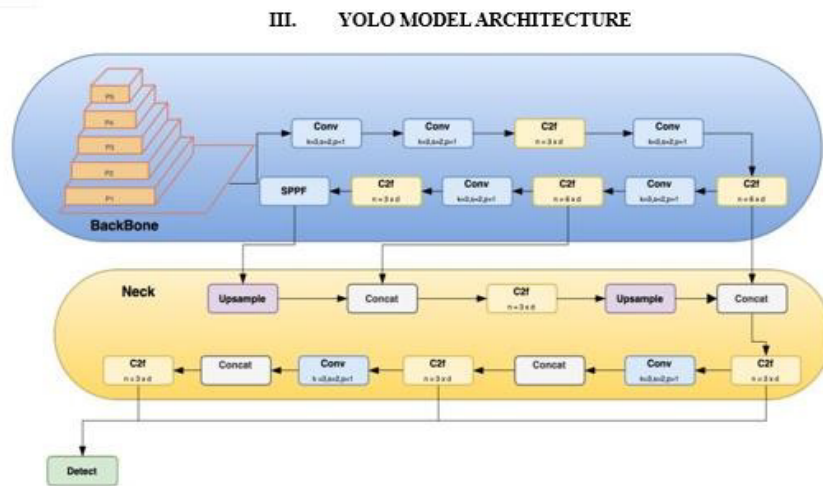


Figure 3: YOLOv8 Model Architecture [6]

The YOLO series has seen several successors, including YOLOv4, YOLOv5, YOLOv6, and YOLOv7 [6]. In 2023, Ultralytics released YOLOv8, depicted in Figure 3. YOLOv8 introduces the C2f module, based on Cross Stage Partial (CSP), in its backbone architecture, different from YOLOv5’s C3 Module. The CSP architecture enhances the model’s learning capacity while reducing computational complexity. The C2f module [6] comprises two Conv Modules and n BottleNeck, connected through Split and Concat operations. The rest of the backbone architecture remains consistent with YOLOv5. Additionally, the SPPF Module is used at the final layer [6] of the backbone in YOLOv8.

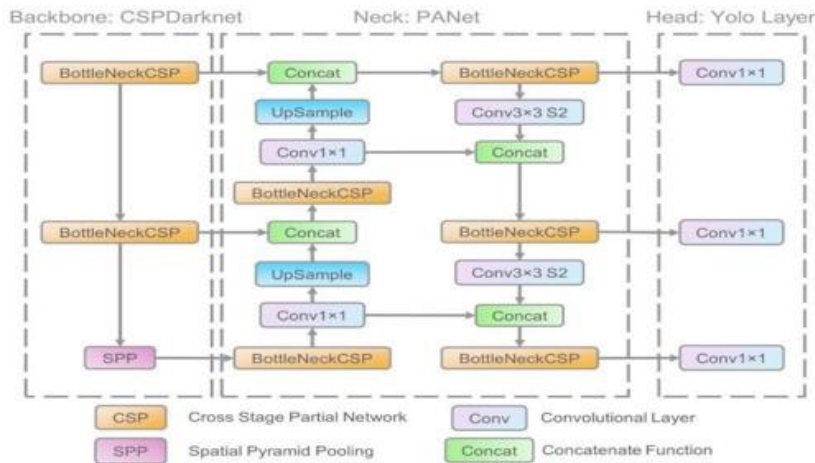


Figure 4: YOLOv5 Model Architecture [7]

The main enhancements in YOLOv8 compared to YOLOv5 consist of increased model efficiency, enhanced feature extraction layers, and upgrades in the backbone and neck networks, resulting in quicker inference and improved performance across various hardware. YOLOv8 also offers better capabilities for training with limited data, making it more effective for situations where extensive annotated datasets are unavailable. Additionally, YOLOv8 includes extra training options and model evaluation metrics, facilitating easier monitoring and refinement of the training process for users. The model is capable of real-time object detection on videos, images, and live webcam feeds, and it can manage more intricate scenarios because of its superior precision and recall.

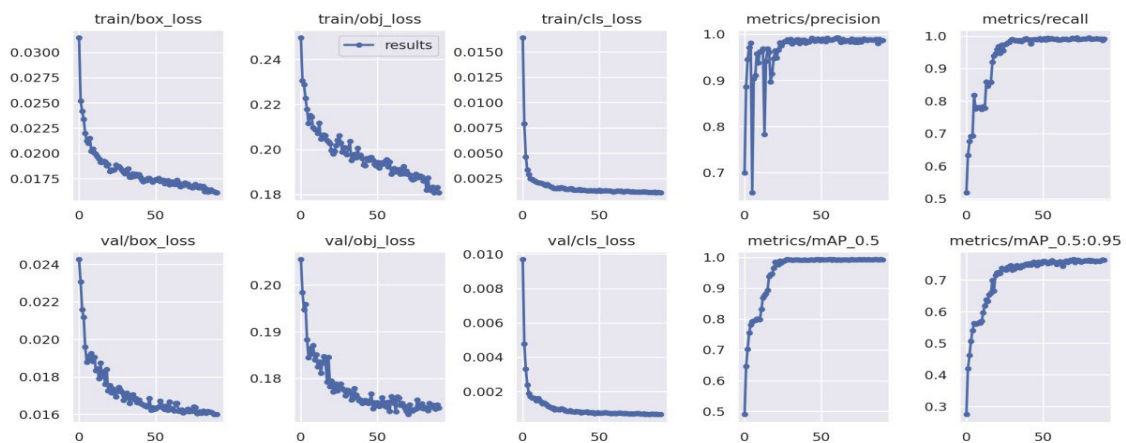
Even with the progress made in YOLOv8, YOLOv5 still remains very popular due to its established ecosystem, robust community backing, and user-friendliness. Numerous developers continue to rely on YOLOv5 owing to its reliability and the fact that it has been extensively validated in numerous fields, from security monitoring to self-driving vehicles. Both models are engineered for speed and accuracy, but YOLOv8 enhances performance with its architectural advancements and optimizations. For brand new projects that demand high precision and quicker inference times, YOLOv8 could be the more favorable choice. Nevertheless, for those already accustomed to YOLOv5 or needing a reliable solution with comprehensive documentation and tutorials, YOLOv5 is still a fantastic option.

RESULTS AND DISCUSSIONS:

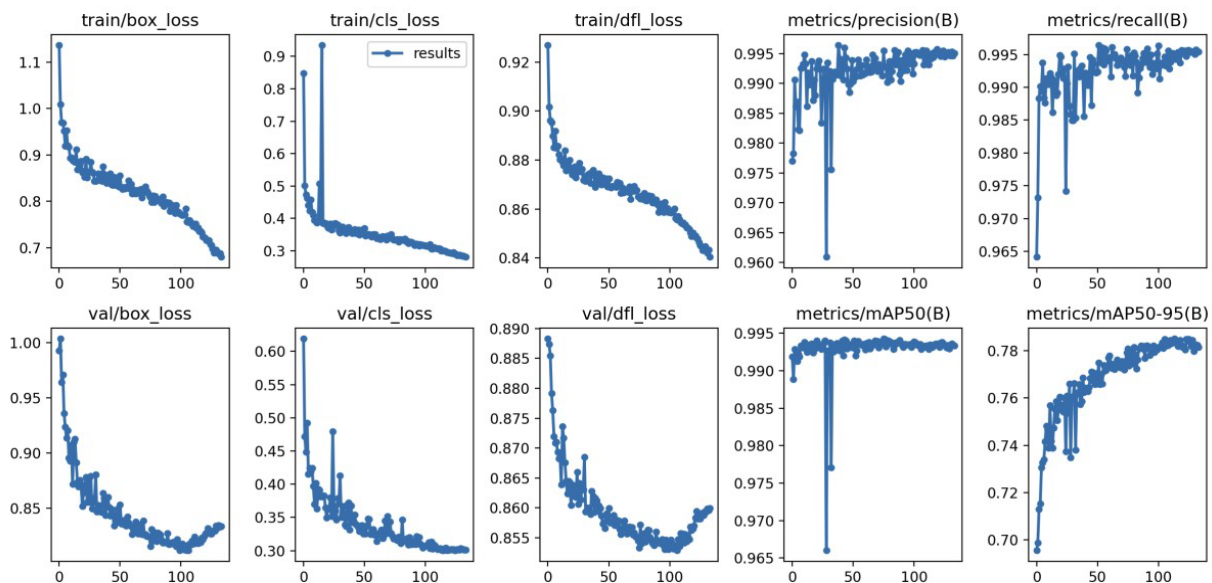
YOLOv5 generally achieves mAP scores ranging from (0.70-0.90), with precision around (80-90%) and recall between (75-90%) for detecting helmets and number plates. It performs competently but can have difficulties with small items or when objects are obstructed. YOLOv8, as an enhanced version, boasts improved mAP (0.85-0.95), precision (90-93%), and recall (80-95%). It is particularly adept at managing smaller objects, occlusions, and challenging environments.

Comparing YOLOv8 and YOLOv5: YOLOv8 demonstrates greater performance, particularly in identifying objects from challenging angles and under diverse lighting conditions. It is both faster and more accurate, making it suitable for real-time applications such as vehicle tracking and safety oversight.

Result-plot YOLOv5:



Result-plot YOLOv8:



Obstacles: Both models encounter difficulties related to occlusion, variations in angles, and environmental issues such as inadequate lighting. YOLOv8 is more proficient at addressing these challenges due to its optimized design.

Use Cases: These models are beneficial in various sectors for helmet detection (ensuring worker safety) and number plate recognition (utilized in traffic control and toll collection). In conclusion, while YOLOv5 is effective, YOLOv8 delivers enhanced detection precision, speed, and resilience, making it better suited for complex detection tasks that require real-time capabilities.

CONCLUSION AND FUTURE WORKS:

In summary, both YOLOv5 and YOLOv8 are exceptionally effective and efficient models for object detection in real-time, each presenting unique benefits based on the specific needs of the task. YOLOv5 continues to be a favoured option owing to its established stability, user friendliness, and robust community backing, making it well-suited for numerous practical applications. Conversely, YOLOv8 enhances the groundwork laid by YOLOv5, offering substantial advancements in architecture, performance, and adaptability, which makes it more appropriate for demanding tasks where higher precision and quicker inference are essential. Although YOLOv5 remains widely utilized across various applications, YOLOv8 signifies the next step in the YOLO series, delivering improved capabilities for tackling more sophisticated and intricate object detection challenges. In the end, the decision between YOLOv5 and YOLOv8 relies on the specific requirements of the project, choosing either the userfriendliness and reliability of YOLOv5 or the advanced performance and enhanced optimization capabilities of YOLOv8.

References

- [1] Yixiao Zhang et al, Research on YOLOv5 Vehicle Detection and Positioning System Based on Binocular Vision, 2024.
- [2] Sunil Kumar et al, Fusion of Deep Sort and Yolov5 for Effective Vehicle Detection and Tracking Scheme in Real-Time Traffic Management Sustainable System, 2023.
- [3] Wojciech Lindenheim-Locher et al, YOLOv5 Drone Detection Using Multimodal Data Registered by the Vicon System, 2023. <https://doi.org/10.3390/s23146396>
- [4] Qing An et al, Research on Safety Helmet Detection Algorithm Based on Improved YOLOv5s, 2023. <https://doi.org/10.3390/s23135824>
- [5] Chenyang Wei et al, Fast Helmet and License Plate Detection Based on Lightweight YOLOv5, 2023. <https://doi.org/10.3390/s23094335>
- [6] Ju Han et al, Safety Helmet Detection Based on YOLOv5 Driven by Super-Resolution Reconstruction, 2023. <https://doi.org/10.3390/s23041822>
- [7] Yiduo Zhang et al, FEFD-YOLOV5: A Helmet Detection Algorithm Combined with Feature Enhancement and Feature Denoising, 2023. <https://doi.org/10.3390/electronics12132902>
- [8] Weipeng Tai et al, DAAM-YOLOV5: A Helmet Detection Algorithm Combined with Dynamic Anchor Box and Attention Mechanism, 2023. <https://doi.org/10.3390/electronics12092094>

- [9] Shuai Chen et al, Helmet Wearing Detection of Motorcycle Drivers Using Deep Learning Network with Residual Transformer-Spatial Attention, 2022. <https://doi.org/10.3390/drones6120415>
- [10] Jianfeng Han et al, EGS-YOLO: A Fast and Reliable Safety Helmet Detection Method Modified Based on YOLOv7, 2024. <https://doi.org/10.3390/app14177923>
- [11] Ahatsham Hayat et al, Deep Learning-Based Automatic Safety Helmet Detection System for Construction Safety, 2022. <https://doi.org/10.3390/app12168268>
- [12] Maged Shoman et al, Enforcing Traffic Safety: A Deep Learning Approach for Detecting Motorcyclists' Helmet Violations Using YOLOv8 and Deep Convolutional Generative Adversarial Network-Generated Images, 2024. <https://doi.org/10.3390/a17050202>
- [13] P. R. Anisha et al, A clever deep feature-based prediction algorithm for metabolic syndrome in sleep disorders, Springer Multimedia Tools and Applications, 2023 <https://doi.org/10.1007/s11042-023-17296-4>.
- [14] C. Kishor Kumar Reddy et al, an intelligent optimized cyclone intensity prediction framework using satellite images, Springer Earth Science Informatics, 2023 <https://doi.org/10.1007/s12145-023-00983-z>.
- [15] Anisha P R et al, Intelligent Systems and Machine Learning for Industry: Advancements, Challenges and Practices, CRC Press, Taylor & Francis, 2022.