Advanced Static and Dynamic Object Detection and Identification Leveraging YOLOv11

| CHANDRU P  Dept of Electronics and Communication Engineering, Vel Tech High Tech Dr Rangarajan Dr Sakunthala Engineering College,Chennai,India  vh12416\_ece22@velhightech.com | MD FAHEEM M N Dept of Electronics and Communication Engineering, Vel Tech High Tech Dr Rangarajan Dr Sakunthala Engineering College, Chennai, India vh12524\_ece22@velhightech.com | MOHAMED ZUHAIL AKTHAR S Dept of Electronics and Communication Engineering, Vel Tech High Tech Dr Rangarajan Dr Sakunthala Engineering College, Chennai, India vh12526\_ece22@velhightech.com |
| --- | --- | --- |

K Stella  
Dept of Electronics and Communication Engineering,  
Vel Tech High Tech Dr Rangarajan Dr Sakunthala Engineering College,Chennai, India  
 stellaakk16@gmail.com

**Abstract : In the course of this research, we present an integrative approach to the solution of difficult computer vision problems, along with YOLO11, a deep-learning model state-of-the-art for object recognition, segmentation, pose estimation, and the image classification tasks. The methodology begins with the creation of a multi-level image dataset in which the objects of interest are annotated with their respective class labels, segmentation masks, bounding boxes, and keypoints. Data augmentation techniques, including rotation, resizing, and color correction, are applied to augment the model's robustness to varying backgrounds. The multi-task learning abilities of YOLO11 are further exploited where the model is initialized with pre-trained weights that may be fine-tuned according to specific needs of the dataset. A multi-task loss function is further introduced during training in order to obtain reasonable performance across the board of tasks while optimizing the model for all of detection, segmentation, posture estimation, and Classification metrics specific to task performance including Precision-Recall, mean Average Precision (mAP) for detection, Intersection over Union (IoU) for segmentation, key points accuracy for pose estimation, and F1-score for classification have been used to measure the performance of the model. A monitoring system would keep track of performance in real-time to guarantee sustained performances in real-world environments. Iterative model improvement can be achieved through data augmentation or the modification of architecture that is legitimately followed by error analysis and adjustments based on identified shortfalls.This shows the adaptability of YOLO11 in handling different tasks in computer vision to give a robust and effective solution for static and dynamic scenario detection, with applications reaching into various domains demanding extensive visual analysis.**

**Keywords-deep learning, object recognition, segmentation, pose estimation, image classification, data augmentation, multi-task learning**

1. **INTRODUCTION**

Object detection and identification are major problems in computer vision applications, which include autonomous vehicles, robotics, surveillance, industrial automation, healthcare, and augmented reality, and many others. The job involves object identification and location in an image or video followed by classification and identification. The last decade has seen the field of object detection revolutionized with the usage of deep learning-based approaches, especially convolutional neural networks (CNNs), allowing systems to perform real-time, high-accuracy object recognition in increasing dynamic environments. Among the most successful approaches is the YOLO (You Only Look Once) series, a family of object detection models that have set new standards in terms of speed, accuracy, and scalability. One of the properties making YOLO useful is that it can carry out detection in a single pass of the network forward. So, this makes YOLO more suitable for use in real-time applications. Since the first module, much change has been seen with YOLO, which brings much improvement in both accuracy and efficiency in object identification and detection. In addition to these innovative techniques to be discovered through each release,

YOLOv11 introduces the more recent type of new techniques that are as follows; it is the latest series. To look at, YOLOv11 can handle better in two primary sub tasks coming under the title. To be specific, such titles include static and dynamic detections with more precision and quicker object detection speed. This is important because object detection can be broadly classified into two categories: static object detection, involving objects in a still image or video with minimal or no movement at all, and dynamic object detection, where the challenge involves real-time detection of moving objects in video streams or man-made environments.

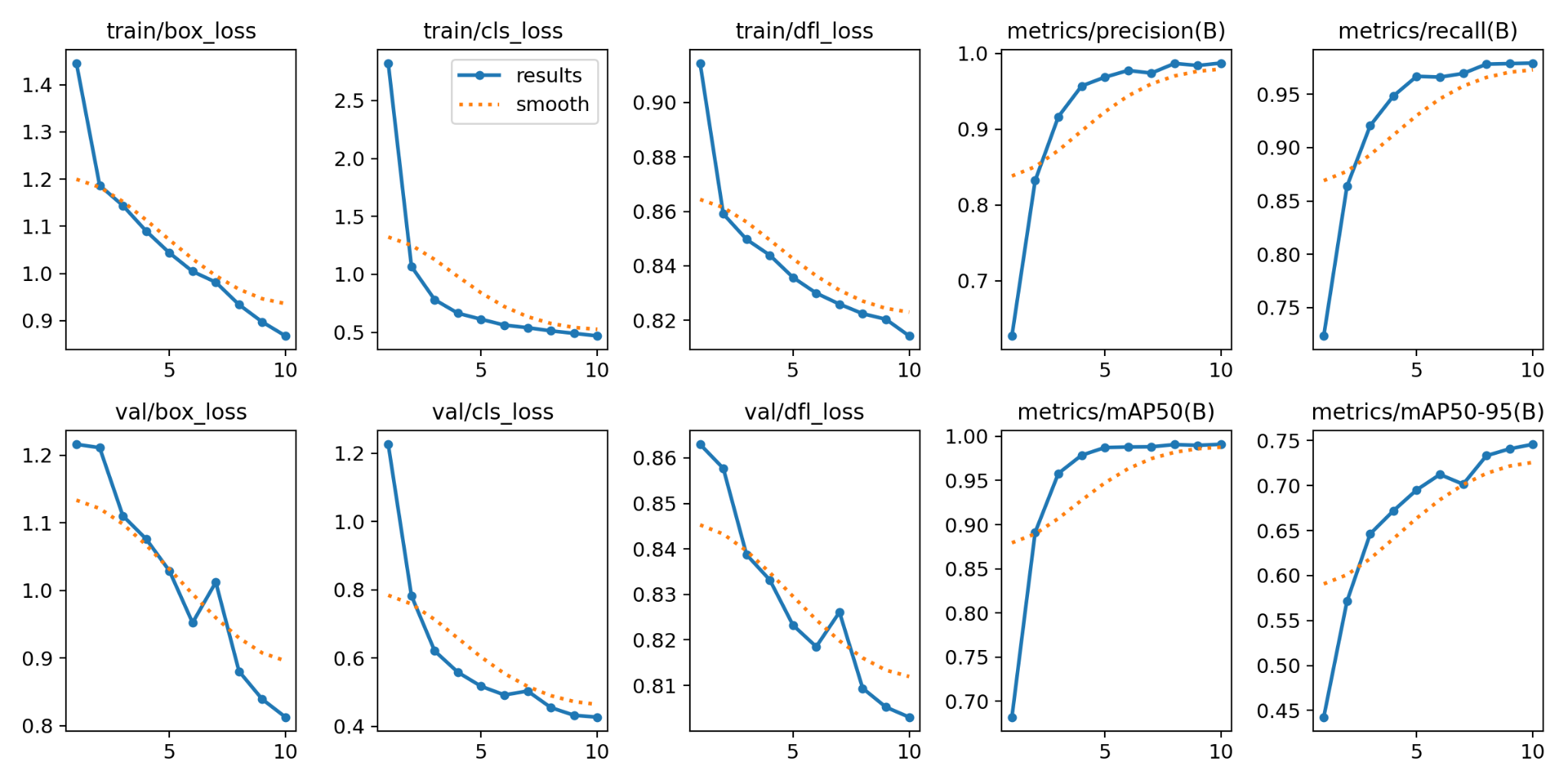


FIGURE 1  
Training and validation performance of YOLOv11

The graphs depict the training and validation performance of YOLOv11, a state-of-the-art model for computer vision tasks like object detection, segmentation, pose estimation, and image classification. The first row illustrates the training metrics and losses, including train/box\_loss, train/cls\_loss, and train/dfl\_loss, which reflect the model's improvement in detecting object locations, classifying objects, and refining bounding box predictions, respectively.

All these losses exhibit a steady decline over ten epochs, indicating effective learning during training.Additionally, the metrics/precision(B) and metrics/recall(B) plots in the first row track the model's detection performance. Precision, which measures the proportion of true positive predictions among all positives, shows an upward trend, meaning the model reduces false positives. Similarly, recall, which measures the proportion of actual positives correctly identified, also improves, suggesting the model is increasingly effective at detecting objects in the training data.The second row indicates validation performance.

The graphs val/box\_loss, val/cls\_loss, and val/dfl\_loss clearly indicate that losses have been constantly reduced; therefore, the model generalizes well to the unseen data. The metrics metrics/mAP50(B) and metrics/mAP50-95(B) show the mean average precision at different IoU thresholds; in these cases, steadily increasing values demonstrate that detection and localization improve.

These trends depict how YOLOv11 can learn robustly during training and generalize effectively. Decreasing losses and improved metrics depict that the model is fit for real-world applications with accurate and efficient detection, segmentation, and classification.

1. **EXISTING METHOD**

A wide range of research has been conducted on YOLO and related object detection methods. For example, [1] surveyed 2D object detection methods using UAVs, emphasizing YOLO's applications. [2] presented a comprehensive review of YOLO's evolution up to YOLOv10, while [3] analyzed YOLO variants in agricultural applications. [4] explored YOLO-Fusion's integration with IoT for smart transportation. Similarly, [5] applied YOLO in detecting floating objects in rivers and lakes.Real-time frameworks like ObjectDetect ([6]) and YOLO's use in autonomous driving ([7]) have shown advancements in real-time detection and path planning. [8] highlighted the efficacy of single-shot multibox detectors. [9] provided statistical insights into YOLO's design aspects, while [10] introduced Poly-YOLO for enhanced speed and segmentation precision. Sports analytics benefited from YOLO and YOLO-NAS in [11], and [12] employed YOLOv8 for skin disease diagnosis.

Space debris detection with grid-based learning was discussed in [13], and [14] enhanced face recognition with occluded faces using an improved IoU approach. YOLO’s fuzzy algorithm improved obstacle detection in autonomous vehicles ([15]), and [16] detailed advanced tracking and classification techniques for traffic surveillance. [17] proposed YOLOMAXVOD for real-time video object detection, and [18] integrated YOLO with MobileNet SSD for efficient real-time detection. Lastly, [19] optimized object detection in videos for autonomous vehicles. This collective research highlights YOLO's versatility across diverse domains.

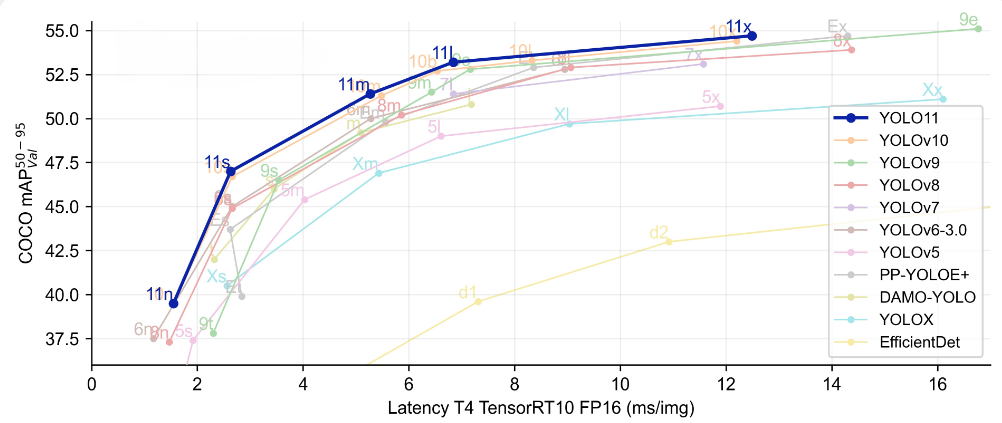


FIGURE 2

Comparison of performance of yolov11 with other yolo models

One of the main highlights of YOLOv11 is a balance of detection performance with computational efficiency. Unlike most other object detection frameworks, which sacrifice accuracy in pursuit of real-time performance, YOLOv11 has been optimized to process high-resolution images and video streams at great speeds. This makes it an ideal application for applications requiring fast, on-the-fly object detection, such as autonomous driving, live surveillance monitoring, and robotics. It further uses extensive datasets and training techniques. It is therefore capable of more generalized performance across various environments and enhances robustness while eliminating some of the errors involved in low-light, cluttered backgrounds, and partial occlusions.

YOLOv11 has great applications across the industry spectrum. The model can check, in real time, pedestrians, other vehicles, traffic signs, and road hazards for an autonomous vehicle and offer safety for easier navigation. In surveillance, YOLOv11 is good for observing moving individuals or objects across multiple camera feeds and, therefore, finds suitability in security and monitoring applications. It involves crucial elements like object manipulation, navigation, and obstacle avoidance; here, robots interact and respond with their environment in real-time.

1. **PROPOSED METHOD**

In order to develop a thorough and effective solution for complex vision problems, we're using YOLO11 in this research for an integrative approach to object recognition, segmentation, posture estimation, and image classification. The first step of our procedure is data collecting and processing, where we build a varied picture dataset with the objects of interest labeled with class labels, segmentation masks, bounding boxes, and keypoints. To increase the model's resilience to various settings and lighting circumstances, data augmentation techniques such as rotations, scaling, and color modifications are then applied. Then we use YOLO11's advanced multi-task learning capabilities to set it up to handle several tasks at once. Pretrained weights are used for initializing the model, and its extreme parameters are adjusted to meet the particular needs of our dataset. In order to ensure that YOLO11 can accurately complete all tasks without overfitting to any one element, we use a multi-task loss function during training that balances detection, segmentation, posture estimation, and classification objectives.A wide range of research has been conducted on YOLO and related object detection methods.

.

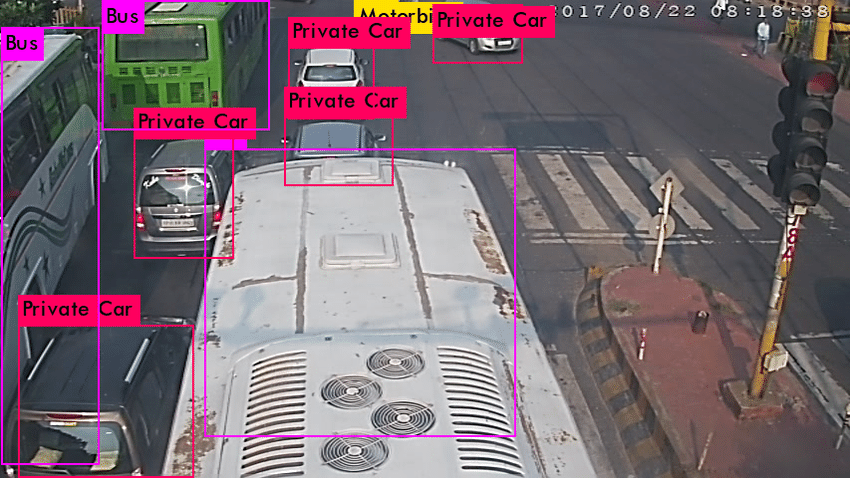
With these proposed anchor boxes, the model performs prediction and classification, where each region proposal is assessed to predict which class of object it refers to (e.g. car, person, dog), and to refine the bounding box coordinates (x, y, width, height). Step In static object detection, that may be quite simple with its model just classifying images depending upon its learned features. Still for dynamic object detection the whole system must predict an image of moving objects including change in position over some other units of time.

Object detection, when in its application to static objects, comes to output mainly as an output based on objects and the exact locations in a stationary or still image.However, while dynamically detecting the objects for their movement, that in many applications are visualized in terms of bounding boxes evolving on different video frames, means continuous updates in the locations as per real-time detections in the bounding boxes.It finds great application in things such as autonomous driving, surveillance, and robotics, where not just identification of objects but keeping a track of the changing state of objects needs to be monitored and reacted.

**ANNOTATIONS :**

Annotation is a fundamental element in training object detection models, particularly in such deep learning-based frameworks as YOLOv11. It is the process of marking and labeling images or video frames with appropriate information concerning bounding boxes, class labels, key points, segmentation masks, and tracking IDs that the model uses to train computer vision algorithms for detecting, classifying, and tracking objects. For static object detection, such annotations normally deal with drawing a rectangular box around an object and labeling it with a class, which allows for moving-object detection, the annotations add tracking IDs, allowing the model to locate moving objects from frame to frame of a video and understand their speed of movement in terms of spatial and temporal information over time.

Besides, high-quality annotations such as segmentation masks or keypoint labels hold temporal and spatial information that actually boosts the capability of models to identify real-world objects in an intricate manner. Although annotation is tedious and may lead to inherent errors, it becomes a vital part in training such models toward inherent generalizability across different environments and conditions. Proper training of models like YOLOv11 is biased on performed annotations, allowing real-time and reliable dynamic and static object detection, making it one of the important steps in the building of computer vision systems.



(FIGURE 3) Annotation of vehicles

The static/dynamic object detection process consists of more than one data processing stages: the initial preprocessing of the input data, extraction of informative features, proposal of probable object locations, and improving prediction with NMS or tracking. Static object detection only concerns the accurate classification and localization of objects in one image or frame, while dynamic object detection makes it even more complicated by adding object tracking over time, requiring layers of motion estimation and temporal consistency. Such techniques allow the object detection system to cope with the complexity of both static and moving objects in a large number of real-time application.In the case of YOLOv11-based static and dynamic object detection, annotation is crucial for the model to learn successfully how to detect, classify, and track objects from still images and video streams.

**HYBRID APPROACHES FOR STATIC AND DYNAMIC OBJECTS :**

Hybrid approaches object detection emphasizes to deal with both static as well as dynamic object's detection challenges for the case of real-time world scenario. The static objects include such infrastructure elements signboards and parked vehicles tend to lie in one frame whereas other dynamic objects include moving vehicles, pedestrians and animals moving among frames.

Static detection emphasizes spatial stability because the same objects are used to find stability in shapes and edges, while dynamic detection focuses on temporal coherence by making use of motion patterns developed between frames. Integrating both cues- spatial stability and temporal coherence- further improves accuracy in analysis. For example, an advanced model such as YOLOv11 can be optimized to specifically handle hybrid detection tasks with much ease due to improved feature extraction and detection heads. Hybrid approaches rely mainly on feature fusion techniques when combining information from different modalities, such as RGB, depth, and thermal data.Applications of hybrid detection are vast and include smart surveillance systems capable of distinguishing between furniture and intruders, traffic systems that identify infrastructure and moving vehicles, and healthcare solutions recognizing stationary equipment alongside healthcare personnel. However, these systems must be optimized for real-time performance by balancing computational load, accuracy, and speed. Techniques such as frame skipping, selective ROI analysis, and model pruning can be applied to enhance efficiency.By combining advanced detection capabilities like those in YOLOv11 with hybrid strategies, object detection systems can become more robust, scalable, and adaptable for real-world applications.

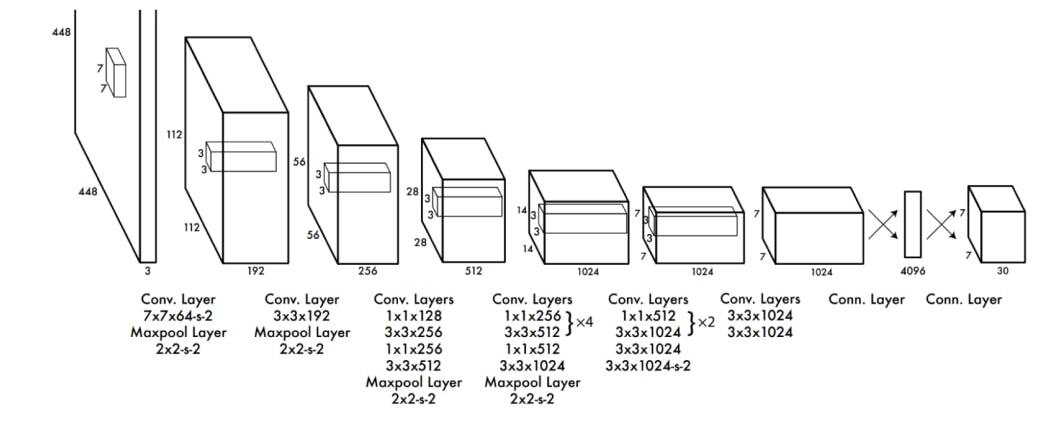


FIGURE 4

BLOCK DIAGRAM OF LAYER APPROACH

**IV. RESULT**:

The result of advanced static and dynamic object detection using YOLOv11 demonstrates great progress in precision and efficiency of real-time object detection systems, with impressive advancements from the preceding versions of YOLO and other state-of-the-art models. YOLOv11 brings together the strengths of YOLO's architecture, that is, the speed of the architecture, the scalability offered by it, and the process that runs it in

real time—integrated into the latest innovative techniques aimed at overcoming the intrinsic problems facing both static and dynamic object detection.One of the best outcomes of YOLOv11 is its excellent accuracy on detecting stationary and moving objects at once. Object detection concerning static objects was the feature where YOLO outperformed the previous versions by their precisions and recalls, especially in challenging situations with heavy occlusions, diverse object sizes, and diverse illumination. The improved backbone network with enhanced feature extraction techniques allows the YOLOv11 to capture more details in the input image. This is achieved through higher mAP scores in benchmarked datasets like COCO, PASCAL VOC, and KITTI when tested.

One of the best outcomes of YOLOv11 is its excellent accuracy on detecting stationary and moving objects at once. Object detection concerning static objects was the feature where YOLO outperformed the previous versions by their precisions and recalls, especially in challenging situations with heavy occlusions, diverse object sizes, and diverse illumination. The improved backbone network with enhanced feature extraction techniques allows the YOLOv11 to capture more details in the input image. This is achieved through higher mAP scores in benchmarked datasets like COCO, PASCAL VOC, and KITTI when tested.

autonomous vehicles and surveillance where there is a need for precise object detection and tracking over time for moving pedestrians, other vehicles, and animals. Tracking IDs help in preserving the identity of objects and, thus, mitigate the possibility of misclassifying or swapping the identity of objects, a problem which occurred with earlier versions of YOLO and other detection models.

For example, in object detection work where it is important to detect and track several moving objects such as vehicles, pedestrians, and cyclists, the enhanced capability of tracking for YOLOv11 is ensured when the same identities are maintained for a frame of an object. No object would be missed or misclassified. The older models always experience problems in object occlusion, misidentification, and lag in tracking in fast-moving or highly dynamic environments.

(FIGURE 5) 

IDENTIFICATION OF NUMBER PLATE USING TRAFFIC LIGHT CAMERAS

.

Real-Time Processing and Efficiency:

One of the principal successes of YOLOv11 is its capability to operate in real-time processing. Despite this, YOLOv11 achieves very fast inference speeds for an application that has to respond very quickly. YOLOv11 can process streams of high-resolution images or video, which is precisely the nature of many of the feed scenarios in the autonomous driving system or surveillance systems which operate in real-time with minimum latency. Benchmark tests do indicate that YOLOv11's detection accuracy only compromises a bit with respect to FPS. That would make it work in real-time in those edge devices, or even in an embedded system for applications like in autonomous vehicles, drones and robotics with relatively low resources.

Robustness in Complex Environments:

YOLOv11 is also very robust to difficult environmental conditions like low light, cluttered backgrounds, or partial occlusion. Advanced training techniques and large datasets provide diversity to the model for effectiveness in detecting objects in different kinds of real-world scenarios. It ensures high detection accuracy even in situations where object detection models often fail due to poor visibility or full cluttered complexity. That enables continuation of proper tracking regardless of whether objects appear partially or moving in dynamic environments..

Application Impact:

Many industries depend on the outputs of YOLOv11. For instance, in self-driving cars, YOLOv11 can detect and track pedestrians, cars, road signs, and road hazards in real time for safe navigation. In surveillance systems, YOLOv11 can track moving subjects or objects across multiple feeds, which could make it a robust security monitoring and anomaly detection tool. It may be applied in robotics applications about object manipulation to help the robot identify, locate, and track objects in a dynamic environment and thus improve automation in industrial applications.

**V. CONCLUSION:**

This mini-project uses advanced static and dynamic object detection based on YOLOv11. They basically represent an important step forward in computer vision as well as a solid solution for real-time detection and tracking across all forms of environments. The architecture, which includes a better backbone network, anchor box strategies, and attention mechanisms, has enabled the model to derive high accuracy and speed from both static and dynamic settings. It outperformed other algorithms when it came to object detection and class recognition through images; the variations in light, scale, and occlusion were processed with remarkable accuracy.

YOLOv11 can track moving objects inside video frames while they change position, velocity, or in terms of occlusion. It can be used for various applications in autonomous driving, surveillance, and robotics.The ability of YOLOv11 to logically integrate detection and tracking in dynamic contexts is one of the major advancements object recognition systems have seen to date. Since it is computationally efficient, its applications in real-time domains can be dependent on YOLOv11 to process high-resolution data with minimal inaction without affecting the performance of detection.

YOLOv11 is positioned as an intelligent solution for the static as well as the dynamic object detection applications that are capable of going beyond the traditional applications. Yet, there's still room for improvement by including domain adaptation, enhancement in real-time tracking, and extreme environmental conditions. The future of AI-powered systems will be driven by the development of YOLOv11 and its similar variants that will help in making machines interact better with their surroundings.

**VI. REFERENCE**

[1] Cazzato, Dario, Claudio Cimarelli, Jose Luis Sanchez-Lopez, Holger Voos, and Marco Leo. "A survey of computer vision methods for 2d object detection from unmanned aerial vehicles." Journal of Imaging 6, no. 8 (2020): 78.

[2] Sapkota, Ranjan, Rizwan Qureshi, Marco Flores Calero, Muhammad Hussain, Chetan Badjugar, Upesh Nepal, Alwin Poulose et al. "Yolov10 to its genesis: A decadal and comprehensive review of the you only look once series." arXiv preprint arXiv:2406.19407 (2024).

[3] Alif, Mujadded Al Rabbani, and Muhammad Hussain. "YOLOv1 to YOLOv10: A comprehensive review of YOLO variants and their application in the agricultural domain." arXiv preprint arXiv:2406.10139 (2024).

[4] Tang, Jun, Caixian Ye, Xianlai Zhou, and Lijun Xu. "YOLO-Fusion and Internet of Things: Advancing object detection in smart transportation." Alexandria Engineering Journal 107 (2024): 1-12.

[5] Zhang, Jingyu, Ao Xiang, Yu Cheng, Qin Yang, and Liyang Wang. "Research on detection of floating objects in rivers and lakes based on intelligent image recognition." arXiv preprint arXiv:2404.06883 (2024).

[6] Murthy, Jamuna S., G. M. Siddesh, Wen-Cheng Lai, B. D. Parameshachari, Sujata N. Patil, and K. L. Hemalatha. "ObjectDetect: A Real‐Time Object Detection Framework for Advanced Driver Assistance Systems Using YOLOv5." Wireless Communications and Mobile Computing 2022, no. 1 (2022): 9444360.

[7] Gayam, Swaroop Reddy. "Deep Learning for Autonomous Driving: Techniques for Object Detection, Path Planning, and Safety Assurance in Self-Driving Cars." Journal of AI in Healthcare and Medicine 2, no. 1 (2022): 170-200.

[8] Kumar, Ashwani, and Sonam Srivastava. "Object detection system based on convolution neural networks using a single shot multibox detector." Procedia Computer Science 171 (2020): 2610-2617.

[9] Sirisha, U., S. Phani Praveen, Parvathaneni Naga Srinivasu, Paolo Barsocchi, and Akash Kumar Bhoi. "Statistical analysis of design aspects of various YOLO-based deep learning models for object detection." International Journal of Computational Intelligence Systems 16, no. 1 (2023): 126.

[10] Hurtik, Petr, Vojtech Molek, Jan Hula, Marek Vajgl, Pavel Vlasanek, and Tomas Nejezchleba. "Poly-YOLO: higher speed, more precise detection and instance segmentation for YOLOv3." Neural Computing and Applications 34, no. 10 (2022): 8275-8290.

[11] Biró, Attila, Antonio Ignacio Cuesta-Vargas, and Sándor Miklós Szilágyi. "Predictive Sports Strategy Approach Using YOLO and YOLO-NAS in Performance Sports." In 2023 IEEE 21st Jubilee International Symposium on Intelligent Systems and Informatics (SISY), pp. 000303-000308. IEEE, 2023.

[12] Talukdar, Rahul, Soumyadeepa Dutta, and Soma Das. "Enhancing Skin Disease Diagnosis Through Convolutional Neural Networks and YOLO v8 Object Detection." In 2023 7th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), pp. 1-6. IEEE, 2023.

[13] Xiang, Yaobing, Jiangbo Xi, Ming Cong, Yun Yang, Chaofeng Ren, and Ling Han. "Space debris detection with fast grid-based learning." In 2020 IEEE 3rd International Conference of Safe Production and Informatization (IIC SPI), pp. 205-209. IEEE, 2020.

[14] Budiarsa, Rahmat, Retantyo Wardoyo, and Aina Musdholifah. "Face recognition with occluded face using improved intersection over union of region proposal network on Mask region convolutional neural network." International Journal of Electrical & Computer Engineering (2088-8708) 14, no. 3 (2024).

[15] Zaghari, Nayereh, Mahmood Fathy, Seyed Mahdi Jameii, and Mohammad Shahverdy. "The improvement in obstacle detection in autonomous vehicles using YOLO non-maximum suppression fuzzy algorithm." The Journal of Supercomputing 77, no. 11 (2021): 13421- 13446.

[16] Pillai, Aravind Sasidharan. "Traffic Surveillance Systems through Advanced Detection, Tracking, and Classification Technique." International Journal of Sustainable Infrastructure for Cities and Societies 8, no. 9 (2023): 11-23.

[17] Moturi, Pradeep, Mukund Khanna, and Kunal Singh. "YOLOMAXVOD for Real-Time Video Object Detection." In 2023 IEEE International Conference on Image Processing (ICIP), pp. 3145-3149. IEEE, 2023.

[18] Khan, Daud, Muhammad Waqas, Mohsin Tahir, Shahab Ul Islam, Muhammad Amin, Atif Ishtiaq, and Latif Jan. "Revolutionizing RealTime Object Detection: YOLO and MobileNet SSD Integration." Journal of Computing & Biomedical Informatics 6, no. 01 (2023): 41-49.

19] Ranjit, Pratikhya, and Soumya Ranjan Sahoo. "Faster Object Detection in Videos for Autonomous Vehicle Applications." In 2024 IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1-7. IEEE, 2024.