International Journal of Informatics and Computation (IJICOM)
Vol. 7, No.1, 2025
ISSN: 2685-8711, E-ISSN: 2714-5263 DOI: 10.35842/ijicom

Vehicle Theft Detection Using YOLO Based on License Plates and Vehicle Ownership

Bradika Almandin Wisesa[1], M. Hizbul Wathan[2], Evvin Faristasari[3], Sirlus Andreanto Jasman Duli[4], Better Swengky[5], Silvia Agustin[6]

Abstract

Detection of vehicle theft requires innovative approaches to address an increasing number of cases in Indonesia. This study presents a YOLOv11-based system for detecting vehicle theft by combining real-time object detection with a vehicle ownership database. The proposed system identifies license plates, detects vehicle owners using facial recognition, and analyzes suspicious activity to determine theft occurrences. The proposed method can produce model effectiveness with an accuracy = 70%. Key improvements in architecture, including enhanced feature fusion and dynamic anchor assignment, contribute to the object's detection in complex environments. This research can be a potential technique to provide efficient, scalable, and real-time security solutions in dynamic surveillance applications.

Keywords:
YOLO, Vehicle Detection, Security, License Plate, Vehicle Ownership

This is an open-access article under the CC BY-SA license

1. Introduction

Vehicle security issues, such as theft continue to rise, causing economic losses and posing a threat to public safety. It is a critical challenge that requires innovative solutions to deal with theft detection in real time. Traditional methods of monitoring and preventing vehicle theft, such as manual patrols and direct inspections, are often ineffective due to limited human resources and slow response times [2]. Detecting connected vehicles plays a vital role in transforming human mobility, tackling road congestion and road safety. However, they rely heavily on the security, accuracy, and stability of sensor readings and network data. When there are anomalies in the data, it is necessary to detect them on time and handle them. Existing detection methods often struggle to identify unknown types of anomalies and are difficult to deploy on computationally limited vehicle terminals [21].

Vehicle detection using conventional techniques, such as handcrafted feature extraction (e.g., Haar-like features, HOG) and traditional machine learning methods (e.g., SVM, AdaBoost), has several limitations that hinder its effectiveness in modern applications. One major issue is the lack of robustness to varying environmental conditions. Conventional techniques often struggle with changes in lighting, weather, and occlusions, leading to reduced accuracy in real-world scenarios. For example, shadows, reflections, or poor lighting can significantly degrade the performance of feature-based detectors. Additionally,

Corresponding Author: Bradika Almandin Wisesa(Bradika@polman-babel.ac.id)
1. Bradika Almandin Wisesa, Politeknik Manufaktur Negeri Bangka Belitung, Bradika@polman-babel.ac.id
2. M. Hizbul Wathan, Politeknik Manufaktur Negeri Bangka Belitung, mhizbul@polman-babel.ac.id
3. Evvin Faristasari, Politeknik Manufaktur Negeri Bangka Belitung, evvin@polman-babel.ac.id
4. Sirlus Andreanto Jasman Dudi, Politeknik Manufaktur Negeri Bangka Belitung, sirlusjasman@gmail.com
5. Better Swengky, Politeknik Manufaktur Negeri Bangka Belitung, Better@polman-babel.ac.id
6. Silvia Agustin, Politeknik Manufaktur Negeri Bangka Belitung, sisilagustia@gmail.com

these methods are less effective in handling complex backgrounds or crowded scenes, where vehicles may overlap or appear in cluttered environments. Another challenge is the limited generalization capability of conventional techniques. Handcrafted features are often designed for specific scenarios and fail to adapt to new or unseen data, making them less suitable for dynamic and diverse environments like urban traffic or highways [22].

Another critical issue is the computational inefficiency of conventional vehicle detection methods. Techniques like HOG detection require significant computational resources, especially when processing high-resolution images or video streams in real-time. This makes them less practical for applications requiring low latency, such as autonomous driving or real-time traffic monitoring. Furthermore, these methods often rely on multi-stage pipelines (e.g., sliding window approaches), which are computationally expensive and time-consuming. The lack of scalability to large datasets is another drawback, as traditional methods struggle to handle the vast amounts of data generated by modern sensors like cameras and LiDAR. This limits their applicability in large-scale systems, such as smart city infrastructure or fleet management [23].

Finally, conventional techniques face challenges in detecting small or partially visible vehicles. In urban environments, vehicles may appear at varying scales or be partially occluded by other objects, such as pedestrians or buildings. Traditional methods often fail to accurately detect such instances due to their reliance on predefined features and rigid detection frameworks. Moreover, these techniques are less effective in distinguishing between different types of vehicles (e.g., cars, trucks, motorcycles) or identifying specific attributes (e.g., color, make, or model). These limitations highlight the need for more advanced approaches, such as deep learning-based methods, which offer greater flexibility, accuracy, and robustness in vehicle detection tasks [24][25][26].

Advancements in deep learning, particularly in computer vision, offer new opportunities to address vehicle detection issues [1]. Recent deep learning methods produced a good performance for various anomaly data, especially when facing large datasets during the training process. It can reduce the computational time of the detection process in large scenario detection tasks. *You Only Look Once* (YOLO), a popular object detection algorithm, enables real-time analysis of images or videos to detect vehicles based on their license plates. It is a widely used technique in object detection, which enables real-time analysis of images or videos. By implementing a YOLO-based system, vehicles can be identified through their license plates and matched with ownership data to detect suspicious or stolen vehicles [2].

Therefore, this research aims to develop a YOLO-based vehicle detection for recognizing and verifying vehicle ownership to provide an innovative solution for improving vehicle security, assisting law enforcement in identifying unauthorized vehicles, and reducing the risk of vehicle theft in various regions. In this study, we use YOLO v11 for object detection with improved architecture for posing estimation and instance estimation.

2. Related Works

To address vehicle detection, conventional techniques utilized Two-dimensional (2-D) LiDAR as an efficient alternative sensor for vehicle detection, which is one of the most critical tasks in autonomous driving. Compared to the fully developed 3-D LiDAR vehicle detection, 2-D LiDAR vehicle detection has much room to improve. Most existing state-of-the-art works represent 2-D point clouds as pseudo-images and then perform detection with traditional object detectors on 2-D images. However, they ignore the sparse representation and geometric information of vehicles in the 2-D cloud points. To address these issues, they present a novel keypoint-aware and keypoint-matching network termed

KAM-Net, which focuses on better detecting the vehicles by explicitly capturing and extracting the sparse information of L-shape in 2-D LiDAR point clouds [27].

Deep learning-based surveillance, crime detection, and vehicle security systems have been the subject of numerous studies. Many communities proposed learning techniques to address vehicle detection which offer insights pertinent to the proposed study. For instance, a paper presented RANSAC and ICP algorithms to detect thief occurrences utilizing CCTV footage-based thief detection with CNN to identify the motion. In the scenario, the owner will receive an alarm message, a captured image, and options to call the police, fire department, or neglect. However, using trained faces, this research has not been able to identify car owners [3]. Another study combined IoT and CNN to construct a detection system for vehicle registration and recognition [29].

The current study explored the most popular real-time detection and recognition called YOLO [2][4][5][10][11]. An article established detection Based on YOLO, R-CNN, Faster R-CNN, and Fast R-CNN. According to the results, it took 60 seconds for R-CNN, 47 seconds for Faster R-CNN, 54 seconds for Fast R-CNN, and 7 seconds for YOLO to discover the burglary. It demonstrates that the fastest algorithm is YOLO. However, this research has not been able to detect vehicle owners based on trained faces.[7]. Another research combined KNN for object detection and R-CNN for image segmentation[6]. A study investigated YOLO to identify odd behaviors of someone covering their face. Although it focuses on person recognition, license plates, and vehicle recognition can also employ its methods. However, this research has not been able to detect motorcycles [8]. Another paper also proposed Real-time video integration of YOLO for real-time surveillance. The study demonstrates YOLO's ability to adapt to different lighting and occlusions in video data. Through real-time monitoring, the system detects suspicious activity and detection of car theft [9].

To harvest more prediction heads, a work presented by EfficientLiteDet to detect multi-scale targets effectively by utilizing Transformer Prediction Heads (TPH) instead of original anchor detection heads. To explore the potential of the self-attention mechanism in TPH, the proposed model integrates the “convolutional block attention model” to locate crucial attention regions in scenarios with denser targets. They applied various data augmentation strategies such as mosaic, mix-up, multi-scale, and random-horizontal-flip during the model training. Extensive experiment results showed that the EfficientLiteDet model has better performance in real-time scenarios. On Pascal Voc-2007, Highway and Udacity datasets, the proposed model achieves mean average precision (mAP) 87.3%, 80.1%, and 77.8%, respectively, which is quite better than Tiny-YOLOv4 state-of-the-art algorithm by + 2.4%, 1.8% and + 2.4%, respectively [30].

An article proposed YOLOv5 with a speed of 140 FPS, and then, the Deep Simple Online and Real-time Tracking (Deep SORT) is integrated into the detection result to track and predict the position of the vehicles. In the first phase, YOLOv5 extracts the bounding box of the target vehicles and it is fed with the output of YOLOv5 to perform the tracking. Additionally, the Kalman filter and the Hungarian algorithm are employed to anticipate and track the final trajectory of the vehicles. To evaluate the effectiveness and performance of the proposed algorithm, simulations were carried out on the BDD100K and PASCAL datasets. The proposed algorithm surpasses the performance of existing DL methods. Finally, the multi-vehicle detection and tracking process illustrated that the precision, recall, and mAP are 91.25%, 93.52%, and 92.18% in videos, respectively [28].

A recent study presented an innovative approach to revolutionize vehicle detection by using a single camera with the YOLOv11 combined with an ensemble OCR technique. The proposed system has achieved a Mean Average Precision (mAP) of 0.895 over a wide range of conditions, demonstrating 98.5% accuracy in license plate recognition, 94.2% accuracy in axle detection, and 99.7% OCR confidence scoring. The architecture incorporates intelligent vehicle tracking across IOU regions, automatic axle counting by way of spatial wheel detection patterns, and real-time monitoring through an extended dashboard interface. The solution shows improved performance while drastically reducing hardware resources compared to conventional systems. This research contributes toward intelligent transportation systems by introducing a scalable, precision-centric solution that improves operational efficiency and user experience [31].

3. Proposed Method

In this study, the first flowchart functions to detect registered objects, while the second flowchart is designed to detect objects that are at risk of being stolen. The flowcharts are shown in Figure 1 and Figure 2.

3.1 Object Detection Flowchart

This is the flowchart for implementing an object detection system using YOLOv11, from data preparation to generating detection results.

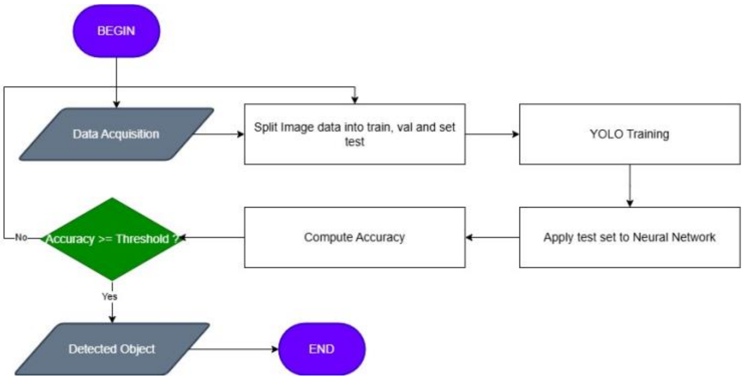

Fig 1. Object Detection Flowchart

3.2 Flowchart for thief detection.

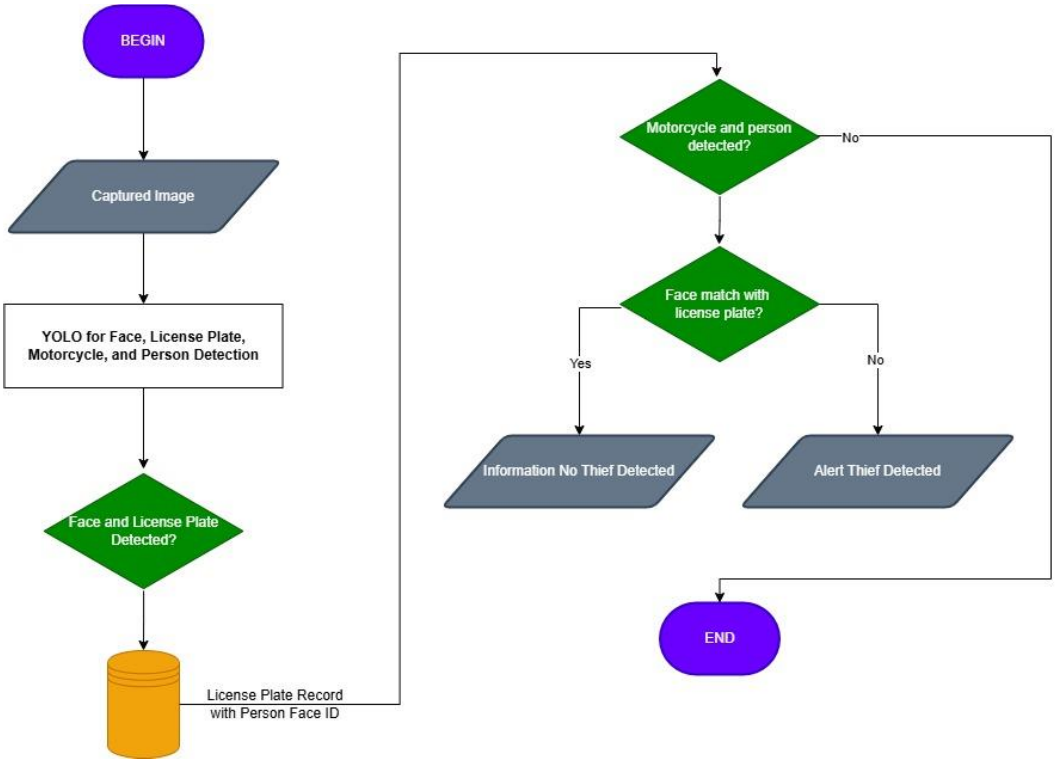

Fig 2. Object Detection Flowchart

3.3 Proposed DL Architecture: YOLOv11

The YOLO framework revolutionizes object detection by integrating bounding box regression and object classification into a single neural network architecture, consisting of a backbone for feature extraction, a neck for scale aggregation, and a head for final prediction, where YOLO v11 builds upon YOLO v8 with architectural innovations and parameter optimizations to enhance detection performance.

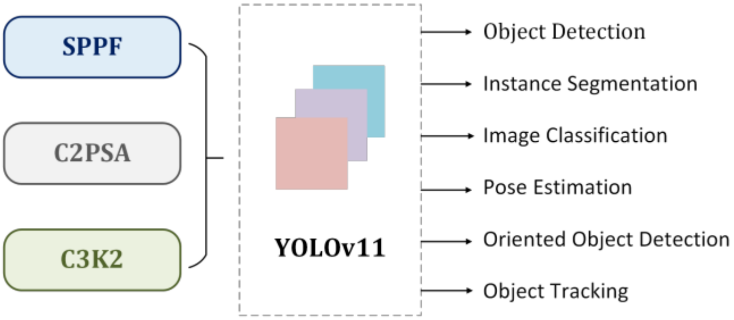

Fig 3. Key architectural modules in YOLO11

YOLO is a deep learning-based object detection algorithm that predicts bounding boxes and class probabilities directly from images in a single forward pass. The general mathematical formulation of YOLOv11 follows the principles of prior YOLO versions but integrates improved architecture and loss functions.

The core function of YOLOv11 can be represented as:

$$\hat{Y} = f_\theta(X) \tag{1}$$

where:

- X represents the input image,
- f_θ is the deep neural network model with learnable parameters θ,
- $\hat{Y}$ denotes the predicted output, including object classes, bounding box coordinates (x, y, w, h), and confidence scores.

The loss function in YOLOv11 is optimized to enhance detection accuracy and is given by:

$$\mathcal{L} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \mathbb{1}_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \lambda_{\text{size}} \sum_{i=0}^{S^2} \mathbb{1}_{ij}^{\text{obj}} \left[(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \right] + \lambda_{\text{conf}} \sum_{i=0}^{S^2} \mathbb{1}_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 + \lambda_{\text{cls}} \sum_{i=0}^{S^2} \mathbb{1}_{i}^{\text{obj}} \sum_{c \in \text{classes}} (p_{i,c} - \hat{p}_{i,c})^2 \tag{2}$$

where:

- The first two terms optimize bounding box predictions x, y, w, h
- The third term optimizes the confidence score C,
- The last term minimizes classification loss,
- λvalues are weighting factors to balance the loss components,
- $\mathbb{1}_{ij}^{\text{obj}}$ an indicator function determining whether an object exists in a given cell.

YOLOv11 enhances real-time object detection by using a highly optimized convolutional neural network (CNN) that divides an image into a grid system. Each grid cell predicts bounding boxes and class probabilities, allowing for fast and accurate detection. Compared to previous versions, YOLOv11 integrates improved transformer-based feature extraction, an enhanced attention mechanism, and a refined loss function to reduce localization and classification errors. By adopting these enhancements, YOLOv11 achieves superior performance in detecting small objects, reducing false positives, and maintaining real-time processing capabilities, making it ideal for applications such as autonomous driving, surveillance, and medical imaging.

Another key improvement in YOLOv11 is its multi-scale prediction strategy, which leverages advanced anchor-free detection to handle varying object sizes more effectively. The inclusion of dynamic label assignment and optimized IoU-based loss calculation ensures a more precise localization of objects. Additionally, it employs an adaptive learning strategy that refines predictions during training, leading to improved robustness against occlusions and complex backgrounds. By combining these architectural advancements, YOLOv11 pushes the boundaries of speed and accuracy in object detection tasks, making it one of the most powerful real-time detection models available.

By integrating YOLOv11 with a motorcycle thief detection system, this research provides a robust solution for real-time monitoring and rapid response to thief incidents. The architecture effectively identifies motorcycles and suspicious human activity, even in challenging environments.

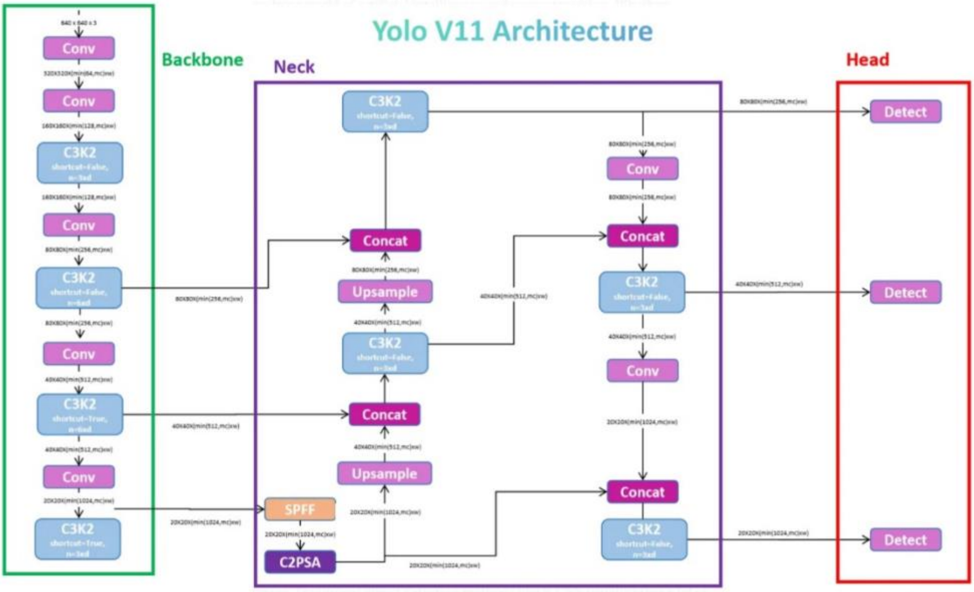

Fig 4. YOLOv11 Architecture

1. Backbone Network
One essential part of the YOLO architecture is the backbone, which is responsible for extracting features at different scales from the input image. Convolutional layers and specialized blocks are stacked to create feature maps at various resolutions. Convolutional layers, SPPF, and C2PSA make up the backbone. By using initial convolutional layers for downsampling images, decreasing spatial dimensions while increasing the number of channels, and replacing C2f with the C3k2 block—a more computationally efficient implementation of the Cross Stage Partial (CSP) Bottleneck—YOLOv11 preserves a structure similar to its predecessors. Unlike YOLOv8, which uses a single large convolution, C3k2 uses two smaller convolutions, allowing for faster processing without sacrificing efficiency [2]. Enhanced feature fusion combines residual connections and attention mechanisms to improve detection in crowded or low-visibility scenarios. CSPXNet is optimized for detecting motorcycles and suspicious activity using dynamic convolution [13], which adjusts kernel weights to capture a variety of motorcycle and human action features. [15].

2. Neck Component
Features from several scales are combined in the neck before being sent to the head for prediction. For the model to best capture multi-scale information, this procedure usually entails sampling and combining feature maps from several levels. The C3k2 Block and Attention Mechanism are located in the neck. By substituting the C3k2 block for the C2f block in the neck, YOLO11 makes a significant improvement. This block is intended to improve speed and efficiency in the feature aggregation process, where its integration leads to more optimal performance following upsampling and concatenation. To improve spatial attention, YOLO11 adds the C2PSA module. This enables the model to concentrate on important areas of the image for better detection, especially for small or partially obscured objects, setting it apart.

The enhanced PANet facilitates multi-scale detection essential for varying distances and perspectives in surveillance with BiFPN [16] efficiently aggregates features from different scales for robust object detection and also attention modules to highlight critical regions, such as motorcycles and nearby individuals, while suppressing irrelevant background noise. [17]

3. Head Module
   The head of YOLOv11 generates final object detection and classification predictions by processing feature maps from the neck to produce bounding boxes and class labels. The head contains C3k2 Block and CBS Blocks. In the head section, YOLOv11 utilizes multiple C3k2 blocks to efficiently process and refine feature maps. These blocks are placed in various pathways to handle multi-scale features at different depths. Their flexibility depends on the c3k parameter, where c3k = False makes it resemble the C2f block with a standard bottleneck structure, while c3k = True replaces it with the C3 module for deeper feature extraction. The key advantages of C3k2 include faster processing through two smaller convolutions and improved parameter efficiency. Additionally, YOLOv11 introduces the C3k block with adjustable kernel sizes to enhance detection accuracy. In the head section, YOLOv11 includes several CBS (Convolution-BatchNorm-SiLU) layers after the C3k2 blocks to refine feature maps by extracting relevant features, stabilizing data flow through batch normalization, and enhancing model performance using the SiLU activation function, ensuring more accurate detection results before proceeding to the bounding box and classification prediction stage [2].
   Key enhancements in the detection head address the challenges of thief detection with decoupled head architecture to separate classification [18] and localization tasks to improve precision in identifying motorcycles and human movements and dynamic anchor assignment Adapts to varying sizes and positions of objects, ensuring accurate localization of targets in real-time footage [19].

4. Final Convolutional Layers and Detect Layer.
   Each detection branch concludes with a series of Conv2D layers that refine the feature maps to produce the required outputs, including bounding box coordinates for object localization, object scores to indicate object presence, and class scores to identify the detected object’s category. The final Detect layer then consolidates these predictions into the model’s final output. To handle real-time requirements, YOLOv11 is optimized with IoU-aware and focal loss to enhance detection accuracy for small or partially obscured objects, quantization-aware training to reduce model size for deployment on edge devices like CCTV systems, and hardware acceleration to ensure smooth operation on GPUs or TPUs in real-time surveillance setups [20].

4. Experimental Setup

Data collection was carried out on January 19th, 2025, with images taken at various levels of light intensity and different viewing angles.

Table 1. Data Collection

No	Item	Pictures
1	License Plate	
2	People	

The training process was conducted on January 19th, 2025, using the YOLO algorithm that consists of Convolutional Neural Network (CNN), Grid Base Prediction, Anchor Boxes, Bounding Box Regression, Intersection over Union (IoU), Loss Function, Non-Maximum Suppression (NMS), Transfer Learning. The experiment was conducted on 2 different subjects with different lighting and viewing angles.

Table 2. Testing Detection of Proposed System

No	Item	Pictures
1	Subject 1	
2	Subject 2	

5. Result and Analysis

This system uses custom object detection to detect vehicle owners using face recognition of the owner and vehicle license plates and combines it with YOLO's built-in object detection [21] namely motorcycle and person object detection. First, vehicle owner detection is performed. This vehicle owner detection is taken from the vehicle_owner.txt file which contains the vehicle license plate and the owner's name, if detected, the vehicle license plate search will continue. If detected, it will enter the next stage, namely motorcycle and person detection. If all these objects are detected, the system will display the message "No Thief Detected" Conversely, if a person's object is detected but the face is not recognized according to the face of the vehicle owner, the system will display the message "Thief Detected".

Fig 5. Result of Subject 1

After testing on 4 different subjects, each with 10 trials with different angles and lighting, an accuracy of 70% was obtained with a confusion matrix as shown in the image below.

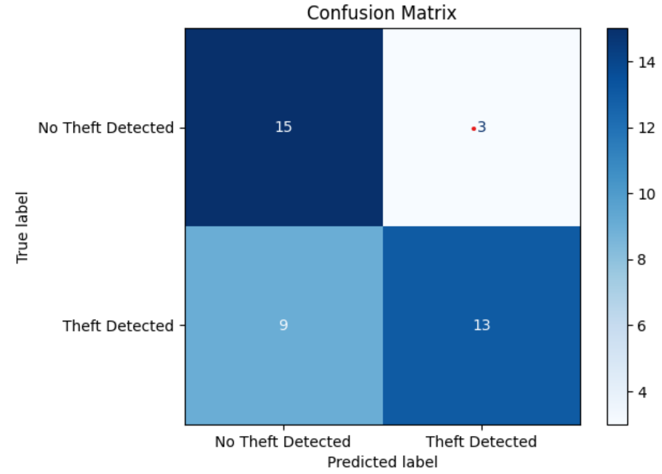

Fig 6. Confusion Matrix

Table 3. Evaluation of a Detection Model

	Precision	Recall	Accuracy	F1-Score	Support
No Theft Detected	0.62	0.83	0.72	0.71	18
Theft Detected	0.81	0.59	0.70	0.68	22

Table 3 provides an evaluation of a classification model detecting theft and non-theft instances using key performance metrics: precision, recall, accuracy, and F1-score. For the "No Theft Detected" class, the model achieves a precision of 0.62, meaning that 62% of the instances predicted as "No Theft" were correct. The recall is 0.83, indicating that the model successfully identified 83% of actual "No Theft" cases. The F1-score of 0.71 balances these two metrics, showing reasonable performance in this category. The support of 18 indicates the number of actual instances labeled as "No Theft."

The model exhibits a higher precision (0.81), meaning most of the theft predictions were correct. However, the recall is only 0.59, suggesting that 41% of actual theft cases were missed. The F1-score of 0.68 reflects a moderate balance between precision and recall. With a support of 22, this class had more actual theft cases to detect. The model appears to favor identifying "No Theft" over "Theft," as shown by its higher recall for non-theft cases. This imbalance suggests a need for improvements, potentially through better data representation or fine-tuning of classification thresholds.

6. Conclusion

This study presents an innovative solution by developing a YOLO-based system for vehicle theft detection. By integrating object detection with a motorcycle ownership database, the proposed model can demonstrate significant potential for real-time monitoring and anomaly detection applications. According to the experimental result, it can achieve 70% accuracy in identifying theft incidents with better precision and recall scores in the training process. However, the proposed approach needs improvement, particularly in handling variations in lighting conditions and extreme camera angles. Thus, it not only highlights the relevance of computer vision in security systems but also serves as a foundation for future theft detection studies. As future research, preventing vehicle theft can be integrated with IoT-based technologies, edge computing, or even automated surveillance networks. Additionally, dataset expansion and inference optimization are crucial steps to enhance efficiency and improve model generalization across real-world scenarios.

References

[1] A. Nikfarjam *et al.*, "On the Use of Quality Diversity Algorithms for the Travelling Thief Problem," *ACM Transactions on Evolutionary Learning and Optimization*, vol. 4, pp. 1–22, 2024.

[2] R. Khanam and M. Hussain, “YOLOv11: An overview of the key architectural enhancements,” *arXiv preprint arXiv:2410.17725*, Oct. 2024. [Online]. Available: http://arxiv.org/abs/2410.17725.

[3] A. Kushwaha, A. Mishra, K. Kamble, R. Janbhare, and A. Pokhare, “Theft detection using machine learning,” *IOSR Journal of Engineering*, 2018. [Online]. Available: www.iosrjen.org.

[4] A. K. Sharma, P. Mittal, R. Ranjan, and R. Chaturvedi, “Bank robbery detection system using computer vision,” in *Advances in Transdisciplinary Engineering*, IOS Press BV, Jan. 2023, pp. 619–625. doi: 10.3233/ATDE221322.

[5] N. Jegham, C. Y. Koh, M. Abdelatti, and A. Hendawi, “Evaluating the evolution of YOLO (You Only Look Once) models: A comprehensive benchmark study of YOLO11 and its predecessors,” *arXiv preprint arXiv:2411.00201*, Oct. 2024. [Online]. Available: http://arxiv.org/abs/2411.00201.

[6] S. Shirole, “Theft detection using deep learning,” *Research Square*, Nov. 2023. doi: 10.21203/rs.3.rs-3540282/v1.

[7] S. Pavithra and B. Muruganantham, “Panoramic video surveillance: An analysis of burglary detection based on YOLO framework in residential areas,” *Journal of Computer Science*, vol. 19, no. 11, pp. 1345–1358, 2023. doi: 10.3844/jcssp.2023.1345.1358.

[8] S. Ramzan et al., “Smart and secure vehicle parking system to avoid theft using deep image recognition,” *Journal of Innovative Computing and Emerging Technologies*, 2024.

[9] R. Kumar et al., “Detection of criminal activities/criminals through CCTV by live footage analysis,” in *2024 OPJU International Technology Conference on Smart Computing for Innovation and Advancement in Industry 4.0 (OTCON)*, IEEE, 2024. doi: 10.1109/OTCON60325.2024.10687573.

[10] J. Gao et al., “Camera-based crime behavior detection and classification,” *Smart Cities*, vol. 7, no. 3, pp. 1169–1198, Jun. 2024. doi: 10.3390/smartcities7030050.

[11] K. Menaka et al., “YOLO algorithm-based suspicious activity detection in ATM surveillance,” in *2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)*, IEEE, 2024. doi: 10.1109/AMATHE61652.2024.10582199.

[12] A. Turki et al., “A sophisticated drowsiness detection system via deep transfer learning for real-time scenarios,” *AIMS Mathematics*, vol. 9, no. 2, 2024. doi: 10.3934/math.2024156.

[13] “Electricity theft detection in smart grids based on deep neural network,” *International Journal of Creative Research Thoughts (IJCRT)*, 2023. [Online]. Available: www.ijcrt.org.

[14] J. Gao et al., “Camera-based crime behavior detection and classification,” *Smart Cities*, vol. 7, no. 3, pp. 1169–1198, Jun. 2024. doi: 10.3390/smartcities7030050.

[15] E. Yang and O. Yi, “Enhancing road safety: Deep learning-based intelligent driver drowsiness detection for advanced driver-assistance systems,” *Electronics*, vol. 13, no. 4, 2024. doi: 10.3390/electronics13040708.

[16] A. Nageswari and V. Usha, “Quick-witted security system,” *International Journal of Research in Applied Science and Engineering Technology (IJRASET)*, vol. 11, no. 7, pp. 816–824, Jul. 2023. doi: 10.22214/ijraset.2023.54663.

[17] N. S. Asl et al., “Toward applicable EEG-based drowsiness detection systems: A review,” *Frontiers in Biomedical Technologies*, vol. 9, no. 4, 2022. doi: 10.18502/fbt.v9i4.10426.

[18] M. Smita et al., “System to detect theft events using Raspberry Pi,” *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, vol. 2, no. 7, 2022. doi: 10.48175/IJARSCT-4447.

[19] A. Bade et al., “GSM-based vehicle theft detection using face recognition,” *MagPi Magazine*, vol. 10, 2022. [Online]. Available: https://magpi.raspberrypi.org/articles.

[20] M. Nighrunkar et al., “Theft detection: An optimized approach using cGAN and YOLO,” in *Proceedings of the International Conference on Advanced Computing and Applications (ICACA)*, Springer, 2023, pp. 325–332. doi: 10.1007/978-3-031-23724-9_30.

[21] W. Wang et al., “A vehicle abnormal behavior detection model in single intelligent vehicle scenarios,” *Journal of Internet Technology*, 2024.

[22] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 886–893.

[23] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001, pp. 511–518.

[24] S. Zhang, R. Benenson, and B. Schiele, “CityPersons: A diverse dataset for pedestrian detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 3213–3221.

[25] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? The KITTI vision benchmark suite,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 3354–3361.

[26] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 580–587.

[27] T. Zou et al., “KAM-Net: Keypoint-aware and keypoint-matching network for vehicle detection from 2-D point cloud,” *IEEE Transactions on Artificial Intelligence*, vol. 3, pp. 207–217, 2022.

[28] S. Kumar et al., “Fusion of Deep Sort and YOLOv5 for effective vehicle detection and tracking scheme in real-time traffic management sustainable system,” *Sustainability*, 2023.

[29] M. A. Jawale et al., “Implementation of number plate detection system for vehicle registration using IoT and recognition using CNN,” *Measurement: Sensors*, 2023.

[30] C. B. Murthy et al., “EfficientLiteDet: A real-time pedestrian and vehicle detection algorithm,” *Machine Vision and Applications*, vol. 33, 2022.

[31] K. Sivakoti, “Vehicle detection and classification for toll collection using YOLOv11 and ensemble OCR,” *arXiv preprint arXiv:2412.12191*, 2024.