Car parking lot availability detection using the faster R-CNN method

Andi Riansyah, Alif Hakim Al Faruq, Badieah
Department of Informatics Engineering, Faculty of Industrial Technology, Universitas Islam Sultan Agung, Semarang, Indonesia

Article Info

Article history:

Received Jun 5, 2025
Revised Apr 29, 2026
Accepted May 17, 2026

Keywords:

Computer vision
Faster R-CNN
Object detection
Parking area
ResNet50

ABSTRACT

Parking demand continues to rise as private vehicle use increases, making timely information about available spaces essential for efficient parking management. Many existing monitoring approaches still rely on fixed slot sensors or visual detectors that report accuracy without examining how confidence settings affect the final availability decision. This work investigates Faster region-based convolutional network (Faster R-CNN) with a ResNet-50 backbone for image-based parking availability detection using a public parking-lot dataset annotated in Pascal visual object classes (VOC) format. The experiment evaluates several confidence thresholds to determine how each setting changes the balance among accuracy, precision, recall, and F1-score. The most balanced setting was obtained at a threshold of 0.5, where the model achieved 95% accuracy and 97.3% for precision, recall, and F1-score. These results show that threshold configuration is an important factor in reducing missed detections and false alarms, although validation using real campus CCTV data and direct comparison with lightweight detectors remain necessary before practical deployment.

Corresponding Author:

Andi Riansyah
Department of Informatics Engineering, Faculty of Industrial Technology
Sultan Agung Islamic University
Kaligawe Raya street, Semarang, Indonesia
Email: andi@unissula.ac.id

1. INTRODUCTION

The continuous increase in private vehicle use has created new challenges for parking management in public and institutional environments [1]. University campuses experience this issue because cars are commonly used by lecturers, employees, students, and visitors for daily mobility. When the number of arriving vehicles is not supported by sufficient parking information, users often spend additional time searching for vacant spaces. This situation can reduce circulation efficiency and contribute to congestion near parking areas.

Parking facilities are designed to temporarily accommodate vehicles, but their effectiveness depends on both physical capacity and the availability of accurate occupancy information [2], [3]. In many locations, drivers still identify empty spaces manually because real-time slot availability is not displayed [4]. This manual search may increase fuel consumption, travel time, and disorder within the parking area. Therefore, automatic parking-space monitoring is needed to support more efficient and reliable parking management.

Computer vision has become a promising approach for parking monitoring because cameras can observe multiple spaces without installing sensors in every slot. Deep-learning-based object detectors, particularly convolutional neural network models, can process images and locate relevant objects automatically. Among these detectors, Faster region-based convolutional network (Faster R-CNN) is frequently adopted because it combines region proposal generation with object classification in a single detection pipeline. This structure enables reliable localization in visual scenes that contain dense objects or complex backgrounds [5].

A related study by [6] used Faster R-CNN for roadside parking-violation detection and reported 77.9% accuracy, 71.1% precision, 73.6% recall, and 77% mAP. Faster R-CNN improves region-based detection by using a region proposal network to generate candidate object areas before classification [7]. This mechanism allows the model to identify image regions that may contain parking-related objects before refining the final prediction. In the present work, the method is adapted to detect empty and occupied parking spaces from annotated parking-lot images.

Faster R-CNN was chosen as the main detector because localization quality is important in parking scenes that may include overlapping vehicles, narrow spaces, and different camera angles. One-stage detectors such as YOLOv8, EfficientDet, and MobileNet-SSD are generally more efficient for real-time inference, but they may require careful tuning when small or partially occluded objects appear. The two-stage design of Faster R-CNN requires more computation, yet it is suitable when the priority is accurate object localization. For this reason, the study uses Faster R-CNN with ResNet-50 while acknowledging that a complete speed-accuracy comparison with other detectors is still required [8].

Previous studies on parking detection often emphasize the final accuracy value and provide limited discussion of how confidence-threshold selection influences the reliability of the output. In real parking applications, this threshold is important because overly strict filtering can miss valid detections, whereas overly relaxed filtering can produce false alarms. This study therefore evaluates several threshold values instead of relying on a single default setting. The analysis focuses on how these values affect accuracy, precision, recall, and F1-score in parking availability detection [9].

The contribution of this paper is organized into three points. First, a Faster R-CNN detector with a ResNet-50 backbone is applied to image-based parking availability detection. Second, the study examines the performance trade-off produced by different confidence-threshold values and identifies the most balanced operating point. Third, the discussion outlines the limitations related to dataset representation, model comparison, and cross-dataset generalization as directions for future work [10].

2. RESEARCH METHOD

2.1. Dataset

The experiment used a public parking-lot image dataset from Roboflow with annotations provided in Pascal visual object classes (VOC) format. It contains 3,000 images, and each image is paired with an XML file describing the class label and bounding-box location of the annotated object. The images include variations in parking layout, vehicle density, illumination, and camera viewpoint. Because the images were not collected from the target campus, the dataset was used as a proxy to represent general CCTV-based parking conditions.

Data preparation was carried out before model training to check the suitability of the images and annotations. Samples with problematic labels, incomplete bounding boxes, or poor object visibility were removed from the training process [7]. The remaining data were separated into training and validation subsets for model development and performance observation. Augmentation was then applied to the training data using flipping, rotation, scaling, brightness and contrast adjustment, blur, and noise to improve robustness against common visual variations [9].

Figure 1 presents a sample image from the parking-lot dataset used in this research. Bounding-box annotations indicate the object position and class information required by the detector during training. The dataset includes scenes captured under different lighting, viewpoints, and parking densities. Such diversity helps the model learn visual patterns that are not limited to a single parking arrangement.

2.2. Experimental design

The experimental procedure was designed to study the influence of confidence-threshold settings on parking availability detection. After annotation preparation, the dataset was divided into training and validation subsets. The Faster R-CNN model using ResNet-50 was trained on the training subset, while validation data were used to observe model behavior during learning [8]. This setup provided a basis for evaluating whether the trained detector could be used for threshold analysis.

After training, the detector was tested with confidence thresholds of 0.9, 0.8, 0.7, 0.6, and 0.5. Each threshold was applied to the same evaluation setting to observe changes in detection outcomes. Model performance was measured using true positive, true negative, false positive, false negative, accuracy, precision, recall, and F1-score [9]. The purpose of this design was to find a threshold that minimizes missed detections while keeping false alarms at an acceptable level.

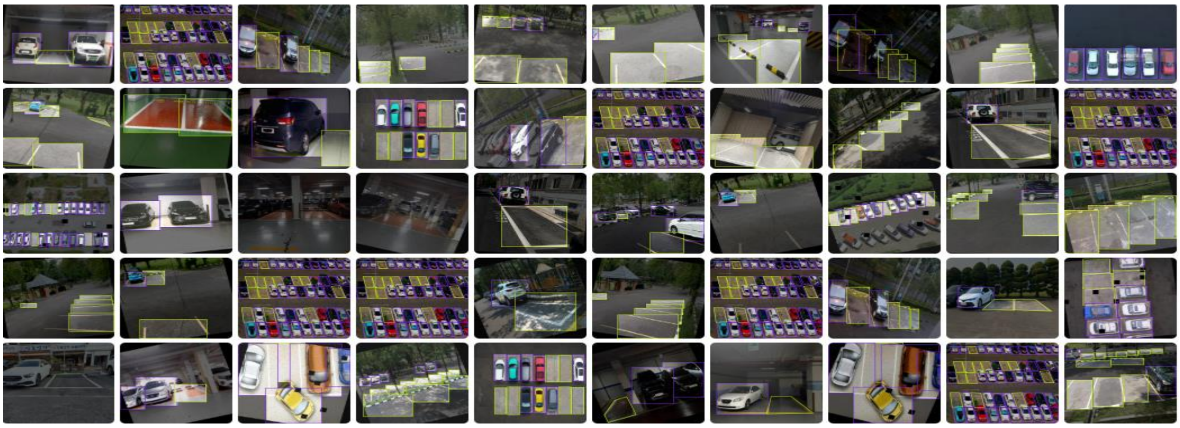

Figure 1. Example of the parking-lot dataset used in this study

2.3. ResNet-50 backbone

ResNet-50 was adopted as the backbone because it can extract deep visual features while maintaining stable learning through residual connections [8]. These connections support gradient flow and help the network learn complex visual representations more effectively [9]–[11]. In the proposed detector, ResNet-50 transforms parking images into feature maps that are processed by the region proposal and detection components [12], [13]. This backbone was selected to support object localization in scenes with variations in parking layout, viewpoint, and object density [14], [15].

2.4. Faster R-CNN model

Faster R-CNN was implemented as the detection framework because it separates candidate-region generation from the final object prediction stage [16]–[18]. The model architecture in this study includes a ResNet-50 feature extractor, a region proposal network that produces candidate regions, and a detection head that performs classification and bounding-box refinement [19]. This configuration was used to identify parking-related objects from annotated images [20], [21]. The final predictions were later filtered using different confidence thresholds during evaluation.

Figure 2 illustrates the Faster R-CNN workflow used in this study. The input image is first converted into feature representations by the ResNet-50 backbone. The region proposal network then proposes candidate regions that are likely to contain parking-related objects [22], [23]. These proposals are processed by the detection head to assign class labels and adjust bounding-box coordinates. The resulting detections are filtered by confidence score to examine the balance between missed detections and false alarms [24].

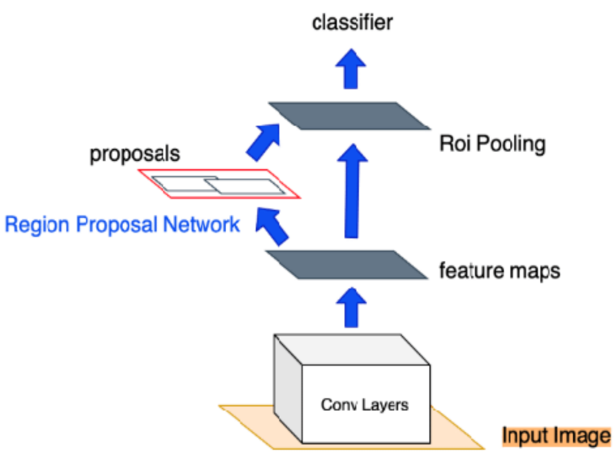

Figure 2. Layer architecture of Faster R-CNN

3. RESULTS AND DISCUSSION

3.1. Preprocessing

Preprocessing was performed to improve the reliability of the dataset before training and evaluation. Each image and Pascal VOC annotation file was reviewed to confirm the consistency of object labels, bounding boxes, and image quality. Images with incomplete annotations, unclear object visibility, or inconsistent labels were excluded [7]. This step reduced annotation noise and helped ensure that the model learned from more reliable samples.

3.1.1. Augmentation

Augmentation was used to expand visual diversity in the training data and reduce overfitting. The applied transformations included flipping, rotation, scaling, brightness and contrast modification, Gaussian noise, motion blur, normalization, and standardization. These operations were selected to imitate variations commonly found in CCTV parking images, such as different viewpoints, lighting conditions, object sizes, and image quality [9]. Examples of augmented samples are presented in Figure 3.

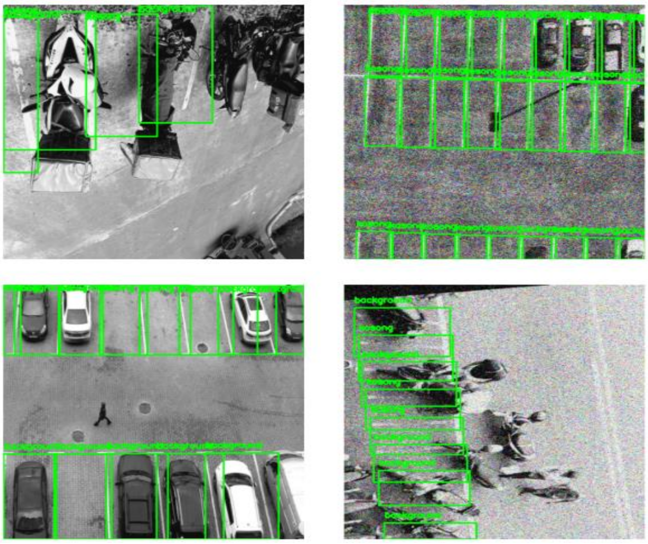

Figure 3. Examples of image augmentation applied to the training dataset

Figure 3 displays several augmented images derived from the original parking-lot dataset. The transformations changed image orientation, scale, brightness, contrast, blur, and noise while keeping the annotation information consistent. These additional visual variations encouraged the detector to learn more general patterns from the training data [7]. As a result, the risk of overfitting to the original image appearance could be reduced.

3.2. Modeling

The Faster R-CNN model with ResNet-50 was trained using the configuration listed in Table 1. The parameters were selected to keep the training process stable while remaining feasible for the available computational resources. This configuration also provided a consistent baseline before testing several confidence-threshold values [11]. The same trained model was used for the subsequent threshold-based evaluation.

Table 1. Hyperparameter

Parameter	Value
Batch size	3
Epoch	40
Learning rate	0.0003
Optimizer	Adam
Threshold	0.25
IoU threshold	0.5

Training was conducted for 40 epochs with the Adam optimizer, a learning rate of 0.0003, and a batch size of 3. An IoU threshold of 0.5 was used to decide whether a predicted box sufficiently matched the ground-truth annotation [25]. The confidence threshold shown in Table 1 was used during initial training and validation. Additional threshold values were applied later during testing to analyze their influence on detection reliability [26].The training behavior obtained from this parameter configuration is shown through the loss curves in Figure 4.

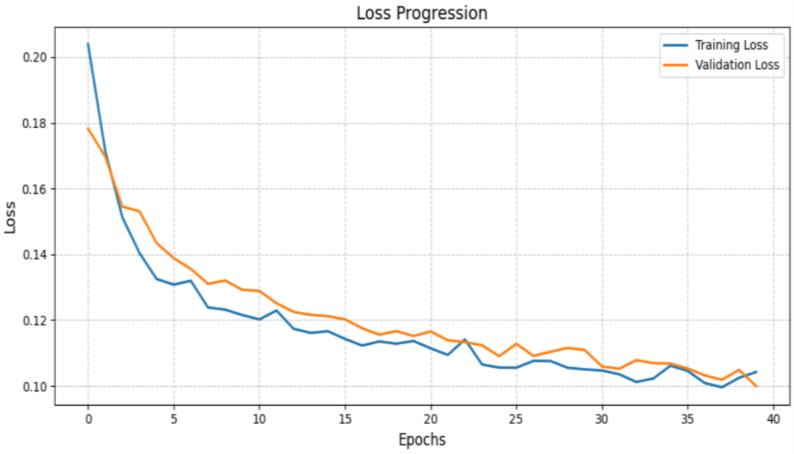

Figure 4. Training and validation loss

Figure 4 presents the training and validation loss of the Faster R-CNN model across 40 epochs. Both curves decrease gradually and approach a low loss value, showing that the model learned relevant visual patterns from the training data [21]. The validation curve follows the training curve closely, which suggests stable generalization during training. Based on this behavior, the trained model was then used for evaluation with different confidence thresholds.

Table 2 reports the mAP values obtained under different IoU criteria. The overall mAP across IoU thresholds from 0.5 to 0.95 was 0.2523, whereas mAP@0.5 and mAP@0.75 reached 0.3729 and 0.2929, respectively. These values show that the detector was able to identify parking-related objects, but bounding-box precision decreased when the IoU requirement became stricter. This result indicates that localization accuracy remains an area for improvement [15].

Table 2. Result mAP

Variable	Value
mAP [0.5:0.95:0.05]	0.2523
mAP @ 0.5	0.3729
mAP @ 0.75	0.2929

The moderate mAP results suggest that the model did not consistently produce highly precise bounding boxes for all test images. Variations in viewpoint, object scale, parking layout, and annotation quality may contribute to this limitation. For this reason, the threshold-based evaluation was further analyzed to assess detection reliability from the perspective of accuracy, precision, recall, and F1-score. This additional evaluation helps explain the operational behavior of the detector for parking availability decisions [17].

3.3. Testing and evaluation

Testing was performed using 40 parking-lot images selected from the validation and testing data. The selected samples contained visible parking boundary markings, including white and yellow lines. They also represented differences in viewpoint, object density, and lighting conditions. This test set was used to observe how the trained model responded to different visual conditions within the dataset.

Figure 5 shows sample detection outputs generated by the trained Faster R-CNN model. The detected parking area is enclosed by a green bounding box and displayed with its predicted class and confidence score [10]. In the example, an empty parking slot is detected with a confidence score of 0.97. This output demonstrates that the model can localize parking-related regions and produce availability information from image input.

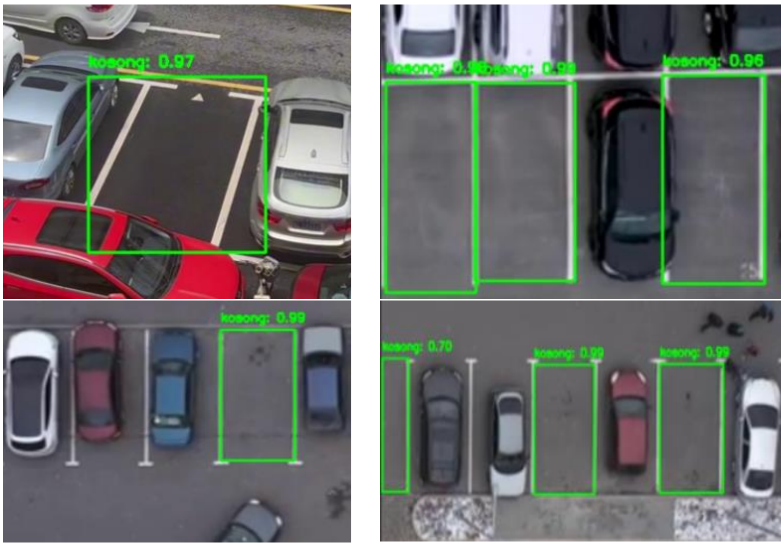

Figure 5. Example of parking-slot detection results

Several confidence thresholds were tested to examine how prediction filtering affects the reliability of parking availability detection. Thresholds from 0.9 to 0.5 were used to represent strict to more permissive detection settings. A high threshold accepts only predictions with strong confidence, which can reduce false positives but may also increase missed detections. A lower threshold accepts more predictions and can improve recall, although it may also introduce additional false positives [12]. Therefore, threshold analysis is necessary because both error types can affect the usefulness of parking information.

Table 3 shows that model performance changed substantially when the confidence threshold was adjusted. At 0.9, precision reached 100%, meaning that every accepted prediction was correct. However, recall was only 68.4% because 12 cases were missed by the strict threshold. When the threshold was reduced from 0.8 to 0.6, recall improved from 78.9% to 92.1% while precision remained unchanged at 100%. This pattern indicates that a moderate reduction in threshold allowed the model to recover more valid detections without adding false positives in that range [21].

Table 3. Confusion matrix

Threshold	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-Score
0.9	26	2	0	12	70%	100%	68.4%	81.2%
0.8	30	2	0	8	80%	100%	78.9%	88.2%
0.7	33	2	0	5	87.5%	100%	86.8%	92.9%
0.6	35	2	0	3	92.5%	100%	92.1%	95.9%
0.5	37	1	1	1	95%	97.3%	97.3%	97.3%

The threshold of 0.5 produced the strongest overall result in this experiment. At this setting, the model obtained 95% accuracy, 97.3% precision, 97.3% recall, and 97.3% F1-score. Although one false positive occurred, the number of false negatives decreased to only one case. This operating point therefore provided the best balance for the evaluated data [11]. In parking applications, such balance is important because missed availability information and false alerts can both reduce user trust.

3.4. Comparative discussion

The main experiment in this paper evaluates Faster R-CNN with ResNet-50 and investigates how confidence-threshold selection affects parking availability detection. To place the result in a broader context, this subsection discusses several alternative detectors, including YOLOv8, MobileNet-SSD, and EfficientDet. These architectures differ in detection strategy, computational demand, and deployment suitability. Such differences are relevant when designing parking monitoring systems for real-world use [20].

Faster R-CNN follows a two-stage detection process in which candidate regions are generated before classification and bounding-box refinement. YOLOv8 and MobileNet-SSD, by contrast, are one-stage detectors that predict classes and bounding boxes in a single pass to obtain faster inference. EfficientDet uses a scalable design intended to balance accuracy and efficiency. Consequently, detector selection should consider whether the application prioritizes localization accuracy, response time, model size, or hardware efficiency [21]. Table 4 provides a conceptual summary of the main characteristics of Faster R-CNN and several alternative object detectors for parking availability detection.

Table 4. Conceptual comparison of object detection architectures for parking availability detection

Model	Detection type	Main strength	Main limitation
Faster R-CNN ResNet-50	Two-stage detector	Strong localization capability and reliable feature extraction	Higher computational cost and slower inference
YOLOv8	One-stage detector	Fast inference and suitable for real-time applications	Requires careful tuning for small or partially occluded objects
MobileNet-SSD	One-stage lightweight detector	Lightweight and suitable for embedded devices	Lower robustness in complex visual scenes
EfficientDet	One-stage scalable detector	Balanced accuracy and efficiency	Performance depends on model scale and input resolution

Table 4 indicates that Faster R-CNN is more suitable for localization-oriented tasks, whereas YOLOv8, MobileNet-SSD, and EfficientDet are attractive for applications that require faster inference or lighter deployment. Because this study did not train and test all architectures on the same dataset, the comparison should be interpreted as conceptual rather than as a direct benchmark. Future research should evaluate these models using the same data split, evaluation protocol, and metrics. Relevant metrics include mAP, accuracy, precision, recall, F1-score, inference time, model size, and memory consumption.

Cross-dataset evaluation was not performed in this study and therefore remains an important limitation. The model was trained and tested on a public parking dataset that may not capture the full diversity of parking environments, camera placements, lighting conditions, vehicle types, and layout designs. Future work should train the model on one dataset and test it on another dataset collected from a different environment. This evaluation would provide stronger evidence about the robustness and generalizability of the proposed approach.

4. CONCLUSION

This study investigated Faster R-CNN with a ResNet-50 backbone for image-based parking availability detection using a public parking-lot dataset. The experimental results show that confidence-threshold selection clearly affects detection performance. Among the evaluated settings, a threshold of 0.5 achieved the most balanced result with 95% accuracy and 97.3% for precision, recall, and F1-score. This finding demonstrates that threshold adjustment is important for controlling missed detections and false alarms in parking monitoring systems.

Several limitations should be considered when interpreting the results. The dataset was obtained from a public source and may not fully represent real campus CCTV conditions. In addition, this study has not yet performed direct quantitative benchmarking against YOLOv8, SSD, EfficientDet, or other lightweight detectors. Future studies should include cross-dataset evaluation, computational efficiency measurement, and deployment testing on embedded or mobile platforms. These extensions are necessary to assess the feasibility of real-time parking availability systems.

ACKNOWLEDGMENTS

The authors express their appreciation to Sultan Agung Islamic University for supporting the completion of this research. The authors also thank all contributors who assisted during data preparation, model evaluation, and manuscript revision. Their support helped improve the quality of this study.

FUNDING INFORMATION

This research did not receive a specific grant from any public, commercial, or non-profit funding agency.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Andi Riansyah	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓	✓
Alif Hakim Al Faruq		✓	✓		✓	✓	✓	✓		✓	✓	✓		
Badieah	✓	✓		✓	✓			✓	✓			✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests related to this work.

DATA AVAILABILITY

The dataset analyzed in this study was obtained from a publicly accessible parking-lot dataset on Roboflow and prepared using Pascal VOC annotations. The data were used for training, validation, and testing of the proposed detector. Additional information about the dataset source is available from the corresponding author upon reasonable request.

REFERENCES

[1] H. Alkhudhayr, "Internet of things based parking slot detection and occupancy classification for smart city traffic management," *Engineering Applications of Artificial Intelligence*, vol. 152, p. 110802, Jul. 2025, doi: 10.1016/j.engappai.2025.110802.

[2] W. A. Jabbar, L. Y. Tiew, and N. Y. Ali Shah, "Internet of things enabled parking management system using long range wide area network for smart city," *Internet of Things and Cyber-Physical Systems*, vol. 4, pp. 82–98, 2024, doi: 10.1016/j.iotcps.2023.09.001.

[3] H. Khalkhali and S. H. Hosseinian, "Multi-stage stochastic framework for simultaneous energy management of slow and fast charge electric vehicles in a restructured smart parking lot," *International Journal of Electrical Power and Energy Systems*, vol. 116, p. 105540, Mar. 2020, doi: 10.1016/j.ijepes.2019.105540.

[4] T. Limbasiya, S. K. Sahay, and D. Das, "SAMPARK: Secure and lightweight communication protocols for smart parking management," *Journal of Information Security and Applications*, vol. 71, p. 103381, Dec. 2022, doi: 10.1016/j.jisa.2022.103381.

[5] K. Vijiyakumar, V. Govindasamy, and V. Akila, "An effective object detection and tracking using automated image annotation with inception based faster R-CNN model," *International Journal of Cognitive Computing in Engineering*, vol. 5, pp. 343–356, 2024, doi: 10.1016/j.ijcce.2024.07.006.

[6] R. A. Hamzah, C. Setianingsih, R. A. Nugrahaeni, S. R. Hanafia, and F. Fuadi, "Parking violation detection on the roadside of toll roads with intelligent transportation system using faster R-CNN algorithm," in *Proceedings - International Conference on Informatics and Computational Sciences*, IEEE, Sep. 2022, pp. 169–174. doi: 10.1109/ICICoS56336.2022.9930590.

[7] S. Wahjuni, Wulandari, and H. Nurarifah, "Faster RCNN based leaf segmentation using stereo images," *Journal of Agriculture and Food Research*, vol. 11, p. 100514, Mar. 2023, doi: 10.1016/j.jafr.2023.100514.

[8] I. Zidaoui, C. Wemmert, C. Joannis, S. Isel, J. Wertel, and J. Vazquez, "Application of ResNet and autoencoder models for anomaly detection in wastewater networks data: a comparative study of supervised and unsupervised approaches: ResNet and autoencoders for anomaly detection in wastewater data," *Flow Measurement and Instrumentation*, vol. 105, p. 102946, Sep. 2025, doi: 10.1016/j.flowmeasinst.2025.102946.

[9] S. Ingoley and J. Bakal, "Efficient models based on deep learning technique for indian sign language," *Procedia Computer Science*, vol. 258, pp. 318–331, 2025, doi: 10.1016/j.procs.2025.04.269.

[10] P. Bohlol, S. Hosseinpour, and M. S. Firouz, "Improved food recognition using a refined ResNet50 architecture with improved fully connected layers," *Current Research in Food Science*, vol. 10, p. 101005, 2025, doi: 10.1016/j.crfs.2025.101005.

[11] W. Wu, L. Huo, G. Yang, X. Liu, and H. Li, "Research into the application of ResNet in soil: a review," *Agriculture (Switzerland)*, vol. 15, no. 6, p. 661, Mar. 2025, doi: 10.3390/agriculture15060661.

[12] S. A. Hasanah, A. A. Pravitasari, A. S. Abdullah, I. N. Yulita, and M. H. Asnawi, "A deep learning review of ResNet architecture for lung disease identification in CXR Image," *Applied Sciences (Switzerland)*, vol. 13, no. 24, p. 13111, Dec. 2023, doi: 10.3390/app132413111.

[13] Madhavi, T. P. Singh, and S. P. Singh, "An improved system for detection of tea leaf disease using hybrid AI models," *Procedia Computer Science*, vol. 258, pp. 2573–2583, 2025, doi: 10.1016/j.procs.2025.04.519.

[14] G. Prethija and J. Katiravan, "Delving into transfer learning within U-Net for refined retinal vessel segmentation: an extensive hyperparameter analysis," *Photodiagnosis and Photodynamic Therapy*, vol. 53, p. 104620, Jun. 2025, doi: 10.1016/j.pdpdt.2025.104620.

[15] A. Alegría-Sala, D. M. Lopez, L. C. Casals, J. Fonollosa, and M. Macarulla, "The dilemma of variables assumptions in thermal comfort calculations for educational buildings: To Simplify or Not?," *Journal of Building Engineering*, vol. 84, p. 108404, May 2024, doi: 10.1016/j.jobe.2023.108404.

[16] K. Velumani *et al.*, "Estimates of maize plant density from UAV RGB images using faster-RCNN detection model: impact of the spatial resolution," *Plant Phenomics*, vol. 2021, p. 9824843, 2021, doi: 10.34133/2021/9824843.

[17] W. Zhang, Z. Sun, J. Mei, and B. Xu, "Intelligent recognition and parameter acquisition of blastholes in rock tunnel based on improved Faster R-CNN," *Tunnelling and Underground Space Technology*, vol. 149, p. 105790, Jul. 2024, doi: 10.1016/j.tust.2024.105790.

[18] C. Zheng, T. Liu, A. Abd-Elrahman, V. M. Whitaker, and B. Wilkinson, "Object-detection from multi-view remote sensing images: a case study of fruit and flower detection and counting on a central Florida strawberry farm," *International Journal of Applied Earth Observation and Geoinformation*, vol. 123, p. 103457, Sep. 2023, doi: 10.1016/j.jag.2023.103457.

[19] J. Cui *et al.*, "Weed identification in soybean seedling stage based on UAV images and faster R-CNN," in *Computers and Electronics in Agriculture*, Dec. 2024, p. 109533. doi: 10.1016/j.compag.2024.109533.

[20] A. Marwanto, A. Riansyah, and M. K. Anwar, "An improvement of apple leaf diseases detection using convolutional neural network methods based on mobile systems," in *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, IEEE, Sep. 2024, pp. 440–447. doi: 10.1109/EECSI63442.2024.10776345.

[21] Q. Fang, A. Jiang, M. Liu, and S. Zhao, "Faster R-CNN model for target recognition and diagnosis of scapular fractures," *Journal of Bone Oncology*, vol. 51, p. 100664, Apr. 2025, doi: 10.1016/j.jbo.2025.100664.

[22] A. Riansyah, A. T. Yulianto, N. A. Iahad, M. Qomaruddin, M. Taufik, and L. K. Wardhani, "Implementation of single shot detector method: smoking detection in no smoking area," in *2025 12th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, IEEE, Sep. 2025, pp. 267–274. doi: 10.1109/eecsi67060.2025.11290283.

[23] R. I. Chowdhury, J. Anjom, and M. I. A. Hossain, "A novel edge intelligence-based solution for safer footpath navigation of visually impaired using computer vision," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 8, p. 102191, Oct. 2024, doi: 10.1016/j.jksuci.2024.102191.

[24] K. A. S. K. and G. Balamurugan, "Faster region based convolution neural network with context iterative refinement for object detection," *Measurement: Sensors*, vol. 31, p. 101025, Feb. 2024, doi: 10.1016/j.measen.2024.101025.

[25] H. Lyu, F. Qiu, L. An, D. Stow, R. Lewison, and E. Bohnett, "Deer survey from drone thermal imagery using enhanced faster R-CNN based on ResNets and FPN," *Ecological Informatics*, vol. 79, p. 102383, Mar. 2024, doi: 10.1016/j.ecoinf.2023.102383.

[26] T. Hou and J. Li, "Application of mask R-CNN for building detection in UAV remote sensing images," *Heliyon*, vol. 10, no. 19, p. e38141, Oct. 2024, doi: 10.1016/j.heliyon.2024.e38141.

BIOGRAPHIES OF AUTHORS

Andi Riansyah is a computer science academic at Sultan Agung Islamic University. He earned his bachelor degree in Computer Science from Sultan Agung Islamic University and his master degree in Information Systems from Diponegoro University. His work focuses on computer science research and the development of applied information technology solutions. He can be contacted at email: andi@unissula.ac.id.

Alif Hakim Al Faruq is an undergraduate student at Sultan Agung Islamic University. His academic interests include computer science and artificial intelligence, particularly the application of deep learning for visual detection tasks. He has been involved in coursework and research activities related to intelligent systems. He can be contacted at email: alifhakim@std.unissula.ac.id.

Badieah is a full-time lecturer in the Department of Informatics Engineering at Sultan Agung Islamic University, Semarang, Indonesia. She holds a master degree in Information Systems from Diponegoro University and has professional experience in computer science and artificial intelligence. Her research interests include natural language processing, recommender systems, and predictive modeling with machine learning and deep learning. She can be contacted at email: badieah.assegaf@unissula.ac.id.