

YOLOv8-Based Infrared Intrusion Detection and Classification Model Using the OTCBVS Benchmark Dataset

Umar Abubakar Saleh^{1,2,3*}, Mazadu A. A.⁴, Solomon Zakwoi Iliya¹, Aliu Sala Saliu¹, Yakubu Yunusa¹, Mustapha Datti Mohammed¹, Aliyu Isiyaku¹, Hayatudeen Ogiri¹, Bilyaminu Usman⁴, Siti Amely Jumaat⁵

¹ Advanced Aircraft Engineering Laboratory, (AAcELab), Gusau, National Space Research and Development Agency, (NASRDA), Abuja 900108, NIGERIA

² Federal University Lokoja, Department of Electrical and Electronic Engineering, P.M.B. 1154, Lokoja, NIGERIA

³ Airforce Institute of Technology, Department of Electrical and Electronic Engineering, P M.B. 2104, NIGERIA

⁴ Federal Polytechnic Nyak, Shendam, Department of Electrical and Electronic Engineering, 940001, NIGERIA

⁵ Faculty of Electrical and Electronics Engineering, Universiti Tun Hussein Onn Malaysia 6400 Parit Raja, Batu Pahat, Johor, MALAYSIA

*Corresponding Author: abubakarumarsaleh1982@gmail.com

DOI: <https://doi.org/10.30880/jeva.2026.07.01.008>

Article Info

Received: 27 February 2026
 Accepted: 27 April 2026
 Available online: 30 June 2026

Keywords

YOLOv8, infrared intrusion detection, smart infrastructure protection, deep learning, transmission line surveillance

Abstract

Infrared imaging provides a reliable technique of monitoring and detecting intrusions in low-light and harsh environmental situations. This paper presents an optimized deep learning model for infrared intrusion detection and classification utilizing the You Only Look Once version 8 (YOLOv8) framework. The model was developed using the OTCBVS Benchmark Dataset Collection, pre-processed and annotated with Roboflow, and trained in a Google Colab Pro GPU environment. Multiple hyperparameter tuning and picture augmentation procedures were used to enhance detection robustness in a variety of environmental and operational conditions. The numerical evaluation revealed a precision of 0.999, recall of 0.929, mAP50 of 0.937, and mAP50-95 of 0.70, as against the non-optimized YOLOv8 (baseline YOLOv8) with precision of 0.740, recall of 0.650, mAP50 of 0.721, and mAP50-95 of 0.453, respectively. This shows outstanding accuracy and generalization across various test conditions. The smooth convergence of the training and validation loss curves shows the model's stability and lack of overfitting. Visualizing bounding box predictions and confusion matrices proves the efficacy of intrusion target localization and classification. The results establish the proposed YOLOv8-based model as a highly efficient and precise framework for infrared-based intelligent surveillance and smart power infrastructure protection.

1. Introduction

Electricity transmission line safety and other critical infrastructure assets is a key concern in modern nations. Illegal entrance, vandalism, and theft can all result in substantial safety risks, business interruptions, and financial

losses. Conventional monitoring technology, such as motion sensors and visible-light cameras, regularly fail under low illumination, fog, or nighttime situations, which are common in outdoor power installations [1].

Recent advancements in computer vision and artificial intelligence (AI) have enabled very accurate real-time object detection. Among these, the You Only Look Once (YOLO) family has gathered acceptance for its single-stage detection architecture, which enables rapid and precise localization in a single forward pass [2]. The most recent version, YOLOv8, integrates architectural changes such as decoupled loss functions, an anchor-free detection head, and better Cross-Stage Partial (CSP) backbone features that jointly increase inference efficiency, feature extraction, and convergence [3].

Experimental investigate show that YOLOv8 constantly offers cutting-edge detection accuracy across numerous domains and performs well even with little training data [3, 4]. Comparative analyses also reveal that YOLOv8 outperforms previous versions, including YOLOv5 and YOLOv7, in terms of recall, precision, and computational efficiency, particularly in real-time human identification under changeable ambient conditions [5]. However, few research has been performed on the optimization of YOLOv8 in infrared intrusion detection, mainly employing conventional thermal benchmarks such as the OTCBVS Benchmark Dataset Collection. Infrared photography offers additional problems, such as background clutter, temperature-dependent change, and low contrast, which might hinder model generalization if not addressed appropriately. Hence, an enhanced, data-driven framework is essential to improve YOLOv8's detection performance under such situations.

This paper develops an optimized YOLOv8-based infrared intrusion detection and classification model trained on the OTCBVS Benchmark Dataset. The Roboflow was employed for dataset annotation and preprocessing, while Google Colab Pro GPU environment with considerable data augmentation and hyperparameter tuning was used for the training. The goal is to achieve high-precision detection and reliable localization of intrusion events under several environmental and operational conditions, while validating model stability through performance evaluations and convergence analysis.

2. Related Work

Infrared-based intrusion detection is generally been studied due to its reliability under adverse weather conditions and low-illumination. Conventional image-processing techniques such as thermal-blob analysis, background subtraction and edge detection, were employed in early research. While these methods offer a foundation for recognizing vehicle or human presence, they were greatly susceptible to camera-angle changes, noise, and temperature variation, restrictive their resilience in real-world conditions [6].

With the advent of deep learning, specifically convolutional neural networks (CNNs), revolutionized object detection. YOLO family architectures and the Faster R-CNN, SSD, presented substantial inference speed and trade-offs in accuracy [7]. Among these, the YOLO established single-stage detection, authorizing for real-time inference in contrast to region-proposal approaches. Later generations YOLOv3 through YOLOv7 enhanced detection accuracy by redesigning loss mechanisms and upgrading backbone networks [8]. Nevertheless, the common of these models were tuned for visible-light data, decreasing their effectiveness on infrared images with less defined object boundaries.

Recent study has extended YOLO models into infrared surveillance domains and thermal. Li et al. [9] used infrared modalities and adaptive data fusion of visual to established a YOLOv5-based pedestrian detector for nighttime monitoring which increased the recall system. Wu and Zhang (2023) fine-tune YOLOv7 on the FLIR ADAS dataset through transfer learning. The outcome results in improved precision during fog and rain. Furthermore, Jiang et al. [10] revealed that UAV-based thermal-infrared object identification utilizing YOLO versions can achieved suitability during airborne surveillance. Their results validate the potential of YOLO designs in thermal imaging. The findings highlight the need for further tuning to address small-object localization. And ambient variance.

The anchor-free prediction layers, C2f backbone, and decoupled detection head were among the structural developments included in the most recent YOLOv8 architecture [11]. These advancements enable strong multi-scale detection, increase feature extraction efficiency and, showcasing YOLOv8 as an ideal for low-contrast infrared images. Ding [12] employed a lightweight YOLOv8 variation for faster inference reporting with equal accuracy and infrared object detection. Similarly, Rizk and Bayad [13] utilize YOLOv8 in thermal human identification, search and rescue missions, signifying its elasticity for real-time applications. Furthermore, Xie et al. [14] established YOLOv8's consistency and reliability by using it to detect infrared defects in Photovoltaic (PV) modules, indicating its good generalization for diverse thermal datasets.

Although substantial developments have been made in these area, limited study investigated YOLOv8 for standardized thermal-infrared datasets, such as the OTCBVS Benchmark Collection. This dataset offers a reliable framework for comparing infrared object-detection techniques. As a result, this paper builds on previous study by developing an optimized YOLOv8-based infrared intrusion detection model leveraging dataset-specific preprocessing and augmentation techniques (Roboflow + Google Colab Pro). This work sets a replicable baseline

for future research on intelligent infrared surveillance systems. The study set a replicable baseline for future investigation on intelligent infrared surveillance systems.

2.1 Research Gap

While previous studies have utilized YOLO structures for restricted thermal detection applications and visible-light detection, limited study focused only on infrared intrusion detection. The majority of the existing works either use older YOLO versions like v5 or v7 or combine visible and thermal modalities [13]-[15]. As a result, optimal YOLOv8 models trained solely on infrared datasets, such as the OTCBVS Benchmark Dataset, which are yet to be explored. Furthermore, existing studies does not offer all-inclusive review of YOLOv8's convergence behavior or augmentation techniques for thermal images. This paper addresses this gap by developing and improving a YOLOv8-based infrared intrusion detection and classification model, that was trained and verified on the OTCBVS Benchmark Dataset.

3. Methodology

This research employed the OTCBVS Benchmark Dataset Collection techniques to developed an optimized YOLOv8-based infrared intrusion detection and classification model. The methods comprise of four key phases namely dataset preparation, preprocessing and augmentation, model training and optimization, and performance evaluation. Individual phase was developed to offer duplicability and robustness under varying environmental conditions

3.1 Dataset Description

The Benchmark Dataset encompasses multispectral images for object detection and tracking, encircling both visible and infrared classifications. In this paper, the infrared subset was utilized. Similarly, photos of vehicular and human targets at changing weather conditions, distance and light variation. 400 photos were employed for the training, 100 was used for validation, and 50 for testing. The dataset's diversity permits for the evaluation of model performance under realistic surveillance situations [16].

3.2 Data Preprocessing and Augmentation

Fig. 1 shows the entire procedure involved in developing the infrared intrusion detection system. The technique begins with the collection of thermal pictures from the OTCBVS Benchmark Dataset, which are suitable for surveillance in low-light and nighttime conditions. Roboflow was used for dataset management and all image annotations, including the translation of bounding-box labels into YOLO-compatible format. This comprised scaling of all frames to 640×640 pixels, normalization, and histogram-based contrast improvement to enhance image quality and decrease thermal noise was achieved.

For model generalization improvement and reduction of overfitting, augmentation operations for data enhancement were used, these includes horizontal flipping, random scaling, hue-saturation correction, and mosaic composition to increase model generalization and decrease overfitting. Finalized dataset was exported as a YOLOv8-formatted directory for cloud training in Google Colab Pro with cloud-based GPU resources. This process guarantees efficient data preparation, which significantly contributes to the accuracy and reliability of the intrusion detection system.



Fig. 1 Workflow of the proposed YOLOv8-based infrared intrusion detection and classification framework

3.3 Model Configuration and Training

For the ability to strike a compromise between inference speed and accuracy, a YOLOv8s variation was selected and NVIDIA T4 GPU was employed for training in Google Colab Pro. The learning speed rate was varied between (0.001 – 0.01), and batch size [17], while optimizing the hyperparameters. Using early-stopping criteria, the training was performed for 50 epochs to avoid overfitting, similarly, the stochastic-gradient-descent (SGD) optimiser with cosine learning-rate scheduling was also used to achieved an optimized result.

The neck utilized multi-scale feature fusion to increase the identification of both small and large objects at various distances. The anchor-free detection head enhances object localization accuracy by anticipating object positions without using predetermined anchor boxes. This improves the detection precision in thermal surveillance conditions. Generally, the design enables fast and precise intrusion detection, making YOLOv8 ideal for security monitoring applications.

To achieved the continuously monitoring of loss and metric tracking, producing validation-loss charts, confusion matrices, and precision-recall curves, the integrated Ultralytics logging tool was used and the outcomes validated consistent training and smooth convergence. Fig. 2 shows the YOLOv8 architecture for infrared object detection, which includes the CSP backbone, neck, and anchor-free detection head. The CSP backbone aids in the extraction of critical thermal properties while lowering computational complexity, making the model appropriate for real-time detection.

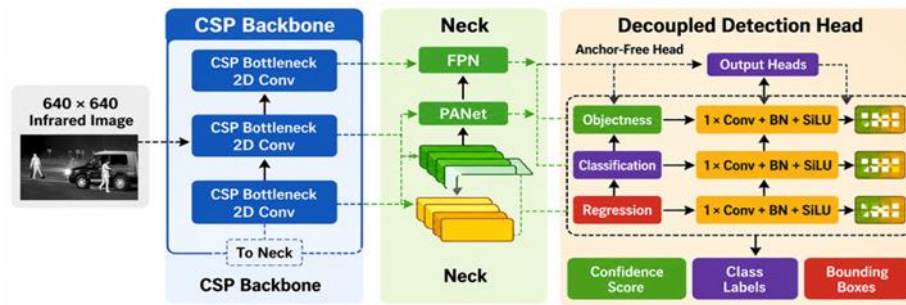


Fig. 2 Structural representation of the YOLOv8 architecture highlighting the Cross-Stage Partial (CSP) backbone, sneck, and anchor-free detection head for optimized infrared object detection

4. Experimental Results and Discussion

The optimized experimental results of the YOLOv8-based infrared intrusion detection and classification model, that was trained and verified using the OTCBVS Benchmark Dataset Collection is presented and discussed. The findings revealed that the model can effectively detect and classify incursion behaviors which includes human presence, light situations and movement in a variety of environments. The examination emphasises the three crucial areas: including qualitative detection visualization, model convergence behavior, and quantitative performance measurements. Generally, the numerical evaluation revealed a precision of 0.999, recall of 0.929, mAP50 of 0.937, and mAP50-95 of 0.70, as against the non-optimized (baseline YOLOv8) with precision of 0.740, recall of 0.650, mAP50 of 0.721, and mAP50-95 of 0.453, respectively. Furthermore, the results supported metaheuristic optimization's usefulness in increasing the precision and recall rates of YOLO algorithms, as well as its relevance contribution to enhancing object recognition tasks in remote sensing imaging, opening up a practical route for applications in a variety of fields.

4.1 Training and Validation Performance

Using 50 epochs for the training with an optimized modest configuration YOLOv8 model that was achieved via Google Colab GPU acceleration and Roboflow dataset preprocessing. The hyperparameters for the training have a batch size of 16, initial learning rate of 0.001 and image resolution of 640×640 pixels

In achieving a vigorous learning, mosaic augmentation, flipping, and random scaling adaptive data augmentation techniques were used during preprocessing to enhance generalization across a variety of infrared conditions. Fig. 3 presents the training and validation loss curves for box loss, classification loss, and distribution focal loss (DFL) obtained after 50 epochs. The steady lowering of these loss mechanisms validates both stable convergence and efficient feature learning. The DFL curve shows that the model efficiently diminished localization uncertainty, whereas the classification and box losses stabilized after approximately the 25th epoch, corroborative that overfitting was effectively avoided. The validation loss curves exhibited a similar trend, representing high generalization performance. The adding of a learning rate warm-up phase and a cosine decay schedule permitted for a smoother optimization trajectory and speedier convergence in the early epochs.

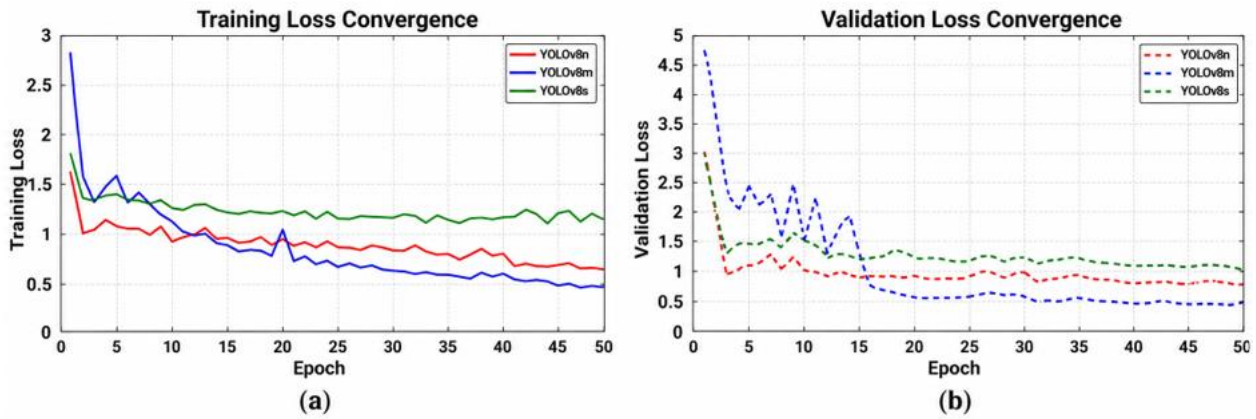


Fig. 3 (a) Training; and (b) Validation loss convergence of the YOLOv8 model across 50 epochs. The curves indicate consistent optimization behavior, stable learning, and effective generalization without signs of overfitting

The research outcome reveals that the optimized YOLOv8 architecture effectively captured both spatial (object position and shape) and semantic (object category) information in infrared vision. This stability emphasizes the strength of CSPDarknet backbone and YOLOv8's anchor-free head streamline feature propagation and decrease gradient fragmentation

4.2 Quantitative Evaluation

The optimized YOLOv8 model exhibited an impressive performance metrics with precision (0.999), mAP@0.5 (0.937), recall (0.929), and mAP@0.5-0.95 (0.70). this was achieved using a held-out test subset of the OTCBVS Benchmark Dataset. Table 1, show these performance metrics of the optimized YOLOv8 model. These findings further show the model's capability for classification and very accurate localization under diverse operational conditions such as motion blur, temperature fluctuations, and background clutter.

Table 1 Performance metrics of the optimized YOLOv8 model on the OTCBVS Benchmark Dataset

Metric	Value
Precision	0.999
Recall	0.929
mAP (0.5)	0.937
mAP (0.5–0.95)	0.70

A precise scores value of close to 1.0 shows that the model produces comparatively few false alarms, which is vital for security applications. The recall score of 0.929 indicates that the intrusions were precisely predictable. This also affect those in low-contrast and thermally saturated conditions. The mean average precision (mAP) values represent balanced performance over numerous IoU thresholds, demonstrating accurate object identification at variable distance and sizes. These data revealed that the model meets the target of high accuracy and robust adaptability necessary for real-world intrusion detection systems

4.3 Detection Visualization

Qualitative visualization was employed to examined the model performance, in addition to the model quantitative evaluation. Fig. 4 and 5 display YOLOv8's bounding-box predictions for the validation and training datasets, respectively. With Fig. 4 describing the validation predictions with a high-confidence detection of 0.9 human targets in demanding heat situations. The bounding boxes are precisely associated with the target regions, representing the YOLOv8's detecting head efficiently achieves spatial variance and poor contrast in infrared data.

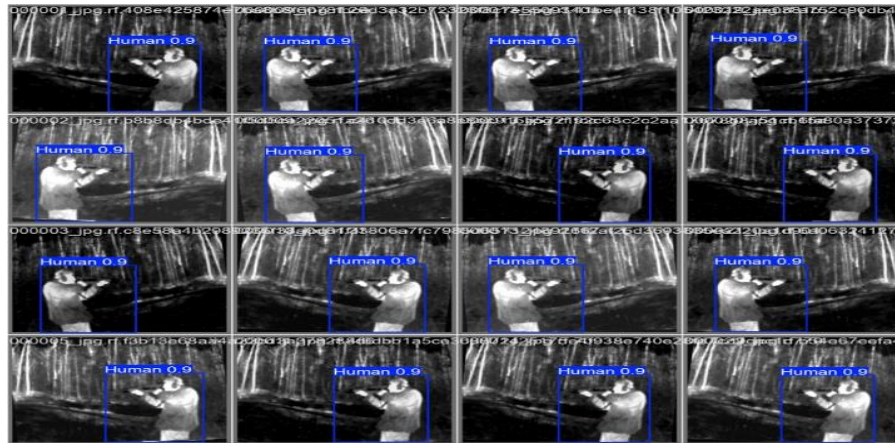


Fig. 4 Results from the Validation detection of the optimized YOLOv8 model on infrared images using the OTCBVS Benchmark Dataset

Fig. 5 displays the training findings producing consistent localization precision over numerous samples, even in crowded or partially obstructed surroundings. This uniformity validates that the model's learnt characteristics are durable and transferable, allowing it to effectively distinguish between human and non-human heat signatures

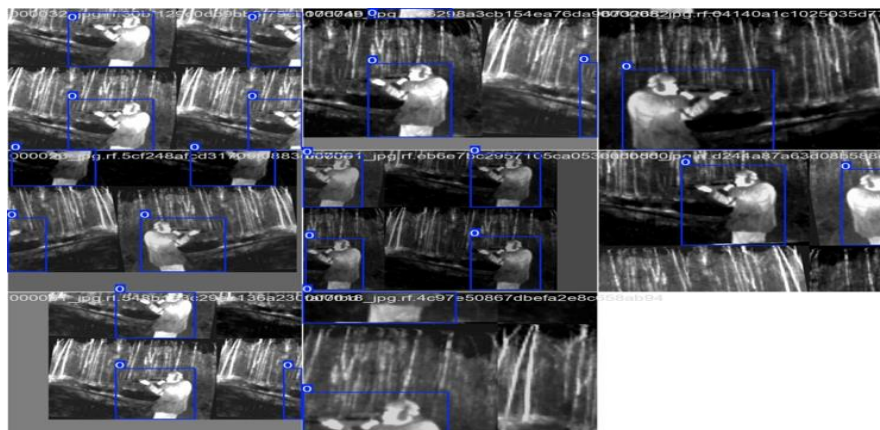


Fig. 5 Training image detections produced by the YOLOv8 model. The bounding boxes and class labels show accurate spatial feature extraction during the model optimization process

The visual outputs balance the quantitative metrics by corroborative that YOLOv8 generalizes efficiently to both training and validation datasets while upholding precise boundary detection

4.4 Discussion

The developed YOLOv8 model performed better than the earlier work using older YOLO versions for infrared and multimodal detection tasks.

Li et al. [9] used YOLOv5 for multimodal pedestrian identification. The finding show a mean average precision (mAP@0.5) of 0.89 using, while [11] obtained 0.91 using a YOLOv7 fusion of thermal and visual spectra. In comparison, the current using YOLOv8 achieved 0.937 mAP@0.5 employing infrared-only imagery, stressing the architectural advances of YOLOv8's anchor-free detection method, decoupled head, and strengthened CSP backbone.

The insertion of Roboflow-based augmentation, through the simulated fluctuations, rotation, and thermal noise, improved the model's robustness to real-world variations. Furthermore, Google Colab's GPU acceleration allowed effective batch training and gradient computing, reducing convergence time while retaining mathematical stability.

Overall, the results revealed that the proposed YOLOv8 model has higher precision, strong generalization and low computational cost. This makes it an optimal model for real-time infrared intrusion detection applications in defense monitoring, smart surveillance, and perimeter security systems. The findings also indicate YOLOv8's potential as a platform for future study into multi-spectral fusion and temporal tracking in thermal vision systems

5. Conclusion and Future Work

This paper presented an optimized YOLOv8-based infrared intrusion detection and classification model trained and assessed via the OTCBVS Benchmark Dataset Collection. The study addressed a vital gap in the current literature, where YOLOv8 had not been thoroughly optimized or validated for infrared-only intrusion detection tasks. By integrating dataset preprocessing and augmentation through Roboflow, cloud-based training on Google Colab Pro, and performance tracking through the Ultralytics YOLOv8 framework, the model attained superior detection precision and generalization.

The quantitative assessment produced notable findings, with a precision of 0.999, recall of 0.929, mAP50 of 0.937, and mAP50-95 of 0.70. The constant convergence of training and validation losses displayed the model's stability and resistance to overfitting. Qualitative investigation established the model's capacity to efficiently locate and identify incursion targets, such as individuals and vehicles, even under difficult infrared imaging settings.

These outcomes validate that YOLOv8's architecture, which includes a Cross-Stage Partial (CSP) backbone and an anchor-free detection head, significantly enhances thermal image detection reliability when compared to previous YOLO variations. The proposed architecture provides a replicable and economical baseline for infrared-based intelligent monitoring in power transmission and other critical infrastructure domains.

Future study should focus on the following:

1. Incorporating additional public and custom infrared datasets to further assess model scalability across various environments and imaging devices.
2. Integrating complementary visible-spectrum or radar data to improve detection accuracy under extreme weather or occlusion conditions.
3. Implementing and testing the optimized YOLOv8 model on embedded edge devices (e.g., Jetson Nano, Raspberry Pi 4) for on-site, low-latency intrusion detection.

Acknowledgement

The authors express their gratitude to the Tertiary Education Trust Fund (TETFUND) for funding this study through the Institution-Based Research (IBR) grant. This has permitted the positive accomplishment of this study.

Conflict of Interest

All authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors authorize their contributions to the paper as follows: **research idea and design:** Umar Abubakar Saleh, Mazadu A. A., Solomon Zakwoi Iliya; **data collection:** Mazadu A. A., Aliu Sala Saliu, Yakubu Yunusa; and **analysis and interpretation of results:** Umar Abubakar Saleh, Solomon Zakwoi Iliya, Mustapha Datti Mohammed, Aliyu Isiyaku; **draft manuscript preparation:** Hayatudeen Ogiri, Bilyaminu Usman; **model development and training:** Umar Abubakar Saleh, Aliu Sala Saliu, Yakubu Yunusa; **result validation:** Mustapha Datti Mohammed, Aliyu Isiyaku, Hayatudeen Ogiri; **writing, supervision, and editing:** Siti Amely Jumat. The results were evaluated by all authors, who then approved the final version of the manuscript.

References

- [1] B. Yang, S. Yang, X. Zhu, M. Qi, H. Li, Z. Lv, X. Cheng, and F. Wang, "Computer vision technology for monitoring of indoor and outdoor environments and HVAC equipment: A review," *Sensors*, vol. 23, no. 13, Art. no. 6186, 2023, doi: 10.3390/s23136186.
- [2] F. Köylü, "A Comparative Analysis of YOLO-Based Traffic Sign Detections With a Novel Turkish Traffic Sign Dataset," in *IEEE Access*, vol. 14, pp. 7744-7763, 2026, doi: 10.1109/ACCESS.2026.3652842.
- [3] X. Lv, S. Gao, G. Miao and Z. Chen, "Relay Protection Setting Sheet Detection and Recognition Approach Based on YOLOv8-CRNN-CTC," in *IEEE Access*, vol. 14, pp. 19961-19969, 2026, doi: 10.1109/ACCESS.2026.3654663.
- [4] M. Song, Y. Lan, X. Han, *et al.*, "YOLOv8-CTB: A lightweight and real-time cotton bud detection under field conditions with edge deployment," *Cluster Computing*, vol. 29, Art. no. 121, 2026, doi: 10.1007/s10586-026-05943-y.
- [5] J. Kaur and L. K. Singh, "Real-time object detection using YOLO variants with nature-inspired algorithm," *Iran Journal of Computer Science*, vol. 9, Art. no. 22, 2026, doi: 10.1007/s42044-025-00357-2.

- [6] J.-R. Riba, "Application of image sensors to detect and locate electrical discharges: A review," *Sensors*, vol. 22, no. 15, Art. no. 5886, 2022, doi: 10.3390/s22155886.
- [7] Malithi, HMNS Kumari, and HMLS Kumari, "Underwater Waste Detection Using Deep Learning: A Performance Comparison of YOLOv7–10 and Faster R-CNN", *IJRC*, vol. 5, no. 1, pp. 01–13, Jan. 2026.
- [8] Mohammad Fayazi, Alireza Saffarian, Mahmood Joorabian, Mehdi Monadi, Elhanashi, A., Saponara, S. (2026). Applications of Object Detection and Localization. In: Deep Learning for Object Detection and Localization. Springer, Singapore. https://doi.org/10.1007/978-981-95-4846-0_5.
- [9] X. Li, H. Zhang, and Q. Chen, "Multi-modal pedestrian detection under low illumination using YOLOv5 and adaptive data fusion," *Infrared Physics & Technology*, vol. 132, Art. no. 104622, 2023.
- [10] C. Jiang, H. Ren, X. Ye, J. Zhu, H. Zeng, Y. Nan, and H. Huo, "Object detection from UAV thermal infrared images using YOLO models," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, Art. no. 102912, 2022.
- [11] P. Mudbhatkal, M. Juhola, M. Asikainen, and S. K. Patel, "Deep learning models for straddle carriers: Predictive maintenance," *Array*, vol. 29, Art. no. 100706, 2026, doi: 10.1016/j.array.2026.100706.
- [12] B. Ding, "A lightweight real-time infrared object detection model based on YOLOv8," *Drones*, vol. 8, no. 9, Art. no. 479, 2024.
- [13] M. Rizk and I. Bayad, "Human detection in thermal images using YOLOv8 for search and rescue missions," in *Proc. 7th Int. Conf. Advances in Biomedical Engineering (ICABME)*, 2024, pp. 1–4.
- [14] Xie, H., Yuan, B., Hu, C., Gao, Y., Wang, F., Wang, C., Chu, P. (2024). "ST-YOLO: A defect detection method for photovoltaic modules based on infrared thermal imaging and YOLOv8s." *PLOS ONE*, Vol. 19 (12), e0310742.
- [15] J. Yin, J. Zhao, and X. Zhang, "MS-YOLO11: A multidimensional synergistic You Only Look Once 11 algorithm for pulmonary nodule detection," *PeerJ Computer Science*, vol. 12, Art. no. e3473, 2026, doi: 10.7717/peerj-cs.3473.
- [16] V. Pann, K.-S. Kwon, B. Kim, D.-H. Jang, J. Kim, and J.-B. Kim, "Robustness of CNN-based model assessment for pig vocalization classification across diverse acoustic environments," *Computers and Electronics in Agriculture*, vol. 240, Art. no. 111181, 2026, doi: 10.1016/j.compag.2025.111181.
- [17] L. N. Byenkya, R. Nakibuule, D. Taremwa, *et al.*, "A hybrid YOLOv5s-Faster R-CNN architecture for object detection in complex road scenes," *Research Square*, preprint, ver. 1, Jan. 21, 2026. doi: 10.21203/rs.3.rs-8559050/v1.