Journal of Engineering, Project, and Production Management
2026, 16(3), 2025-221

Improved YOLO-Based Abnormal Behavior Recognition System for Civil Aviation Security

Fengwei Chen[1] and Tianyou Wu[2]
[1] Professional Instructor, Airport Management College, Guangzhou Civil Aviation College, Guangzhou 510403, China
[2] Instructor, Airport Management College, Guangzhou Civil Aviation College, Guangzhou 510403, China, E-mail: tianyouwuty@126.com (corresponding author).

Project Management
Received October 11, 2025; revised November 26, 2025; accepted December 11, 2025
Available online April 8, 2026

Abstract: As the global aviation industry continues to expand rapidly, the nature of security threats has grown more intricate. Under the influence of the international security situation, aviation screening systems have emerged as a key defense layer. In response, this paper puts forward a composite model that integrates multiple algorithms. First, the model combines a focal loss function, which is suitable for detecting dense targets, with a target detection algorithm. Following that, the particle swarm optimization algorithm parameters are fine-tuned using a Bayesian optimization approach. Finally, the model integrates the focal loss-based hybrid target detection algorithm with the optimized particle swarm optimization algorithm to construct a behavior recognition model. Experimental results show that this model's loss curve is smoother and has lower loss than the three traditional models. On the University of Central Florida dataset and the University of Rochester dataset, the model achieves mean average precision values of 97.675% and 98.246%, respectively, outperforming the other three models. In addition, the model reaches a maximum accuracy of 98.652% and a maximum FPS of 140 on two aviation screening datasets. These results demonstrate that the model offers both high accuracy and real-time performance, indicating its potential to support the development of aviation screening systems and effectively meet the demands of security screening.

Keywords: YOLO, abnormal behavior detection, aviation security, deep learning, real-time recognition.

DOI 10.32738/JEPPM-2025-221

1. Introduction

With the rapid growth of the aviation industry, threats to aviation security have become increasingly complex and diverse. As a critical safeguard, aviation screening relies on the precise and efficient recognition of abnormal human behavior (Walter et al., 2024). However, current technologies struggle with complex scenarios due to challenges such as difficulty detecting small targets, imbalanced class distribution, and insufficient parameter optimization. These issues result in low recognition accuracy for sparse abnormal behavior samples and fail to meet the high-security demands of civil aviation (Bello and Oladipo, 2024; Mathew and Mahesh, 2022). Therefore, developing an abnormal behavior recognition system that balances precise accuracy, and real-time performance is essential for improving screening efficiency and strengthening aviation safety defenses (Mittal et al., 2023). The You Only Look Once (YOLO) algorithm, a single-stage detection technique, reformulates the detection task as a regression problem. By performing one forward pass through a neural network, it simultaneously predicts the location and category of objects, significantly improving detection speed while maintaining high accuracy (Chitraningrum et al., 2024). Particle Swarm Optimization (PSO) is a global optimization technique modeled after the collective behavior observed in biological swarms. It simulates the collaboration and information sharing observed in bird flocks during foraging, using iterative search to find optimal solutions. PSO is simple to implement, converges quickly, and offers strong global search capabilities. Combining YOLO and PSO allows the strengths of both methods to complement each other. This paper introduces a personnel abnormal behavior recognition system based on an improved YOLO and enhanced PSO algorithm. The model incorporates multiple algorithms in a hybrid structure and focuses on recognizing small targets and dense crowd behaviors. The objective is to boost both the precision and real-time efficiency of recognition systems, advancing safety standards in aviation.

2. Related Works

The YOLO algorithm has been widely applied in industrial, security, and intelligent transportation fields due to its advantages in speed and real-time performance. Many researchers at home and abroad have carried out in-depth studies on this algorithm. Bhavana (2024) proposed an improved real-time object detection algorithm, i.e., You Only Look Once version 8 (YOLOv8), for real-time detection of road potholes. This algorithm utilizes edge segmentation to improve YOLOv8, preprocesses the graphics, and finally uses the improved YOLOv8 for real-time detection (Bhavana et al., 2024). Lee and Hwang (2022) proposed a target detection model based on adaptive frame control to solve the problem of low real-time performance. They first applied the target detection algorithm to an artificial intelligence system, then optimized it through adaptive frame control, and finally constructed a hybrid model to detect data and verify its real-time performance (Lee and Hwang, 2022). Sangaiah and Lin (2024) raised a model based on a target detection algorithm and deep learning to address the lack of technology in aerial agriculture. They introduced pooling, attention, and feature extraction modules into the target detection algorithm, combined it with deep learning, and constructed a hybrid model. The model was trained on a rice leaf disease dataset and showed high performance (Sangaiah, Yu, and Lin, 2024). Dewi et al. (2024) designed a recognition model based on multiple target detection algorithms to solve the challenge of accurately detecting traffic signs in smart vehicles. The model first fused various target detection algorithms to form a new optimized method, then introduced pyramid pooling into this new method to build a recognition model (Dewi et al., 2022). Pham and Chang (2023) developed a real-time monitoring system based on a target detection algorithm to improve the low efficiency of factory operations. This system first classified data using a target detection algorithm, then detected the data with deep learning, and finally applied the system to production lines to verify its efficiency (Pham and Chang, 2023).

Human behavior recognition, as a key focus within computer vision and pattern recognition research, had not only advanced artificial intelligence toward human-like cognition but also expanded its applications in public service industrial production, and other fields, enhancing overall social intelligence. Savchenko et al. (2022) proposed a facial recognition network based on neural networks to address the difficulty of recognizing student behaviors using traditional models. The network first extracted student face sequences, then used neural networks to capture features, and finally trained and tested the network using a dataset to verify its high performance. Xu et al. (2023) put forward a recognition model based on multi-view cross-information and confidence measurement to solve the challenge of identifying all human poses. They first extracted skeletal data from a human body database, then input the data into a target detection system for classification and finally conducted simulation training using a public dataset. Singh and Kumar (2023) designed a generative adversarial network model to overcome the difficulty of detecting behaviors in diverse scenes. They first integrated conditional autoregression and amoeboid swarm optimization, then applied the optimized algorithm to the network to build a crowd recognition model. To solve the problem of feature extraction from time-series data, Challa et al. (2023) proposed a classification model based on convolutional neural networks and experimentally confirmed its effectiveness. Gkournelos et al. (2024) developed an artificial intelligence-based recognition model to enable efficient human-machine collaboration. The team first built a practical software framework, trained and tested it using an Artificial Intelligence (AI) system, and finally applied the model on an assembly line for human motion recognition.

In summary, although the YOLO algorithm had demonstrated excellent performance in real-time object detection, its accuracy remained limited when detecting small or densely packed targets. While human behavior recognition had made notable progress, it still lacked both precision and real-time capability. Therefore, this study proposed a personnel abnormal behavior recognition model for aviation screening based on an improved YOLOv8 algorithm and a global optimization algorithm. This model introduced the Focal Loss function into YOLOv8 and applied a Particle Swarm Optimization algorithm enhanced by Bayesian Optimization (BO-PSO) to optimize the YOLOv8 model. With a hybrid architecture that combines multiple algorithms and a targeted design for complex scenarios, the proposed model aimed to improve the robustness of abnormal behavior recognition systems.

3. Abnormal Behavior Recognition for Aviation Screening Based on Improved YOLO

3.1. Improved Algorithm Design Combining Focal Loss and YOLO

Deep learning has expanded behavior recognition from simple scenes to complex dynamic environments. In aviation screening, behavior recognition faces challenges, such as small targets, occlusion, weak cross-scene generalization, and difficulty in data annotation. Therefore, the system requires more accurate recognition capabilities (Ma et al., 2022). Focal Loss-YOLOv8 offers clear advantages in addressing class imbalance and focusing on hard-to-distinguish samples, such as small or blurred targets (Gao et al., 2024). Therefore, this study introduces a hybrid algorithm that combines Focal Loss and YOLOv8 into the behavior recognition system. YOLOv8 is a target detection algorithm and part of the YOLO series. Its loss function is defined in Eq. (1).

$$Loss = \lambda_{cls}\mathrm{Loss}_{cls} + \lambda_{loc}Loss_{loc} + \lambda_{conf}Loss_{conf} \quad (1)$$

In Eq. (1), $Loss_{cls}$ is classification loss, $Loss_{loc}$ is the localization loss, and $Loss_{conf}$ is the confidence loss. λ_{cls},

λ_{loc}, and λ_{obj} are weight coefficients. The localization loss function is given in Eq. (2).

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \tag{2}$$

In Eq. (2), IoU stands for the Intersection over Union, $\rho^2(b, b^{gt})$ is the Euclidean distance between the predicted and actual center points, α and v are the weight term and aspect ratio consistency measure, and L_{CIoU} is the localization loss function. The objectness function is shown in Eq. (3).

$$L_{obj} = -[y \log(p) + (1-y)\log(1-p)] \tag{3}$$

In Eq. (3), y and p represent the label for object existence and the predicted probability of object presence. Finally, YOLOv8 performs data augmentation, as described in Eq. (4).

$$I_{mosaic} = Mix(I_1, I_2, I_3, I_4) \tag{4}$$

In Eq. (4), (I_1, I_2, I_3, I_4) refers to the cropped regions from different images, and Mix denotes the proportion used for random stitching. YOLOv8 optimizes behavior recognition through feature extraction, detection head design, and loss function design. The workflow of YOLOv8 in the behavior recognition system is illustrated in Fig. 1.

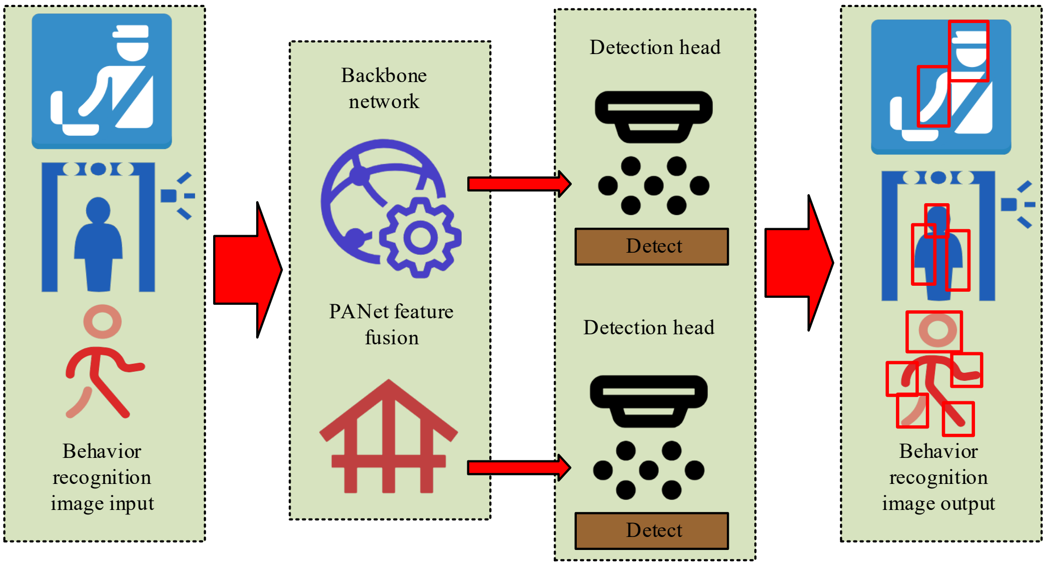

Fig. 1. Flowchart of YOLOv8 behavior recognition (Icon source from: https://icon.sucai999.com/)

In Fig. 1, the input image is first passed into the backbone of the YOLOv8 network. A key feature fusion model within the network then processes these features. Then, the Path Aggregation Network (PANet) performs further feature fusion. Finally, the processed image with recognized behavior is output. Focal Loss is a modulated cross-entropy loss designed to down-weight easy examples, enabling the model to concentrate more on challenging samples. The conventional binary cross-entropy loss function is shown in Eq. (5).

$$CE(p, y) = \begin{cases} -\log(p) & if \;\; y = 1 \\ -\log(1-p) & if \;\; y = 0 \end{cases} \tag{5}$$

In Eq. (5), p and y indicate the model's predicted probability for the positive class and the corresponding ground truth label. Focal Loss adds a modulation factor to the standard binary cross-entropy function, as shown in Eq. (6).

$$FL(p_t) = -\alpha_t (1-p_t)^{\gamma} \log(p_t) \tag{6}$$

In Eq. (6), p_t is the predicted probability, γ is the focusing parameter, and α_t is the balancing factor. In summary, Focal Loss effectively mitigates the class imbalance problem in YOLOv8 through its loss function. Therefore, the Focal Loss algorithm is introduced into the YOLOv8 framework to improve its performance. The workflow of the Focal Loss-YOLOv8 hybrid algorithm is illustrated in Fig. 2.

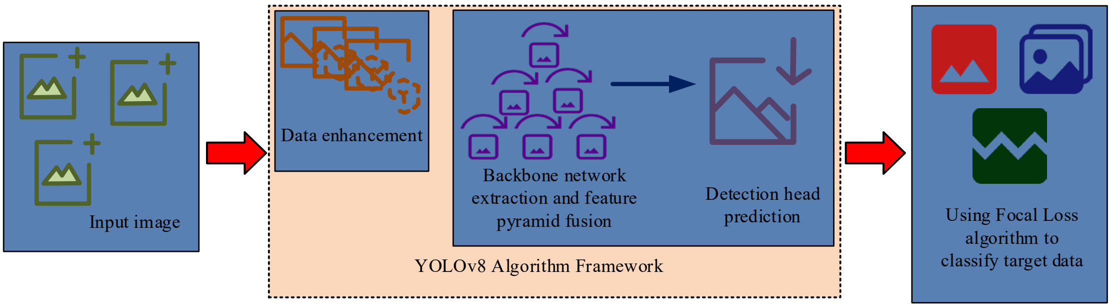

Fig. 2. Flowchart of Focal Loss-YLOv8 hybrid algorithm (Icon source from: https://icon.sucai999.com/)

As shown in Fig. 2, the Focal Loss and YOLOv8 algorithm first inputs image data into the YOLOv8 framework for data augmentation. The data are then processed by the backbone network to extract features and perform feature pyramid fusion. The detection head predicts the bounding box, objectness, and class, and finally, the data are passed to the Focal Loss layer for classification loss calculation to achieve accurate classification. The Focal Loss function for behavior recognition is defined by Eq. (7).

$$FL = \sum_{i=1}^{K} FL_i = \sum_{i=1}^{K} [-y_i \alpha_i (1-p_i)^{\gamma} \log(p_i) - (1 - y_i \alpha_i^{'} p_i^{\gamma} \log(1-p_i) \quad (7)$$

In Eq. (7), M and i represent the total number of behavior classes and the class index. p_i is the predicted probability, ranging from 0 to 1. α_i and $\alpha_i^{'}$ are the positive and negative sample balancing factors, used to balance the number of "existing behavior" samples (Bhavana et al., 2024). γ is the focusing parameter. In conclusion, Focal Loss-YOLOv8 enhances the model's ability to recognize sparse target classes. Therefore, the Focal Loss-YOLOv8 hybrid algorithm is applied to the abnormal behavior recognition system for aviation screening. The process of this algorithm within the system is shown in Fig.3.

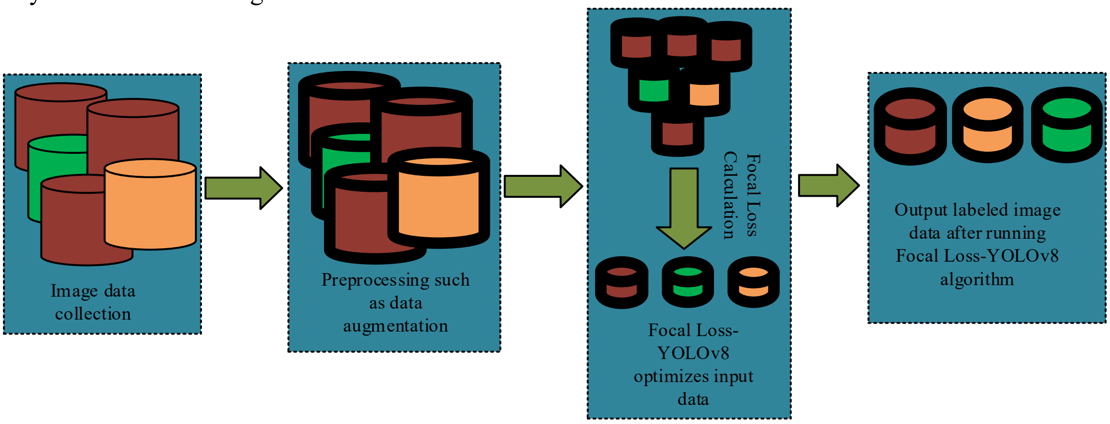

Fig. 3. Operation of Focal Loss-YOLOv8 algorithm in behavior recognition system

As shown in Fig. 3, the behavior recognition system follows a structured workflow. First, image data are collected through high-definition cameras and other equipment, covering human actions and environmental information in aviation screening scenarios. Then, these raw data are labeled. In the data preprocessing stage, augmentation removes blurred or redundant information. The high-quality data are then input into the Focal Loss-YOLOv8 algorithm layer. This layer first uses the backbone network to extract deep features, capturing edges, textures, and motion trajectories. Then, the feature pyramid fuses multi-scale feature data to enhance recognition of complex behaviors. Finally, the Focal Loss function calculates classification errors and dynamically optimizes model parameters to ensure accurate classification of fine-grained behaviors. The classification results are then output to complete the behavior recognition process.

3.2. Behavior Recognition Model Integrating BO-PSO and Improved YOLO Algorithm

Although the Focal Loss-YOLOv8 algorithm shows advantages in addressing the imbalance disparity in the distribution of positive and negative samples, it still faces limitations such as parameter sensitivity and increased computational complexity (Kumar et al., 2023). Therefore, this study proposes a hybrid model that integrates BO-PSO and the improved YOLO algorithm. In this model, BO narrows the search range for parameter combinations, while PSO performs a fine-grained search. Together, they optimize the Focal Loss-YOLOv8 algorithm. PSO is a population-based global optimization algorithm that mimics how bird flocks search for food in order to locate the optimal solution (Rajamohanan and Latha, 2023). The velocity update equation for PSO is given in Eq. (8).

$$v_i^d(t+1) = w \cdot v_i^d(t) + c_1 \cdot r_1^d(t) \cdot (pbest_i^d - x_i^d(t)) + c_2 \cdot r_2^d(t) \cdot (gbest^d - x_i^d(t)) \quad (8)$$

In Eq. (8), $v_i^d(t)$ denotes the velocity of the t-th particle. w is the inertia weight. c_1 and c_2 are learning factors. $pbest_i^d$ and $gbest^d$ represent the best individual and global position, respectively. $x_i^d(t)$ represents the position of the i-th particle. Although PSO has global search ability, it often requires many particles, resulting in high computational cost. BO, as a global optimization algorithm, is based on probabilistic models and learns the probability distribution of its objective function to find the optimal solution, which helps reduce computational load. The Gaussian process used in BO is defined in Eq. (9).

$$\begin{cases} f(x) \sim GP(m(x), k(x, x')) \\ m(x) = E[f(x)] \\ k(x, x') = E[f(x) - m(x)]E[f(x') - m(x') \end{cases} \tag{9}$$

In Eq. (9), $f(x)$, $m(x)$, and $k(x,x')$ represent the objective, the mean, and the kernel function, respectively. GP stands for the Gaussian Process. The covariance function is expressed in Eq.(10).

$$k(x, x') = \sigma_f^2 \exp(-\frac{1}{2l^2}\left\|x - x'\right\|^2) \tag{10}$$

In Eq. (10), σ_f^2 denotes the signal variance, l is the length of scale, and $\left\|x - x'\right\|^2$ is the squared Euclidean distance. After defining the prediction set, the mean and variance of the predictive distribution are calculated as shown in Eq. (11).

$$\begin{cases} \mu(x_t) = k(x_t, x)^T [k(x, x) + \delta_n I]^{-1} y \\ \delta^2(x_t) = k(x_t, x_t) - k(x_t, x)^T [k(x, x) + \delta_n I]^{-1} k(x_t, x) \end{cases} \tag{11}$$

In Eq. (11), $\mu(x_t)$ and $\delta^2(x_t)$ are the mean and variance of the predictive distribution. x_t represents the prediction set. I is the identity matrix and y is the target output corresponding to the input x. In summary, BO can optimize PSO parameters using the Gaussian Process. The workflow of the BO-PSO algorithm is illustrated in Fig. 4.

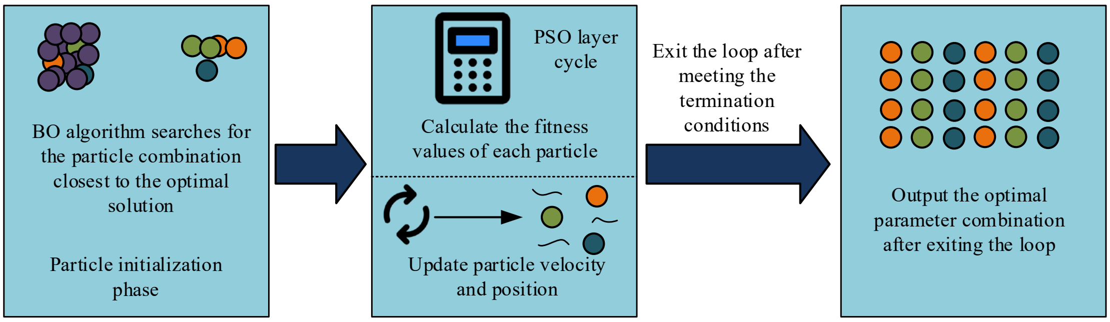

Fig. 4. Operation flowchart of BO-PSO algorithm

As shown in Fig. 4, BO first generates a set of initial particles that are positioned nearer to the optimal solution. This set is then passed into PSO as the starting point. Since PSO enters a loop to update particle velocities and positions, the BO-generated near-optimal particles help avoid excessive computational cost caused by random initialization. Finally, PSO outputs the optimal parameter combination. The calculation expression for BO-PSO's optimal solution search is shown in Eq. (12).

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \tag{12}$$

In Eq. (12), $x_i^d(t+1)$ is the position. The number of iterations $t+1$ decreases as BO accelerates the search for the optimal solution. Therefore, this study applies the BO-PSO optimization algorithm to the Focal Loss-YOLOv8 algorithm to build a composite model for abnormal behavior recognition. The recognition process of the BO-PSO-Focal Loss-YOLOv8 composite model in aviation screening is illustrated in Fig. 5.

As shown in Fig. 5, the model first improves the base YOLOv8 algorithm by replacing its original classification loss function with the Focal Loss function. In the BO-PSO optimization layer, the hybrid strategy combines BO and PSO. BO uses the Gaussian Process model to accurately determine the key parameter combinations of PSO, which significantly accelerates the convergence rate and avoids local optima. The PSO algorithm, optimized by BO, performs global tuning of YOLOv8 hyperparameters, quickly locating optimal solutions in high-dimensional parameter space. After these steps, the model processes the collected and augmented behavior image data through multi-scale feature extraction in the backbone network, hierarchical feature fusion in the feature pyramid network, and precise classification with the Focal Loss function to complete abnormal behavior recognition.

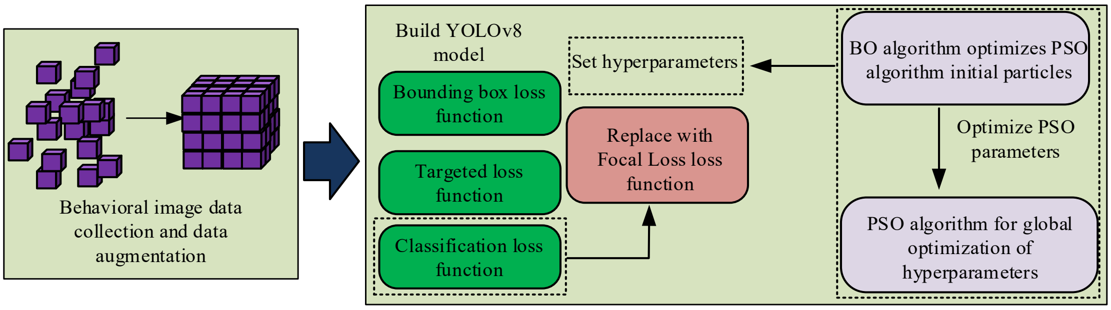

Fig. 5. BO-PSO-Focal Loss-YOLOv8 composite model recognition process diagram

4. Performance of the Recognition System based on Focal Loss-YOLOv8 and BO-PSO

4.1. Validation of the Focal Loss-YOLOv8 Algorithm

To assess the classification performance of Focal Loss-YOLOv8 during the recognition process, the study compared its performance with three other algorithms: GIoU Loss-YOLOv8, CE-YOLOv8, and Varifocal-YOLOv7. The experiment was conducted on a Windows 11 system with an NVIDIA RTX 5090 GPU and an i7-12700 CPU. The datasets used were the University of Central Florida (UCF101) and Human Motion Database 51 (HMDB51). Three indicators—Average Precision for Small Objects (AP-S), Average Precision for Medium Objects (AP-M), and mean Average Precision (mAP)—were selected for comparison. The scatter plots after classification by the four algorithms are shown in Fig. 6.

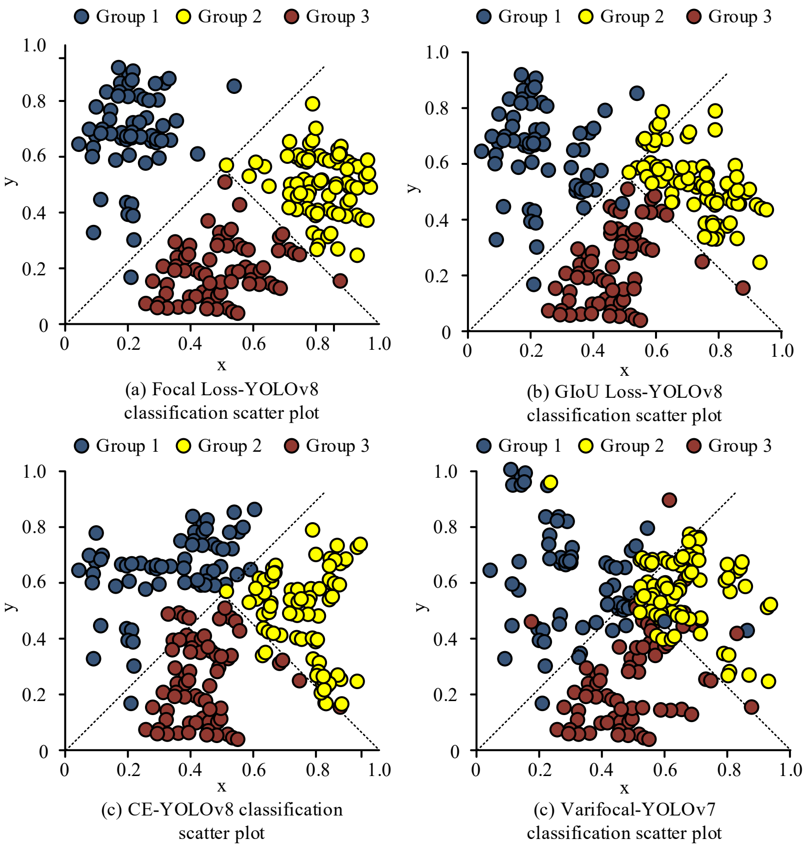

Fig. 6. Classification scatter plots of four algorithms

As shown in Fig. 6(a), the scatter points of Focal Loss-YOLOv8 mostly fell within the correct regions. According to Fig.6(b), (c), and (d), GIoU Loss-YOLOv8, CE-YOLOv8, and Varifocal-YOLOv7 all showed varying degrees of category confusion, with Varifocal-YOLOv7 performing the worst due to the lack of a suitable loss function for accurate classification. These results demonstrate that Focal Loss-YOLOv8 achieved the highest classification accuracy. To further evaluate classification accuracy, the study tested images of 32×32 and 64×64 pixels using all four algorithms. The resulting Precision-Recall (PR) curves are presented in Fig. 7.

As shown in Fig. 7(a), for 32×32 images, the Area Under the Curve (AUC) of Focal Loss-YOLOv8 reached 0.984, while the AUC values of GIoU Loss-YOLOv8, CE-YOLOv8, and Varifocal-YOLOv7 were 0.886, 0.742, and 0.686, respectively. Fig.7(b) showed that for 64×64 images, the AUC of Focal Loss-YOLOv8 was 0.962, higher than all other models. This demonstrated that Focal Loss-YOLOv8 achieved accurate recognition by effectively leveraging an improved loss function. The results indicated that Focal Loss-YOLOv8 outperformed the other three algorithms in both small and medium object classification tasks. To further evaluate its accuracy and generalizability, the study tested the four algorithms on mAP, AP-S, and AP-M. The results are listed in Table 1.

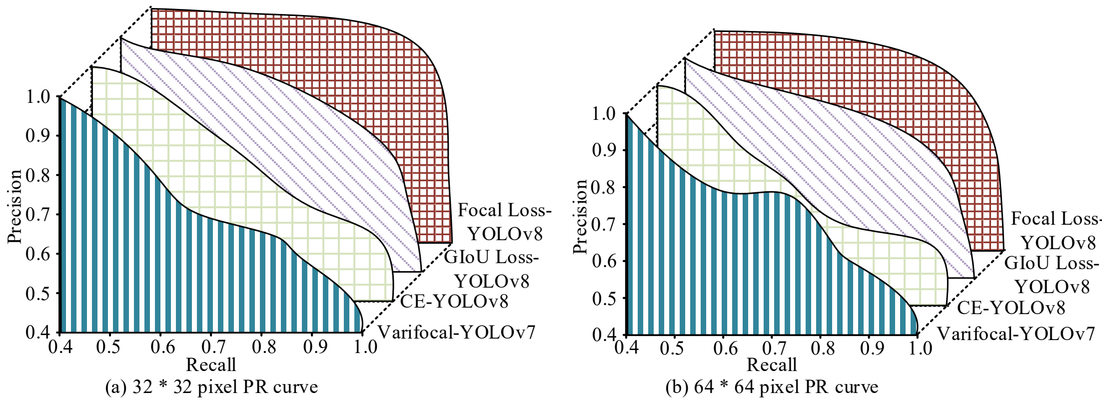

Fig. 7. The PR curves of all four algorithms under different objectives

Table 1. Test results of various indicators for the four algorithms

	Algorithm	Evaluation indicators			
UCF101 data set		mAP (%)	AP-S (%)	AP-M (%)	Parameter quantity
	Focal Loss-YOLOv8	96.254%	97.675%	98.634%	22.5
	GIoU Loss-YOLOv8	84.125%	89.215%	88.627%	22.5
	CE-YOLOv8	78.321%	77.024%	79.325%	22.5
	Varifocal-YOLOv7	70.463%	68.471%	75.694%	22.5
	Algorithm	Evaluation indicators			
HMDB51 data set		mAP (%)	AP-S (%)	AP-M (%)	Parameter quantity
	Focal Loss-YOLOv8	98.246%	97.789%	99.034%	22.5
	GIoU Loss-YOLOv8	84.255%	88.138%	89.462%	22.5
	CE-YOLOv8	79.256%	78.524%	80.025%	22.5
	Varifocal-YOLOv7	72.168%	70.351%	77.245%	22.5

According to Table 1, Focal Loss-YOLOv8 achieved a mAP of 96.254%, AP-S of 97.675%, and AP-M of 98.634%, all significantly higher than the other three algorithms. These results showed that Focal Loss effectively addressed the lack of attention to small and blurred objects in traditional loss functions, while retaining the efficient feature extraction architecture of YOLOv8. In summary, the key advantage of Focal Loss-YOLOv8 lies in its improved handling of class imbalance, allowing the model to focus more on hard examples and thereby significantly improving detection accuracy for small objects. At the same time, it preserved the fast detection speed of YOLOv8, making it suitable for complex scenarios in aviation screening.

4.2. Verification of the Behavior Recognition Model in Aviation Screening

After verifying the performance of the improved YOLO algorithm, the study introduced the BO-PSO optimization method into the improved YOLO framework to build the BO-PSO-Focal Loss-YOLOv8 model for abnormal behavior recognition. The study conducted performance comparison experiments between this final model and three traditional behavior recognition models: Faster R-CNN-Focal Loss, GA-YOLOv8, and EfficientDet-GHM Loss. The experiments were configured with a Frames Per Second (FPS) value of 200 and input image sizes of 36×36 and 90×90 pixels. The loss curves from the four models are shown in Fig. 8.

In Fig. 8(a), the BO-PSO-Focal Loss-YOLOv8 model converged after only 10 iterations with a loss of 0.08, remarkably close to zero. In contrast, the Faster R-CNN-Focal Loss, EfficientDet-GHM Loss, and GA-YOLOv8 models required 20, 40, and 50 iterations to converge, with loss values of 0.14, 0.16, and 0.21. These results confirmed that the final recognition model achieved faster convergence and lower loss. Fig. 8(b) further demonstrated that the model performed better on medium objects. To assess behavior recognition accuracy, the four models were tested and trained on the UCF101 and HMDB51 datasets. The results are shown in Fig. 9.

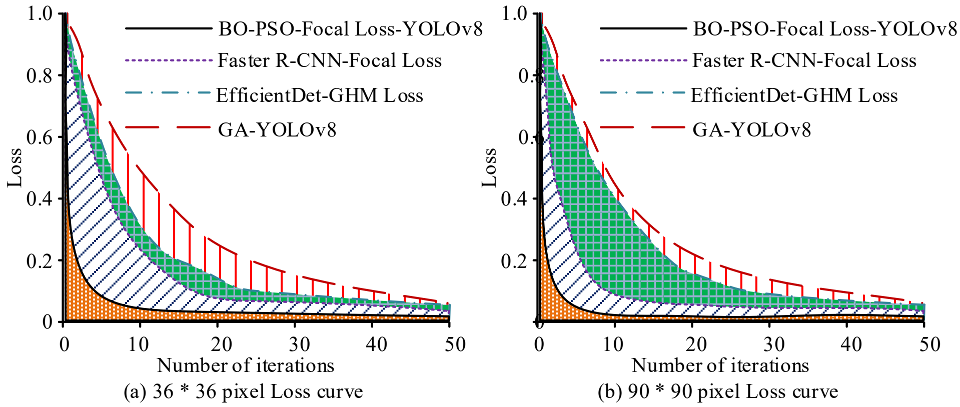

Fig. 8. Loss curve for processing two types of image data

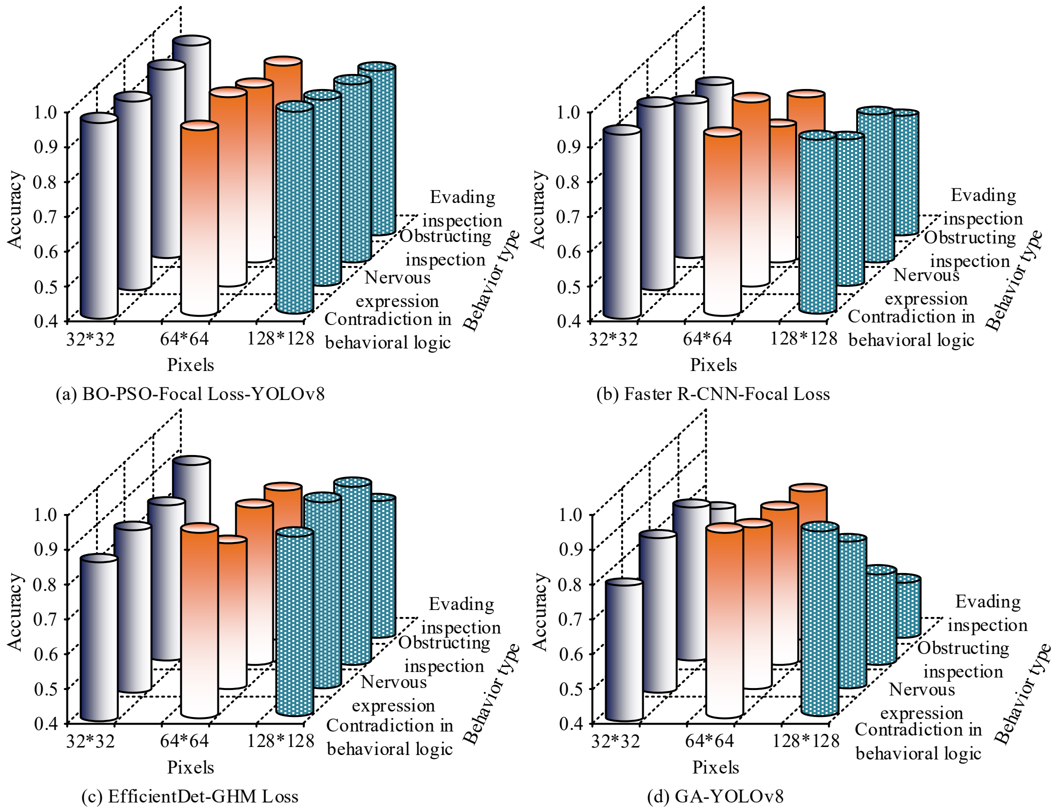

Fig. 9. Experimental results on the accuracy of behavior recognition

According to Figure 9(a), the BO-PSO-Focal Loss-YOLOv8 model consistently achieved a recognition accuracy of around 0.984 across different image resolutions and behavior details, approaching the ideal value of 1.000. Figures 9(b), (c), and (d) show that the other three models exhibit lower recognition accuracy, particularly in identifying fine-grained behaviors. These results confirm that the proposed composite model achieved superior recognition performance in complex environments. To further verify classification performance, the study tested all four models using diverse behavior categories. The results are shown in Fig. 10.

Fig. 10(a) shows that the BO-PSO-Focal Loss-YOLOv8 model maintained a high classification accuracy of 0.967 even across 100 behavior categories, while the other three models performed significantly worse. Fig. 10(b) demonstrates that the proposed model also outperforms others in classifying large-scale objects. These results confirmed that the BO-PSO-Focal Loss-YOLOv8 model delivered superior classification performance in both test scenarios. To thoroughly evaluate the model's effectiveness, this study employed two aviation screening datasets and tested the four models on four different metrics: mAP, accuracy, recall, and FPS. The results are presented in Table 2.

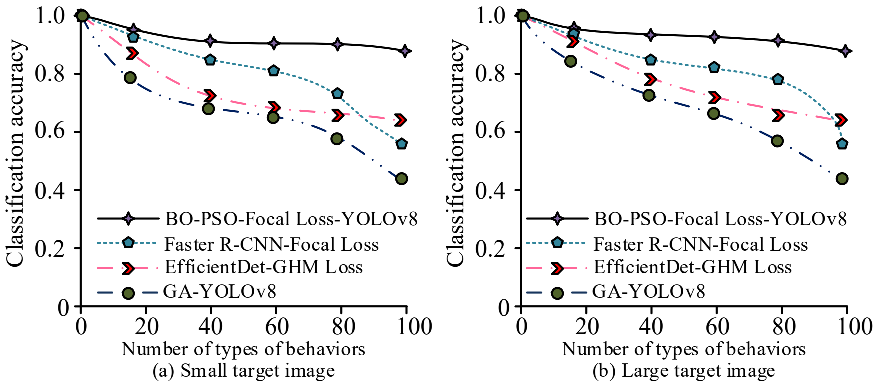

Fig. 10. Experimental results of classification performance of recognition model

According to the information in Table 2, it can be seen that on the two civil aviation security inspection datasets, the BO-PSO Focal Loss-YOLOv8 final personnel abnormal behavior recognition model performs outstandingly in all core indicators, with the highest mAP reaching 98.145%, peak accuracy of 98.652%, and recall rate of 99.354%, which is far superior to the other three comparison models. This significant advantage stems from the dual-layer Focal Loss optimization mechanism, which effectively alleviates the class imbalance problem caused by sparse abnormal behavior samples in security check scenarios. Furthermore, the BO-PSO algorithm enhances the model's adaptability to complex scenarios through dynamic parameter optimization. This model is significantly better than the two-stage detector Faster R-CNN in FPS compared to GA-YOLOv8. Although the number of parameters is similar, the model optimizes the network's hyperparameters through the BO-PSO algorithm, achieving the highest detection accuracy while maintaining a high throughput of 140 FPS. In contrast, although Faster R-CNN has decent accuracy, its lower FPS cannot meet the demanding real-time requirements of civil aviation security checks. From this, the BO-PSO-Focal Loss-YOLOv8 model balances high precision and strong robustness while ensuring high real-time performance and can accurately capture various abnormal behaviors in civil aviation security checks. In summary, in real-time video security detection, 98% mAP means that the system has almost human-level precision insight. It can reliably identify abnormal behaviors such as retrograde, stranded, and climbing from massive video streams at a speed of 140 frames per second in dense crowds and complex lighting environments at airports, completely transforming the security check mode from traditional post verification to in-process intervention. This precision indicator enables security personnel to free themselves from passive surveillance and focus on precise response, thus building an all-weather, exceptionally reliable, and proactive intelligent security defense line. This result validates the high accuracy of the BO-PSO Focal Loss-YOLOv8's final abnormal behavior recognition model in complex environments. Finally, to verify the applicability and robustness of the proposed model in real civil aviation security check environments, the BO-PSO-Focal Loss-YOLOv8 model was further deployed in the security check area of the T2 terminal of an international airport. When the model is deployed in the actual airport, it uses the edge computing architecture to conduct real-time analysis on the edge servers deployed in key locations such as security check areas through direct access to the existing CCTV network video stream of the airport. The model will provide real-time feedback on detected abnormal behavior to the visual interface of the security operator and provide prompts through highlighted boxes, sound, and light alarms to assist in rapid response. To ensure privacy compliance, the system design adopts localized processing. The original video stream is immediately discarded after being analyzed at the edge, and only alarm events and related clips are uploaded. Real-time anonymization of facial and other biometric features can be performed to improve security inspection efficiency while strictly protecting personal privacy. The testing environment includes multiple lighting conditions and scenarios with high crowd density. The mAP of the BO-PSO Focal Loss-YOLOv8 model was tested on-site in the aforementioned environment, and the test results are shown in Table 3.

From a regulatory perspective, this model requires a careful balance between aviation safety regulations and data privacy ethics. In terms of cost, the model needs to meet strict aviation safety compliance requirements, and the initial investment includes hardware upgrades and system certification costs. At the same time, to comply with data ethics and privacy protection regulations, it is necessary to integrate real-time anonymization processing mechanisms at the technical level and establish data minimization principles in the process. In terms of benefits, the model can significantly enhance the proactive identification and response capabilities to safety threats required by Aviation Law. This effectively prevents safety accidents through real-time abnormal behavior detection, reduces the risk of human negligence, and strengthens the safety subject responsibility of aviation operators. Its core values in ensuring public safety, while its embedded considerations for privacy protection, help to gain public trust and regulatory recognition. Doing this improves security levels, respects individual rights, and achieves the unity of security benefits and compliance.

5. Conclusion

Aviation screening serves as a critical barrier ensuring aviation safety, and the accurate identification of abnormal human behaviors is essential to this objective. However, challenges such as mediocre performance in small object recognition, imbalanced class distribution, and insufficient parameter optimization limited the effectiveness of existing recognition techniques under complex scenarios. To enhance the performance of civil aviation security systems, this study put forward a model named BO-PSO-Focal Loss-YOLOv8. This model first combined Focal Loss with YOLOv8 to dynamically adjust sample weights, thereby improving the class imbalance issues while focusing on small and blurred samples that were difficult to distinguish. Then, it incorporated the BO-PSO algorithm to optimize model parameters. In this framework, BO reduced the search space, while PSO conducted fine-tuning, which helped lower computational complexity and accelerate convergence. Experimental results showed that on the UCF101 and HMDB51 datasets, the model achieved an mAP of 96.254%, with AP-S and AP-M reaching 97.675% and 98.634%, respectively, outperforming three traditional models. On the civil aviation security datasets, the model achieved an mAP of 98.145%, a maximum accuracy of 98.652%, a recall rate of 99.354%, and an FPS of 140, demonstrating both high precision and real-time performance. Overall, the performance improvement of the BO-PSO-Focal Loss-YOLOv8 model stems from its ingenious collaborative design. The Focal Loss function effectively solves the problem of class imbalance commonly found in monitoring scenarios by focusing on difficult samples, significantly enhancing the detection sensitivity for small targets and sparse abnormal behaviors. The BO-PSO optimization algorithm guides particle swarm search through BO, automatically finding the optimal hyperparameter combination for the model, fully unleashing the potential of YOLOv8 infrastructure. This paradigm of precise focus and intelligent optimization has fundamentally transformed intelligent monitoring from passive recording to active perception. Its high precision and real-time performance provide a reliable solution for edge deployment, making it possible to build an all-weather, low false alarm automated security system, and setting a successful example of algorithm fusion innovation for the entire industry. Although the model still has domain limitations and overfitting issues, future work will continuously improve its robustness through the application process.

Table 2. Experimental results of comprehensive indicators

	Model	Test indicators			
		mAP (%)	Accuracy (%)	Recall (%)	FPS
Civil Aviation Dataset 1	BO-PSO-Focal Loss-YOLOv8	98.145	98.652	99.354	140±2.5
	Faster R-CNN-Focal Loss	88.462	90.364	94.120	134±3.1
	EfficientDet-GHM Loss	86.315	84.336	87.036	42±1.8
	GA-YOLOv8	79.423	80.157	84.647	58±2.2
	Model	Test indicators			Test indicators
		mAP (%)	Accuracy (%)	Recall (%)	FPS
Civil Aviation Dataset 2	BO-PSO-Focal Loss-YOLOv8	97.125	97.652	98.414	140±2.4
	Faster R-CNN-Focal Loss	87.122	92.146	93.235	132±2.9
	EfficientDet-GHM Loss	81.335	86.456	86.123	43±1.9
	GA-YOLOv8	74.463	82.156	82.036	57±2.1

Table 3. mAP of BO-PSO Focal Loss-YOLOv8 model in various environments (%)

Time period	Models	Natural light	Bright light	Weak light	Sparse	Moderate	Dense
Morning	BO-PSO-Focal Loss-YOLOv8	96.227	96.213	96.141	96.224	96.189	96.134
	Faster R-CNN-Focal Loss	92.145	91.876	90.234	92.001	90.845	89.673
	EfficientDet-GHM Loss	88.934	87.652	86.123	88.745	87.112	85.994
	GA-YOLOv8	85.672	84.321	82.456	85.501	83.987	82.134
Night	BO-PSO-Focal Loss-YOLOv8	95.884	95.762	95.621	95.801	95.432	95.551
	Faster R-CNN-Focal Loss	90.123	89.654	88.112	90.002	88.745	87.331
	EfficientDet-GHM Loss	86.445	85.123	83.678	86.331	84.992	83.451
	GA-YOLOv8	83.214	81.987	80.112	83.001	81.456	80.023

Author Contributions

Both authors contributed equally, jointly participating in the article's conception, methodology, data processing, and writing.

Funding

The research is supported by: An Empirical Study on the Industry-University "Integration, Interconnection and Co-construction of Mechanisms" Talent Training System for the Civil Aviation Safety Technology Management Major in the Guangdong-Hong Kong-Macao Greater Bay Area, Grant Number: 2024GXJK174; A Study on the Collaborative Governance Mechanism of Security Rules for the Airport Cluster in the Guangdong-Hong Kong-Macao Greater Bay Area Under the Background of "One Country, Two Systems and Three Jurisdictions", Grant Number: 2025WQNCX136.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The authors used Grammarly solely for language editing and readability improvement. The authors reviewed and verified all content and take full responsibility for the accuracy and integrity of the manuscript.

Reference

Bello, R. W., and Oladipo, M. A. (2024). Mask YOLOv7-based drone vision system for automated cattle detection and counting. *Artificial Intelligence and Applications,* 2(2), 115–125.

Bhavana N., Kodabagi M. M., and Kumar B. M (2024). POT-YOLO: Real-time road potholes detection using edge segmentation-based YOLO V8 network. *IEEE Sensors Journal,* 24(15): 24802-24809.

Challa, S. K., Kumar, A., and Semwal, V. B. (2022). A multibranch CNN-BiLSTM model for human activity recognition using wearable sensor data. *The Visual Computer,* 38(12), 4095–4109.

Chitraningrum, N., Banowati, L., and Herdiana, D. (2024). Comparison study of corn leaf disease detection based on deep learning YOLO-v5 and YOLO-v8. *Journal of Engineering and Technological Sciences,* 56(1), 61–70.

Dewi, C., Chen, R. C., and Jiang, X. (2022). Deep convolutional neural network for enhancing traffic sign recognition developed on YOLO V4. *Multimedia Tools and Applications,* 81(26), 37821–37845.

Gao, G., Shuai, C., and Wang, S. (2024). Using improved YOLO V5s to recognize tomatoes in a continuous working environment. *Signal, Image and Video Processing,* 18(5), 4019–4028.

Gkournelos, C., Konstantinou, C., and Angelakis, P. (2024). Praxis: A framework for AI-driven human action recognition in assembly. *Journal of Intelligent Manufacturing,* 35(8), 3697–3711.

Kumar, A., Kalia, A., and Sharma, A. (2023). A hybrid tiny YOLO v4-SPP module based improved face mask detection vision system. *Journal of Ambient Intelligence and Humanized Computing,* 14(6), 6783–6796.

Lee, J., and Hwang, K. (2022). YOLO with adaptive frame control for real-time object detection applications. *Multimedia Tools and Applications,* 81(25), 36375–36396.

Ma, D., Fang, H., and Wang, N. (2022). Automatic detection and counting system for pavement cracks based on PCGAN and YOLO-MF. *IEEE Transactions on Intelligent Transportation Systems,* 23(11), 22166–22178.

Mathew, M. P., and Mahesh, T. Y. (2022). Leaf-based disease detection in bell pepper plant using YOLO v5. *Signal, Image and Video Processing,* 16(3), 841–847.

Mittal, U., Chawla, P., and Tiwari, R. (2023). EnsembleNet: A hybrid approach for vehicle detection and estimation of traffic density based on faster R-CNN and YOLO models. *Neural Computing and Applications,* 35(6), 4755–4774.

Pham, D. L., and Chang, T. W. (2023). A YOLO-based real-time packaging defect detection system. *Procedia Computer Science,* 217, 886–894.

Rajamohanan, R., and Latha, B. C. (2023). An optimized YOLO v5 model for tomato leaf disease classification with field dataset. Engineering, *Technology and Applied Science Research,* 13(6), 12033–12038.

Sangaiah, A. K., Yu, F. N., and Lin, Y. B. (2024). UAV T-YOLO-rice: An enhanced tiny YOLO network for rice leaf disease detection in paddy agronomy. *IEEE Transactions on Network Science and Engineering,* 11(6), 5201–5216.

Savchenko, A. V., Savchenko, L. V., and Makarov, I. (2022). Classifying emotions and engagement in online learning based on a single facial expression recognition neural network. *IEEE Transactions on Affective Computing,* 13(4), 2132–2143.

Singh, J. P., and Kumar, M. (2023). Conditional autoregressive-tunicate swarm algorithm based generative adversarial network for violent crowd behavior recognition. *Artificial Intelligence Review,* 56(2), 2099–2123.

Walter S., Schmutz J. B., and Grote G. (2024). A safety approach for improving security-effective coordination strategies at the airport security screening checkpoint. *Human factors,* 66(4): 1302-1314.

Xiao B., Nguyen M., and Yan W. Q. (2024). Fruit ripeness identification using YOLOv8 model. *Multimedia Tools and Applications,* 83(9): 28039-28056.

Xu, M., Guo, L., and Wu, H. C. (2023). Robust abnormal human-posture recognition using OpenPose and multiview cross-information. *IEEE Sensors Journal,* 23(11), 12370–12379.

Fengwei Chen obtained his master's in Software Engineering from Beijing Institute of Technology in 2015. Currently, he serves as a professional instructor at Guangzhou Civil Aviation College. With rich industry experience in civil aviation safety, his priority is teaching and research on aviation safety, civil aviation safety inspection, and counter-terrorism technology. As an industry expert and consultant, Chen was invited to deliver technical lectures on topics including civil aviation safety and security inspection technology. He has participated in provincial and ministerial-level projects and contributed to the development of teaching standards for civil aviation safety inspection. Chen has published articles in both domestic and international journals and conference proceedings. Additionally, he is frequently invited to serve as a judge in civil aviation safety inspection competitions. His research interests include aviation safety, security inspection technology, pattern recognition, machine learning, and virtual simulation.

Tianyou Wu obtained his master's in Business Administration from the Chinese University of Hong Kong in 2017. Mr. Wu currently serves as an instructor specializing in security inspection at Guangzhou Civil Aviation College. He has consistently worked in the field of civil aviation safety, encompassing areas such as civil aviation security inspection, civil aviation emergency rescue, aviation English, management studies, and organizational behavior. He has published numerous articles related to civil aviation safety technology in domestic academic journals and has been involved in multiple provincial and ministerial-level projects. His professional research interests include aviation safety, security inspection technology, airport emergency rescue, and improving the quality of security inspection services.