Duan *et al. Intell. Robot.* **2025**, *5*(1),70-87
DOI: 10.20517/ir.2025.05

Intelligence & Robotics

Research Article

Open Access

A novel real-time intelligent detector for monitoring UAVs in live-line operation on 10 kV distribution networks

Haibo Duan[1], Fanrong Shi[1], Bo Gao[2], Yingyue Zhou[1], Qiushi Cui[3]

[1]School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, Sichuan, China.
[2]Electric Power Research Institute of State Grid Ningxia Electric Power Co., Ltd., Yinchuan 750011, Ningxia, China.
[3]School of Electrical Engineering, Chongqing University, Chongqing 400044, China.

Correspondence to: Prof. Fanrong Shi, School of Information Engineering, Southwest University of Science and Technology, No.59, Middle Section of Qinglong Avenue, Fucheng District, Mianyang 621010, Sichuan, China. E-mail: sfr_swust@swust.edu.cn

How to cite this article: Duan, H.; Shi, F.; Gao, B.; Zhou, Y.; Cui, Q. A novel real-time intelligent detector for monitoring UAVs in live-line operation on 10 kV distribution networks. *Intell. Robot.* **2025**, *5*(1), 70-87. http://dx.doi.org/10.20517/ir.2025.05

Received: 21 Oct 2024 **First Decision:** 11 Dec 2024 **Revised:** 22 Dec 2024 **Accepted:** 3 Jan 2025 **Published:** 21 Jan 2025

Academic Editor: Simon Yang **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

The live-line operation of 10 kV distribution networks is critical for ensuring uninterrupted and high-quality power supply. However, operational sites face challenges such as insufficient intelligent monitoring and suboptimal real-time performance. To address these issues, this study proposes the FEM-YOLOv8 algorithm, specifically designed for protective equipment detection in live-line operation scenarios. The proposed algorithm is deployed on edge devices compatible with unmanned aerial vehicles (UAVs), enabling remote, autonomous, and intelligent monitoring. Key improvements include the introduction of an enhanced FAST-C2f module, replacing the original C2f module in the Backbone to improve feature extraction efficiency while reducing model complexity. Additionally, a lightweight efficient channel attention (ECA) mechanism is incorporated into the Backbone and Neck to enhance target feature detection and representation capabilities. The bounding box regression loss function is replaced with metric preserving distance intersection over union (MPDIoU) to further boost detection accuracy and robustness. The FEM-YOLOv8 model is implemented on the Atlas 200I DK A2 edge device, which is suitable for UAV deployment. Experimental results demonstrate that the improved FEM-YOLOv8 model achieves 93.1% precision (P), 85.9% recall (R), and 92.3% mean average precision (mAP), surpassing the baseline model by 2.8, 3.2, and 2.2 percentage points, respectively. With a detection speed of 83 frames per second (FPS) and a power consumption of only 10.2 W, the model satisfies real-time performance and detection accuracy requirements, providing significant contributions to grid intelligence and power operation safety.

Keywords: 10 kV distribution networks, YOLOv8s, object detection, real-time, live-line operation

1. INTRODUCTION

The live-line operation of 10 kV distribution networks is widely recognized as an effective approach to avoiding maintenance-related power outages and ensuring uninterrupted power supply. However, as the demand for live-line operation continues to grow, the increasing complexity and difficulty of these tasks have also led to heightened risks[1]. At present, live-line operations are primarily supervised manually at distribution sites, which incurs significant labor costs. Additionally, maintaining prolonged concentration during monitoring is challenging for personnel, often resulting in subjective judgments that lead to false or missed detections. These issues make it difficult to address and prevent safety incidents promptly, thus significantly hindering the automation and intelligent management of power systems[2]. Risk prevention and control have consequently become critical areas of safety research for live-line operations, playing a pivotal role in ensuring worker safety and advancing the development of distribution networks[3]. Although video surveillance and similar technical measures are employed on-site to assist in monitoring[4], such as supervising the wearing of personal protective equipment (PPE) and observing operational behaviors, existing cloud-based control systems often lack real-time capabilities and adequate intelligence[5,6]. This limitation becomes particularly evident in scenarios involving unsafe behaviors, such as improper usage of insulated helmets, capes, or gloves, where on-site systems fail to provide timely alerts. To address these challenges, it is imperative to develop intelligent methods capable of real-time detection of PPE to ensure worker safety and enhance the safety management of live-line operations. The integration of advanced detection techniques with intelligent monitoring systems offers a promising solution for improving the safety and efficiency of live-line operations in distribution networks.

Edge computing has been recognized as an innovative paradigm designed to perform real-time data analysis and computation closer to sensing terminals[7,8]. This approach offers notable advantages, including low latency, high efficiency, reduced network bandwidth consumption, and enhanced privacy protection[9]. In recent years, significant advancements in unmanned aerial vehicle (UAV) technology have facilitated its widespread application across various domains, such as border patrol, fire detection, traffic monitoring, and power line inspection. UAVs, equipped with high-definition cameras, are capable of real-time image data collection. When combined with edge computing, this capability enables efficient data processing and analysis, allowing timely detection of faults and potential risks[10]. Simultaneously, deep learning technology has demonstrated remarkable advantages in intelligent detection tasks. Through the use of multi-layer neural networks, high-level features can be extracted from image data, substantially enhancing detection accuracy and real-time processing capabilities. In complex and dynamic environments, deep learning has shown superior performance, making it widely applicable to tasks such as power line inspections[11]. Additionally, distributed communication and event-triggered strategies have proven effective in reducing communication overhead and improving response efficiency in power system management[12,13], providing theoretical support for real-time intelligent detection. Based on these advancements, an intelligent detection model for PPE, FEM-YOLOv8, is proposed in this study. The model leverages deep learning technology and is deployed on edge devices compatible with UAVs to enable real-time intelligent detection of PPE at operation sites. This approach offers significant contributions to the digitalization and intelligent management of power grids.

The key contributions are summarized as follows:

(1) To minimize model complexity while maintaining detection accuracy, the partial convolution (PConv) operation was introduced into the Backbone, forming the FAST-C2f module to replace the original C2f module. This effectively reduces resource consumption. Moreover, a lightweight efficient channel attention (ECA) mechanism was incorporated into both the Backbone and Neck, enabling the model to dynamically learn

channel importance and significantly enhance detection accuracy.

(2) Metric preserving distance intersection over union (MPDIoU) was employed as the bounding box regression loss function, accelerating network convergence and further improving detection performance and robustness.

(3) The FEM-YOLOv8 model was deployed on edge devices equipped with Huawei's Ascend 310 AI accelerator chip. Notably, the acronym FEM reflects the initials of the three key improvements proposed in this study: FAST-C2f, ECA, and MPDIoU, clearly highlighting the algorithm's core innovations. Additionally, a real-time intelligent inspection solution integrating UAVs for live-line operations was developed, establishing a robust foundation for intelligent management at operation sites.

2. RELATED WORKS

In recent years, deep learning object detection algorithms have found extensive use in intelligent detection tasks for power applications[14]. These algorithms can be mainly categorized into two-stage detection methods, represented by fast region-based convolutional neural network (R-CNN)[15] and faster R-CNN[16], and one-stage detection methods, represented by You Only Look Once (YOLO)[17–21] and single shot multibox detector (SSD)[22]. Two-stage detection methods first extract candidate boxes and then classify these regions while performing bounding box regression. For example, Kong *et al.* used the Mask R-CNN model to detect whether construction workers at power sites were wearing safety helmets[23]. Lei *et al.* proposed a Faster R-CNN-based method to detect broken power line insulators and bird nests[24]. Yang *et al.* utilized the Faster R-CNN model to detect oil leakage incidents in substations, providing robust technical support for the safe operation of substations[25]. While two-stage detection methods are known for their high accuracy, they suffer from poor real-time performance and high memory consumption[26]. As a result, researchers have shifted their focus to one-stage detection algorithms.

One-stage detection methods operate end-to-end, directly extracting features and predicting their locations and other attributes, which effectively improves detection efficiency. Compared to two-stage algorithms, one-stage methods are simpler, more efficient, and better suited for engineering applications[27]. Consequently, they have been widely adopted for detection tasks at power operation sites[28]. For instance, Li *et al.* utilized the YOLOv5s detection model and deployed it on UAVs to detect bird nests on transmission lines[29]. Zhao *et al.* applied an improved YOLOv5s algorithm to detect missing bolts in transmission line inspection images captured by UAVs[30]. He *et al.* conducted research on detecting equipment such as safety helmets and protective clothing worn by power workers, developing an object detection algorithm based on the YOLOv7-Tiny model, which enhanced detection performance under complex backgrounds and target occlusion conditions[31]. Panigrahy *et al.* used the YOLOv8n object detection model in conjunction with UAVs equipped with edge computing devices to effectively identify the condition of insulators (e.g., damaged or normal), significantly enhancing the intelligence level of power line inspection[32]. Wang *et al.* proposed a YOLOv3-based object detection and tracking method for the safety detection of live operations in power distribution networks[33]. This method achieved high detection accuracy and real-time performance by detecting the wearing of helmets and identifying hazardous area intrusions in distribution network scenarios, effectively improving the safety management of power systems. However, YOLOv3 still has limitations in detection capability under complex scenarios, particularly with small objects or in complex backgrounds. Li proposed a power operation site risk identification technique based on the YOLOv8 model with the introduction of the BoTNet module[34]. This method optimized existing object detection algorithms, significantly improving detection accuracy. However, the addition of the BoTNet module increased the complexity of the detection model, leaving room for improvement in terms of efficiency and real-time performance.

In summary, the combination of deep learning-based intelligent detection models and UAV technology has achieved remarkable results in power inspection scenarios. However, in the context of live operations in power distribution networks, many methods still face challenges such as high model complexity, insufficient real-time performance, and the lack of deployment on edge computing platforms to achieve real-time detection. To address these issues, we propose an improved FEM-YOLOv8 intelligent detection model for power operation protection. This model is optimized based on the YOLOv8 algorithm and deployed on edge devices compatible with UAVs, providing robust technical support for the real-time intelligent management and control of power operations.

3. FEM-YOLOV8 ALGORITHM

Released in 2023 as part of the YOLO series, YOLOv8 is recognized as the successor to YOLOv5, offering improved capabilities in object detection. The model provides five scales: N (Nano), S (Small), M (Medium), L (Large), and X (Extra Large). Among these, the lightweight N and S models feature fewer parameters and faster inference speeds but lower detection accuracy, whereas the L and X models prioritize higher accuracy at the cost of increased computational demands, making them suitable for performance-critical applications. To meet the real-time, accuracy, and edge deployment requirements of high-risk live-line operation scenarios, the YOLOv8s model is selected in this study as the baseline for further optimization. Structurally, YOLOv8 consists of three primary components: Backbone, Neck, and Head, as illustrated in Figure 1. Compared to YOLOv5, significant optimizations and improvements have been incorporated into these components. In the Backbone, the cross stage partial (CSP) concept is retained, with the C3 modules from YOLOv5 refined into more lightweight C2f modules. The number of channels has also been adjusted to further reduce complexity. For the Neck, the path aggregation network (PAN)–feature pyramid network (FPN) structure is preserved [35,36], but the convolutional layers before the Upsample layers have been removed. Furthermore, all C3 modules in the Neck are replaced with the improved C2f modules. In the Head, a decoupled head structure has been introduced, separating the classification head from the detection head. Additionally, the Anchor-Free concept [37] has been integrated to align the predicted boxes more closely with the ground truth, thereby improving detection accuracy and overall model performance. These enhancements are designed to achieve an optimal balance between lightweight design and performance, making YOLOv8 a more effective solution for diverse object detection tasks.

The original YOLOv8s model is characterized by relatively high resource consumption, rendering it unsuitable for deployment on resource-constrained edge devices. Furthermore, its detection accuracy does not meet the stringent precision requirements of live-line operation scenarios. To overcome these limitations, an improved model, FEM-YOLOv8, is proposed in this study based on YOLOv8s. The improved model achieves a balance between lightweight design and detection performance, enabling efficient deployment on edge devices while satisfying the demands for real-time detection and high precision. FEM-YOLOv8 is particularly suited for complex scenarios such as live-line operations. The detailed structure of FEM-YOLOv8 is depicted in Figure 2, with the proposed improvements visually highlighted in red. First, a new FAST-C2f structure is introduced using the PConv method to replace the original C2f modules in the feature extraction stage. This adjustment effectively reduces the model's complexity, making it more suitable for edge applications. Second, an ECA attention mechanism is incorporated at the end of the Backbone network and after the Upsample modules in the Neck, enhancing feature selection across channels and improving overall feature representation. Finally, MPDIoU is adopted as the bounding box regression loss function, accelerating the convergence of the algorithm and improving the localization accuracy of detected object boundaries.

3.1. FAST-C2f module

To enhance the deployability of models on resource-constrained edge devices, commonly used lightweight model architectures such as MobileNet [38] and ShuffleNet [39] have been developed. These architectures ex-

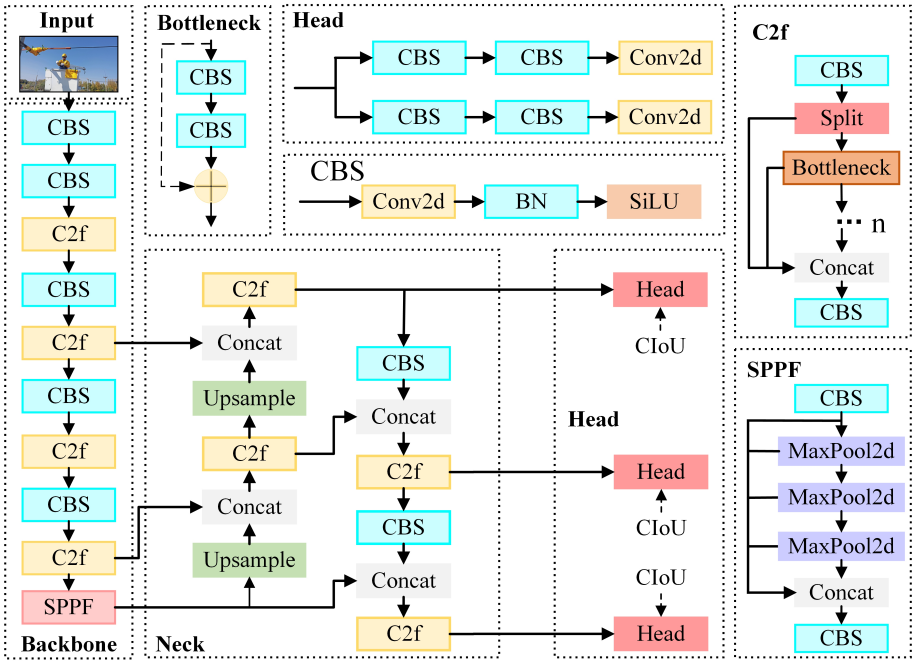

Figure 1. YOLOv8s network structure.

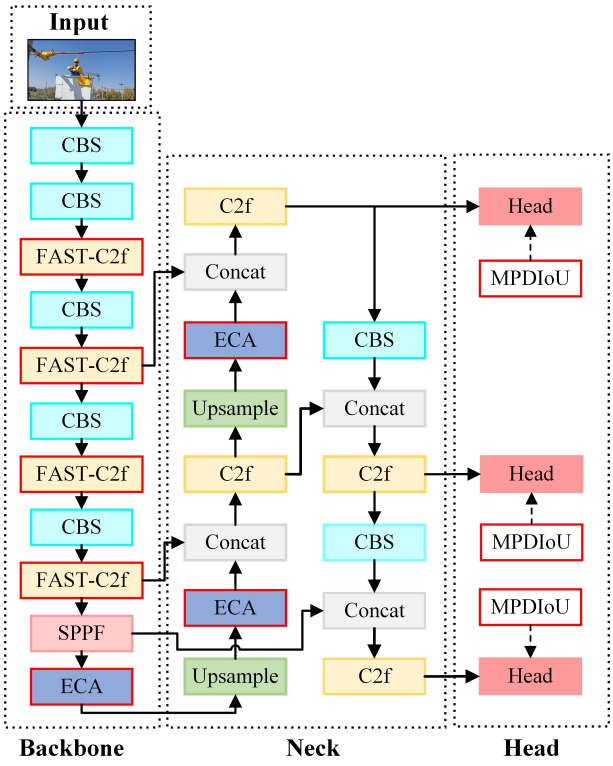

Figure 2. Structure of the FEM-YOLOv8 model.

tract features from input data using depthwise convolution or group convolution. However, despite effectively reducing giga floating-point operations per second (GFLOPs), these methods often encounter increased mem-

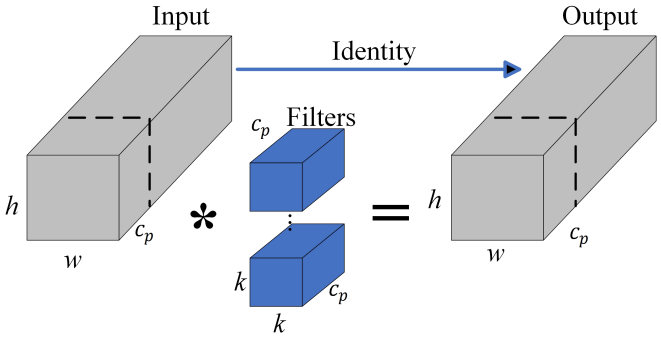

Figure 3. PConv structure diagram. PConv: Partial convolution.

ory access requirements, which negatively affect detection accuracy.

To address this limitation, a method called PConv was proposed by Chen *et al.*[40]. PConv performs standard convolution on a subset of the input feature map's channels, while the remaining channels are preserved for subsequent 1×1 pointwise convolution to retain and utilize additional information. This approach significantly reduces computational redundancy and memory access, improving both the detection accuracy and speed of convolutional neural network models. The structure of PConv is illustrated in Figure 3.

The number of floating-point operations for PConv can be calculated as follows:

$$FLOPs = h \times w \times k^2 \times (c_p)^2 \tag{1}$$

The memory access characteristics of PConv are as follows:

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \tag{2}$$

Where h, w, and k represent the height, width, and kernel size of the feature map, respectively. c_p denotes the number of channels used for conventional convolution, while c represents the total number of input channels. Since typically only $1/4$ of the input channels are involved in the convolution operation, $c_p/c = 1/4$, the computational cost of PConv is approximately $1/16$ that of standard convolution, and the memory access requirement is reduced to about $1/4$. The remaining $c - c_p$ channels are excluded from computation, requiring no memory access.

In this study, the concept of PConv is employed to design the FasterBlock, which replaces the Bottleneck in the original C2f module, resulting in the new FAST-C2f module. The structure of the FAST-C2f module is shown in Figure 4. FasterBlock consists of one PConv and two 1×1 convolutional layers, forming a reverse residual block structure. Batch normalization (BN) and the rectified linear unit (ReLU) activation function are applied only after the middle 1×1 convolutional layer to preserve feature diversity while achieving lower latency. The FAST-C2f module is integrated into the Backbone for feature extraction and processing, enhancing detection performance while significantly reducing the number of parameters and computational complexity.

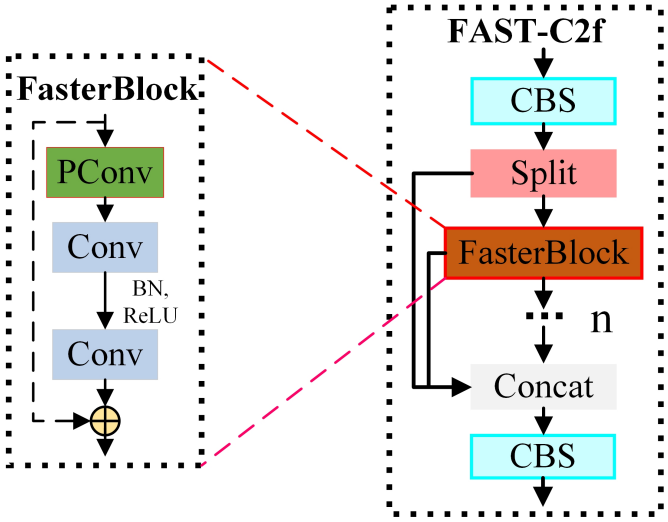

Figure 4. FAST-C2f structure diagram.

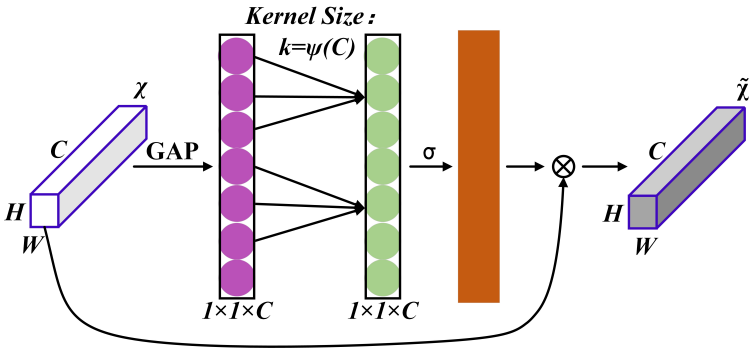

Figure 5. Structural diagram of the ECA module. ECA: Efficient channel attention.

3.2. ECA attention mechanism

In the task of detecting protective equipment in live-line operation scenarios, the large working range of personnel leads to highly complex and variable background features, including interference from power lines, buildings, towers, and other intricate structures. These factors significantly increase the difficulty of detection. To address this issue, the ECA attention mechanism is incorporated into the network to enable the model to focus more effectively on the critical features of the target objects. ECA[41] was proposed based on the observation that dimensionality reduction in squeeze-and-excitation (SE) models[42] can adversely affect prediction performance. Unlike traditional attention mechanisms, ECA offers a lightweight and efficient design that enhances model performance without significantly increasing computational complexity. The structure of the ECA module is depicted in Figure 5.

The module first applies global average pooling to the input data to aggregate spatial information. A 1×1 convolutional layer is then employed to perform cross-channel information interaction, avoiding dimensionality reduction. The size of the convolution kernel k is adaptively determined using

$$k = \psi(C) = \left| \frac{\log_2(C) + 1}{2} \right|_{odd} \tag{3}$$

Where C represents the channel dimension, and the $|*|_{odd}$ operator denotes the nearest odd number.

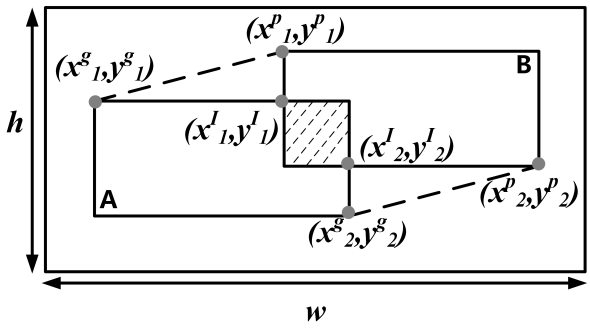

Figure 6. Calculation method of MPDIoU. MPDIoU: Metric preserving distance intersection over union.

Subsequently, k is applied through one-dimensional convolution. Finally, a Sigmoid activation function is used to calculate the weight of each channel, and the resulting weights are multiplied element-wise with the original input feature map to generate the final feature map.

3.3. MPDIoU loss function

In YOLOv8, the complete intersection over union (CIoU) loss function is employed to optimize object detection. CIoU considers three factors: the overlap area between the predicted box and the ground truth box, the distance between their center points, and their aspect ratio. However, the optimization effect is reduced when the predicted box and the ground truth box share the same aspect ratio but differ in width and height, which limits the accuracy of model training[43]. To overcome this limitation, a new bounding box regression loss function, MPDIoU[44], is introduced in this study to replace the original CIoU. The structure and workflow of MPDIoU are illustrated in Figure 6.

In Figure 6, the outer rectangular box represents the input image, with h and w denoting the height and width of the input image. Box A indicates the ground truth box, while Box B represents the predicted box by the model. The dashed line in the top-left corner shows the distance between the top-left corners of the ground truth and predicted boxes, and the dashed line in the bottom-right corner indicates the distance between the bottom-right corners of the ground truth and predicted boxes. In this figure, (x_1^g, y_1^g) and (x_1^p, y_1^p) represent the top-left coordinates of the ground truth and predicted boxes, respectively, while (x_2^g, y_2^g) and (x_2^p, y_2^p) denote the bottom-right coordinates of the ground truth and predicted boxes. Additionally, (x_1^I, y_1^I) and (x_2^I, y_2^I) indicate the bottom-left and top-right coordinates of the intersection between the ground truth and predicted boxes, respectively.

The loss function is written as follows:

$$d_1{}^2 = (x_1{}^p - x_1{}^g)^2 + (y_1{}^p - y_1{}^g)^2 \tag{4}$$

$$d_2{}^2 = (x_2{}^p - x_2{}^g)^2 + (y_2{}^p - y_2{}^g)^2 \tag{5}$$

$$L_{MPDIoU} = 1 + (d_1^2 + d_2^2)/(h^2 + w^2) - IoU \tag{6}$$

Here, d_1 and d_2 represent the distances between the top-left corners and the bottom-right corners of the predicted box and the ground truth box, respectively, and IoU denotes the intersection over union between the

Figure 7. Sample images from the dataset. All images are original, and copyright is retained by the authors.

Table 1. Detailed distribution of target bounding boxes in the sample library (Unit: Number of instances)

Category	Training set	Validation set	Test set	Total
helmet	1,546	201	197	1,944
shawl	1,480	190	191	1,861
gloves	1,799	219	227	2,245

ground truth box and the predicted box.

4. EXPERIMENTS

4.1. Experimental environment and dataset

The model training experiments were conducted on the Huawei Cloud ModelArts platform. The "pytorch1.8-cuda10.2-cudnn7-ubuntu18.04" image was used, with resource specifications of "GPU: 1*P100 (16 GB) | CPU: 8 cores, 64 GB". Due to the absence of publicly available datasets for live-line operations, a custom sample library was constructed using images collected from a live-line operation training base and real operation sites in an autonomous region. Data collection devices included the Xiaomi 13 smartphone and DJI Mavic 3 UAV. Following the removal of blurry and duplicate images, a total of 1,346 images were retained in the sample library. To visually demonstrate the collected image data, Figure 7 provides examples of images from the sample library. These images were collected from live-line operation training bases and real operation sites in the autonomous region, covering various typical live-line operation scenarios and fully reflecting the diversity of the dataset and the complexity of the operation site backgrounds.

The collected data were filtered and annotated using the MakeSense labeling tool, and the images were categorized into three classes: insulated helmets (helmets), insulated shawls (shawls), and insulated gloves (gloves). For the experiments, the dataset was randomly divided into training, validation, and test sets in an 8:1:1 ratio. Table 1 presents the detailed distribution of target bounding boxes for each class across the three subsets, with the unit indicating the number of instances. Specifically, the training set contains 1,546 helmets, 1,480 shawls, and 1,799 gloves. The validation set includes 201 helmets, 190 shawls, and 219 gloves, while the test set comprises 197 helmets, 191 shawls, and 227 gloves. This dataset serves as a reliable foundation for training, validation, and testing of the FEM-YOLOv8 detection algorithm developed in this study.

To enhance the model's generalization capability and reduce the risk of overfitting, a comprehensive online data augmentation strategy was applied during training. The augmentation process included two primary steps. First, basic transformations such as translation, scaling, and flipping were applied to the input images.

Figure 8. The rendering after data augmentation. The image is original, and copyright is retained by the authors.

Second, mosaic data augmentation was implemented by randomly cropping and combining four images into one. This approach effectively enriched the diversity and comprehensiveness of the dataset samples, contributing to improved training performance. The results of the data augmentation process are presented in Figure 8.

4.2. Evaluation indicator and parameter configuration

The training accuracy of the model is evaluated using the widely adopted metrics of precision, recall, and mean average precision (mAP). Considering the deployment of the model on resource-constrained edge devices, additional metrics such as the number of parameters (Params), computational cost (GFLOPs), model size (Model Size), and frames per second (FPS) are introduced to assess the model's complexity and computational efficiency. These metrics provide a comprehensive basis for comparing the proposed approach with existing methods.

The precision (P) and recall (R) are expressed as follows:

$$P = \frac{TP}{TP + FP} \times 100\% \tag{7}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{8}$$

Where TP denotes true positives, where positive samples are correctly identified; FP refers to false positives, where negative samples are incorrectly identified as positive; and FN represents false negatives, where positive samples are incorrectly identified as negative.

The mAP is used to evaluate the overall detection accuracy across all categories. The task studied in this paper

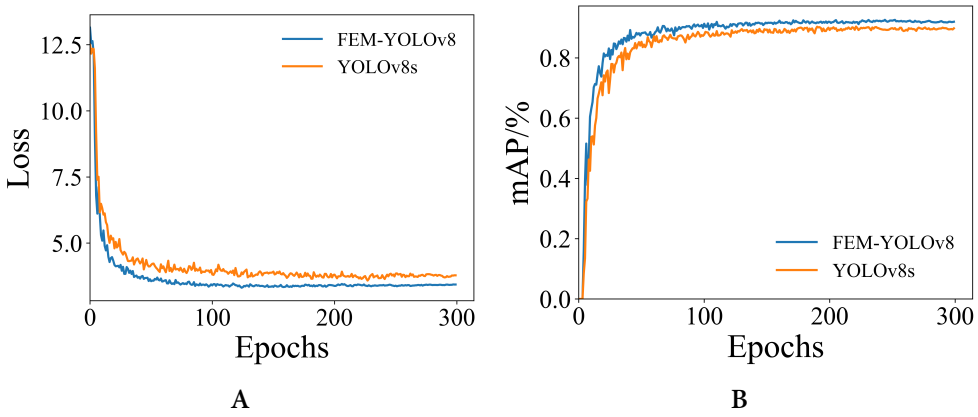

Figure 9. (A) Comparison of loss curves between YOLOv8s and FEM-YOLOv8; (B) Comparison of mAP curves between YOLOv8s and FEM-YOLOv8. mAP: Mean average precision.

is a multi-class classification problem. For each category, precision and recall are calculated, and a precision-recall (P-R) curve is plotted. The area under the curve represents the average precision (AP) for that category. The mAP is then obtained by averaging the AP values of all categories, as expressed in:

$$AP = \int_0^1 P(R)dR \tag{9}$$

$$mAP = \frac{\sum_{i=1}^{n} AP_i}{n} \times 100\% \tag{10}$$

where n denotes the total number of categories, and AP_i represents the AP of the i-th category.

During model training, the input image size was uniformly set to 640 × 640, with a batch size of 64 and eight worker threads. The training process was conducted for 300 epochs. The initial learning rate was set to 0.01, and stochastic gradient descent (SGD) was employed as the training optimizer. The loss function and mAP curves for both the FEM-YOLOv8 model and the baseline YOLOv8s model are presented in Figure 9. As illustrated in Figure 9A and B, the loss function values for both models exhibit a decreasing trend, while the mAP values increase during training. After 200 iterations, the loss values and mAP stabilize, indicating that the network has converged. Compared to YOLOv8s, the FEM-YOLOv8 model demonstrates faster convergence, lower loss values, and higher mAP, highlighting its superior detection performance.

4.3. Experimental results and analysis

4.3.1 Ablation experiments

To further evaluate the effectiveness of the proposed improvements, ablation experiments were conducted, and the results are presented in Table 2. In the table, "✓" indicates the inclusion of specific improvements. As illustrated in Table 2, the proposed enhancements effectively improve the model's performance while optimizing its lightweight design. The incorporation of the FAST-C2f module resulted in a 0.2% reduction in P, accompanied by increases of 0.2% in R and mAP. Additionally, the parameter count and computational cost were reduced by 12.6% and 14.4%, respectively, achieving a significant reduction in complexity without compromising accuracy. The introduction of the lightweight ECA attention mechanism and the MPDIoU loss function further enhanced the model's performance. The ECA attention mechanism improved the model's ability to focus on critical targets within input images, achieving increases of 2.5%, 1.6%, and 1.5% in P, R, and

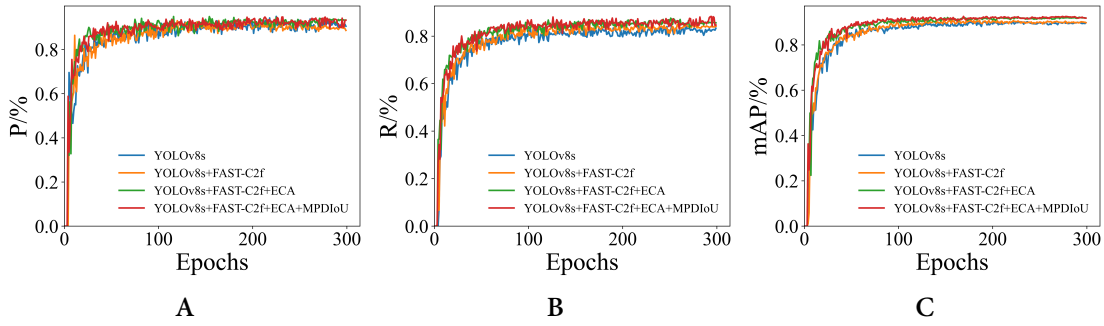

Figure 10. Performance comparison of P, R, and mAP in ablation experiments. (A) Precision curve; (B) Recall curve; (C) mAP curve. P: Precision; R: recall; mAP: mean average precision.

Table 2. Ablation experiment results

Methods	FAST-C2f	ECA	MPDIoU	P/%	R/%	mAP/%	Parameters/10^6	GFLOPs
1				90.2	83.1	90.3	11.1	28.4
2	✓			89.5	84.3	90.5	9.7	24.3
3	✓	✓		92.0	85.9	92.0	9.7	24.3
4	✓	✓	✓	92.3	86.4	92.4	9.7	24.3

ECA: Efficient channel attention; MPDIoU: metric preserving distance intersection over union; P: precision; R: recall; mAP: mean average precision; GFLOPs: giga floating-point operations per second.

mAP, respectively, compared to using only the FAST-C2f module, without increasing the parameter count or computational cost. Overall, the proposed improvements enabled the final FEM-YOLOv8 model to achieve increases of 2.1%, 3.3%, and 2.1% in P, R, and mAP, respectively, relative to the original YOLOv8s model, while reducing the parameter count and computational cost by 12.6% and 14.4%. These enhancements not only ensure a lightweight design but also significantly improve detection accuracy and real-time performance, fully satisfying the requirements for efficient real-time detection in live-line operation scenarios. The proposed improvements have been demonstrated to be both effective and practical.

Figure 10 illustrates the variation curves for key metrics, including P, R, and mAP, obtained during the ablation experiments. As observed in the figure, all proposed improvements contributed to an increase in mAP, with the most significant enhancement attributed to the integration of the ECA attention mechanism. All models were trained for 200 epochs, after which the curves stabilized, indicating convergence. Compared to the baseline YOLOv8s model, the improved FEM-YOLOv8 model achieved faster convergence and superior detection performance, as reflected in the metric trends.

4.3.2 Comparison experiments

To evaluate the performance of the proposed FEM-YOLOv8 model, comparisons were conducted with several mainstream one-stage object detection models suitable for edge deployment, as presented in Table 3. The results demonstrate that the FEM-YOLOv8 model offers significant advantages in detecting protective equipment. Specifically, the params, GFLOPs, and model size of FEM-YOLOv8 are 9.7×10^6, 24.3, and 18.7 MB, respectively, representing reductions of 12.6%, 14.4%, and 12.6% compared to the original YOLOv8s model, thereby achieving a lightweight design. In terms of detection accuracy, the mAP of FEM-YOLOv8 reaches 92.4%, exhibiting improvements of 6.5%, 4.8%, 11.9%, 3.7%, and 3.5% over SSD, YOLOv3, YOLOv3-tiny, YOLOv5s, and YOLOv10s, respectively. Furthermore, the edge deployment tests, as shown in Table 4, indicate that FEM-YOLOv8 meets real-time detection requirements ($\geq$ 30 FPS) and demonstrates exceptional efficiency and real-time capability. Overall, the improved FEM-YOLOv8 model surpasses other algorithms in terms of mAP while maintaining lower model complexity, making it well-suited for high-risk live-line op-

Table 3. Comparison experiment results

Model	mAP/%	Params/10^6	GFLOPs	Model size/MB
SSD	85.9	23.6	61.1	92.2
YOLOv3	87.6	61.5	154.6	118.0
YOLOv3-tiny	80.5	8.7	12.9	16.8
YOLOv5s	88.7	7.0	15.8	14.4
YOLOv8s	90.3	11.1	28.4	22.0
YOLOv10s	88.9	8.0	24.8	16.4
FEM-YOLOv8	**92.4**	9.7	24.3	18.7

mAP: Mean average precision; GFLOPs: giga floating-point operations per second.

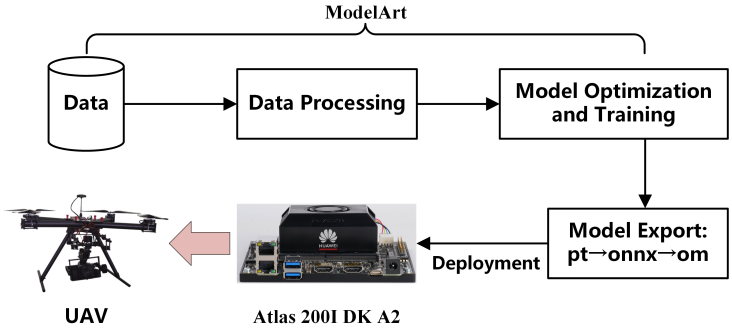

Figure 11. Application scheme.

eration scenarios. It satisfies the dual requirements of detection accuracy and real-time performance, thereby offering significant practical application value.

5. DEPLOYMENT

The edge deployment platform utilized in this study is the Atlas 200I DK A2, an AI development platform designed by Huawei for edge computing scenarios. The platform is applicable to pre-research and development in fields such as robotics, UAVs, and audio-video processing. It comprises three primary components: the Atlas 200 AI acceleration module, a multimedia processing chip (Hi3559C), and a local area network (LAN) switch. The Atlas 200 AI integrates the Ascend 310 AI processor, a high-performance, low-power chip tailored for tasks such as image recognition, video processing, inference computation, and machine learning. A maximum computing power of 10 TOPS (FP16) is achieved by the device. The platform operates on the Ubuntu 22.04 LTS operating system, and its neural network computing architecture is based on compute architecture for neural networks (CANN) version 7.0.RC1.

Figure 11 illustrates the application scheme of the proposed method. First, the model is optimized and trained on the ModelArts platform. Then, the trained model undergoes a format conversion process to ensure compatibility with edge devices. This includes three stages: exporting the model in the pt format (the original training model generated by the PyTorch AI framework), converting it to the onnx format (an open neural network exchange format developed by Facebook and Microsoft, facilitating model conversion and optimization across diverse frameworks), and finally transforming it into the om format (optimized for deployment on the Atlas 200I DK A2 hardware platform). Finally, the model is deployed on the Atlas 200I DK A2, which is suitable for UAV applications. During the application phase, the edge device, integrated with a UAV equipped with a camera, enables real-time intelligent detection of protective equipment for live-line operations. This setup effectively supports safety monitoring in high-risk live-line operation scenarios, showcasing its practicality and efficiency.

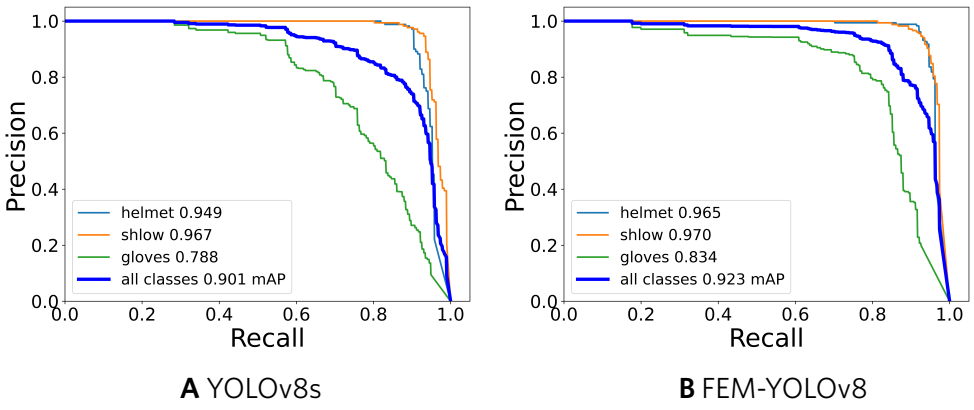

Figure 12. P-R curve of the model on the edge platform. P-R: Precision-recall.

Table 4. Model deployment experiment results

Model	P/%	R/%	mAP/%	Model size/MB	FPS
YOLOv8s	90.3	82.7	90.1	22.0	81
FEM-YOLOv8	**93.1**	**85.9**	**92.3**	**19.6**	**83**

P: Precision; R: recall; mAP: mean average precision; FPS: frames per second.

To assess the performance of the proposed model on the edge platform, both the baseline YOLOv8s model and the improved FEM-YOLOv8 (pt) model developed in this study were converted into an edge-compatible format (om) and deployed on the Atlas 200I DK A2, which supports 10 TOPS of computing power and FP16 precision. The P-R curves of the models are displayed in Figure 12, and the performance results are summarized in Table 4.

As shown in Figure 12, the improved FEM-YOLOv8 model exhibited superior detection performance compared to the original YOLOv8s model, as evidenced by the larger area under the P-R curve. In particular, for the "gloves" category, the mAP value increased by 4.6 percentage points, with other categories also demonstrating varying degrees of improvement. However, by comparing Tables 2 and 4, a slight decrease in mAP was observed during the model conversion process due to network structure fusion, with the mAP of YOLOv8s and the FEM-YOLOv8 model decreasing by 0.2% and 0.1%, respectively. As detailed in Table 4, the proposed FEM-YOLOv8 model achieved P, R, and mAP values of 93.1%, 85.9%, and 92.3%, respectively. These represent improvements of 2.8%, 3.2%, and 2.2% compared to the baseline YOLOv8s model. Additionally, the model size was reduced to 19.6 MB, representing a 10.9% decrease, along with a slight improvement in detection speed. These results confirm the efficiency and outstanding detection capabilities of the improved FEM-YOLOv8 model on edge devices, fully meeting the dual requirements of real-time performance and detection accuracy in resource-constrained environments.

To visually highlight the enhanced detection capabilities of the proposed FEM-YOLOv8 model, both the FEM-YOLOv8 and baseline YOLOv8s models were deployed on the Atlas 200I DK A2 platform for testing. For evaluation, a subset of test images was selected, and the corresponding visualization results are presented in Figure 13.

As illustrated in Figure 13, the FEM-YOLOv8 model demonstrates superior detection performance compared to the baseline YOLOv8s model. The FEM-YOLOv8 model exhibited higher detection confidence and produced no false detections, indicating stronger anti-interference capability, enhanced detection accuracy, and

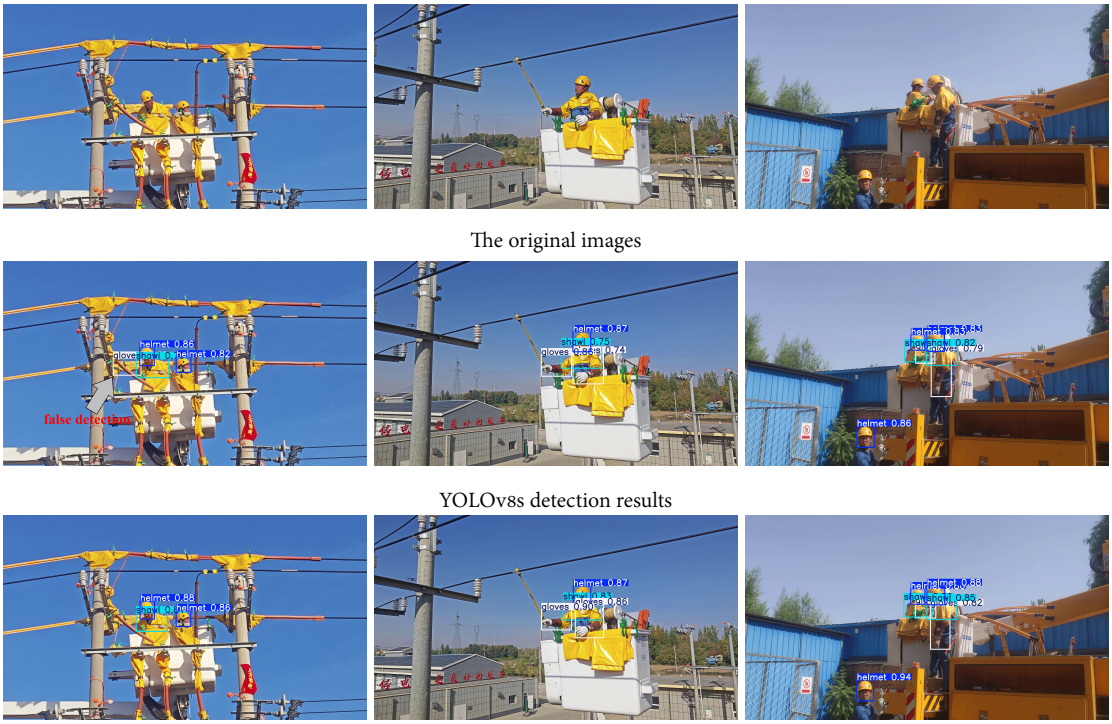

Figure 13. Deployment testing. All images are original, and copyright is retained by the authors.

improved robustness in target detection. The results further validate the effectiveness of the FEM-YOLOv8 algorithm in accurately detecting protective equipment worn by workers, such as insulated helmets, insulated shawls, and insulated gloves. Additionally, in future applications, the FEM-YOLOv8 algorithm can be integrated with UAV platforms and alarm systems to enable real-time monitoring at work sites. If workers are found not wearing the required protective equipment, the system can automatically trigger an alert, providing timely reminders to enhance operational safety. This integration will promote the intelligent development of live-line operations, contributing to improved safety management and digital transformation in the power industry.

6. CONCLUSIONS

In this study, a real-time intelligent detection method for UAV-based live-line operations was proposed, specifically designed to meet the requirements for real-time detection of insulated protective equipment at live-line operation sites. The proposed method is built upon an improved YOLOv8s model, which incorporates FasterBlock modules, ECA attention mechanisms, and the MPDIoU loss function to enhance detection accuracy and robustness. These improvements have been shown to significantly enhance the model's detection performance. The improved FEM-YOLOv8 model was successfully deployed on edge intelligent devices capable of being mounted on UAV platforms. With a computing power of 10 TOPS and a power consumption of just 10.2 W, the system achieves a detection speed of 83 FPS, effectively demonstrating its suitability for real-time intelligent management and control of live-line operation sites. In future work, efforts will focus on further optimizing the model architecture to improve real-time detection performance and increase robustness under complex environmental conditions. Additionally, future work will focus on improving UAV technology by adopting tethered UAVs integrated with advanced sensors, such as high-definition and thermal imaging cameras, to enhance the efficiency and precision of intelligent management and control tasks at live-line operation sites. These improvements are expected to optimize real-time monitoring capabilities and

further contribute to the automation and intelligent development of power grid operations, ensuring greater safety and reliability in power management.

DECLARATIONS

Authors' contributions

Made substantial contributions to the conception and design of the study and performed data analysis and interpretation: Duan, H.

Performed data acquisition and provided administrative, technical, and material support: Shi, F., Gao, B., Zhou, Y., Cui, Q.

Availability of data and materials

The dataset used in this study is self-collected and proprietary, and it is not publicly available. Copyright of the dataset is retained by the authors and their affiliated institutions. Access to the data is not permitted due to confidentiality restrictions.

Financial support and sponsorship

This work was supported by the National Natural Science Foundation of China Regional Innovation and Development Joint Fund (Grant No. U23A20651) and the Science and Technology Project of State Grid Ningxia Electric Power Co., Ltd. (Grant No. 5229DK24000M).

Conflicts of interest

Shi, F. is a Junior Editorial Board Member of the journal *Intelligence & Robotics* but is not involved in any steps of editorial processing, notably including reviewer selection, manuscript handling and decision making. Gao, B. is affiliated with the Electric Power Research Institute of State Grid Ningxia Electric Power Co., Ltd. The other authors declare that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. Deng, D.; Fang, Y.; Zhou, B.; et al. Risk and prevention control analysis of live line operation of distribution network bypass. *J. Phys. Conf. Ser.* **2024**, *2795*, 012008. DOI
2. Zhang, W. Research on safety detection technology of power live Working based on deep learning. 2021. DOI
3. Ye, L. Live working risk assessment analysis and countermeasure. 2017. https://kns.cnki.net/kcms2/article/abstract?v=i12nDYUbXpK_qFT2vWx-FtybcbqytCk7CAtXBz7R639GSYnC9KX1PGgipQFbCB21RSLeBTbnKxDeDNjk09rUX8IPyxJf49C0H-KqTxLmwO4RQQvGNcHxNiiTdXEaxKccn1aagToP8kd6gVnrWAOTiwErk-OeFbBNzULHYWwPcyXbyy2qANi4XHafx8bHKY8WzKe4cwRD-xBnAMLBnPkTJA==&uniplatform=NZKPT&language=CHS. (accessed 14 Jan 2025).
4. Guo, J.; Li, X. Very low-resolution object detection algorithms for electric intelligent safety supervision. *Comput. Eng. Design.* **2020**, *41*, 3188-92. DOI
5. Peng, G.; Lei, Y.; Li, H.; Wu, D.; Wang, J.; Liu, F. CORY-Net: contrastive Res-YOLOv5 network for intelligent safety monitoring on power grid construction sites. *IEEE. Access.* **2021**, *9*, 160461–70. DOI
6. Ma, F.; Wang, B.; Dong, X.; Yao, L.; Wang, H. Safety image interpretation of power industry: basic concepts and technical framework. *Proc. CSEE.* **2022**, *42*, 458-75. DOI
7. Ma, F.; Wang, B.; Dong, X.; Luo, P.; Wang, H.; Zhou, Y. Receptive field vision edge intelligent recognition for ice thickness identification of transmission line. *Power. Syst. Technol.* **2021**, *45*, 2161-9. DOI

8. Xu, L. Y.; Zhao, Y. F.; Zhai, Y. H.; Huang, L. M.; Ruan, C. W. Small object detection in UAV images based on YOLOv8n. *Int. J. Comput. Intell. Syst.* **2024**, *17*, 223. DOI
9. Sindjoung, M. L. F.; Velempini, M.; Djamegni, C. T. A data security and privacy scheme for user quality of experience in a mobile edge computing-based network. *Array.* **2023**, *19*, 100304. DOI
10. Ma, M. Y.; Shen, S. E.; Huang, Y. C. Enhancing UAV visual landing recognition with YOLO's object detection by onboard edge computing. *Sensors.* **2023**, *23*, 8999. DOI
11. Luo, P.; Wang, B.; Wang, H.; Ma, F.; Ma, H.; Wang, L. An ultrasmall bolt defect detection method for transmission line inspection. *IEEE. Trans. Instrum. Meas.* **2023**, *72*, 1-12. DOI
12. Li, Y.; Gao, D. W.; Gao, W.; Zhang, H.; Zhou, J. Double-Mode energy management for multi-energy system via distributed dynamic event-triggered newton-raphson algorithm. *IEEE. Trans. Smart. Grid.* **2020**, *11*, 5339–56. DOI
13. Liu, L. N.; Yang, G. H.; Wasly, S. Distributed predefined-time dual-mode energy management for a microgrid over event-triggered communication. *IEEE. Trans. Ind. Informat.* **2024**, *20*, 3295–305. DOI
14. Li, K.; Qin, L.; Li, Q.; Zhao, F.; Xu, Z.; Liu, K. Improved edge lightweight YOLOv4 and its application in on-site power system work. *Glob. Energy. Intercon.* **2022**, *5*, 168–80. DOI
15. Girshick, R. Fast R-CNN. *arXiv* **2015**, arXiv:1504.08083. Available online: https://doi.org/10.48550/arXiv.1504.08083 (accessed 14 Jan 2025).
16. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE. Trans. Pattern. Anal. Mach. Intell.* **2017**, *39*, 1137–49. DOI
17. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: unified, real-Time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Las Vegas, USA, June 27-30, 2016. IEEE, 2016; pp. 779-88. DOI
18. Redmon, J.; Farhadi, A. YOLO9000: better, faster, stronger. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* Honolulu, USA, July 21-26, 2017. IEEE, 2017; pp. 7263–71. DOI
19. Redmon, J.; Farhadi, A. YOLOv3: an incremental improvement. *arXiv* **2018**, arXiv:1804.02767. Available online: https://doi.org/10.48550/arXiv.1804.02767 (accessed 14 Jan 2025).
20. Bai, Q.; Gao, R.; Li, Q.; Wang, R.; Zhang, H. Recognition of the behaviors of dairy cows by an improved YOLO. *Intell. Robot.* **2024**; *4*:1-19. DOI
21. Zhuang, T.; Liang, X.; Xue, B.; Tang, X. An In-Vehicle real-time infrared object detection system based on deep learning with resource-constrained hardware. *Intell. Robot.* **2024**; *4*:276-92. DOI
22. Liu, W.; Anguelov, D.; Erhan, D.; et al. SSD: single shot multibox detector. In *Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science,* Springer, Cham, 2017; pp. 21-37. DOI
23. Kong, Y.; Wang, W.; Zhang, K.; Qi, Y. Object detection method of electric power site based on improved Mask R-CNN. *Sci. Technol. Eng.* **2020**, *20*, 3134–42. https://kns.cnki.net/kcms2/article/abstract?v=XRdBcB-NO4SX1qe6ikZV25DYo6aOIz8NCQl_QG8Y5aXyV7FLn5heHXsTgzxytFuheInZE0MyXxtj1QEu20RVZmzfh9jpO4Ukmo-7n2TcCdu234pFct3hhdCbyF3uFrZM-QAY_nJEGdHs-jintp3aGIBnquSkxHsOgo1zl4fIzcVfc3n051-D_FLY7AASc9VfjT-AZTs8oss=&uniplatform=NZKPT&language=CHS. (accessed 14 Jan 2025).
24. Lei, X.; Sui, Z. Intelligent fault detection of high voltage line based on the Faster R-CNN. *Measurement* **2019**, *138*, 379–85. DOI
25. Yang, Q.; Ma, S.; Guo, D.; Wang, P.; Lin, M.; Hu, Y. A small object detection method for oil leakage defects in substations based on improved Faster-RCNN. *Sensors* **2023**, *23*, 7390. DOI
26. Han, H.; Xue, X.; Li, Q.; et al. Pig-ear detection from the thermal infrared image based on improved YOLOv8n. *Intell. Robot.* **2024** *4*, 20-38. DOI
27. Yuan, L.; Tang, H.; Chen, Y.; Gao, R.; Wu, W. Improved YOLOv5 for road target detection in complex environments. *Comput. Eng. Appl.* **2023**, *59*, 212-22. https://link.cnki.net/urlid/11.2127.TP.20230607.1151.004 (accessed 2024-12-19).
28. Chang, Z.; Peng, Q.; Chen, Y. Safety supervision method for power operation site based on machine learning and image recognition. *Electric. Power.* **2020**; *53*(04):155-160. https://link.cnki.net/urlid/11.3265.TM.20200403.0900.006 (accessed 14 Jan 2025).
29. Li, H.; Dong, Y.; Liu, Y.; Ai, J. Design and implementation of UAVs for bird's nest inspection on transmission lines based on deep learning. *Drones* **2022**, *6*, 252. DOI
30. Zhao, W.; Jia, M.; Zhai, Y.; Zhao, Z. Detection method for Pin-Losing bolts in transmission lines based on improved YOLOv5s. *J. North. China. Electric. Power. Univ.* **2024** *51*, 92-100. https://link.cnki.net/urlid/13.1212.TM.20220920.0949.002 (accessed 14 Jan 2025).
31. He, M.; Qin, L.; Zhao, F.; et al. Intelligent detection algorithm of security risk management and control for power system on-site operation. *High. Volt. Eng.* **2023**, *49*, 2442-57. DOI
32. Panigrahy, S.; Karmakar, S. Real-time condition monitoring of transmission line insulators using the YOLO object detection model with a UAV. *IEEE. Trans. Instrum. Meas.* **2024** *73*, 1-9. DOI
33. Wang, T.; Liu, R.; Zhang, T.; Wu, Y.; Kong, W. Safety detection of live work in distribution networks based on object detection and tracking. *Electron Eng Prod World* **2024**, *31*, 51–3+68. https://kns.cnki.net/kcms2/article/abstract?v=XRdBcB-NO4SqfKuRSLUWbplAffDbnT3gaSWwRAzeLuCT_H2lxQYzObcTkFM2rT-St_l1qwXkDK7798ZThHRf8q-hgt77UlDdeJlihJK9JDUbxaADEPmZHfDtrJVlHvrfPaJRGGGAL40CoqXog6x76tU9Eq5azvpd-sJ5qp_AfablYHB3bvqDzszdeodJchTK5m6fQEyDCks=&uniplatform=NZKPT&language=CHS (accessed 14 Jan 2025).
34. Li, H. Risk identification technology for power operation site based on deep learning. *Yunnan. Electric. Power.* **2024**, *52*, 46–9. https://kns.cnki.net/kcms2/article/abstract?v=XRdBcB-NO4TtYcdD_A6AfmNzlwJMK3vfEtKcqEOGIQPkQVqenMt1zENl0wr46EnnW1V5j822fAbCwn4upf1Y_hxAlNaOIOvRE4yFaotsyhGZRzToIwvUNaHS-FOzztv_B8r1R1phBQEbM_uPSRaHL7F07Y1Q8Z5HgHd

Cfu-vP_9Fzp4ukXlQJwygF_wRotgZMsKSpJNv38Q=&uniplatform=NZKPT&language=CHS (accessed 14 Jan 2025).

35. Lin, T. Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honululu, USA, July 21-26, 2017; IEEE, 2017; pp. 2117-25. DOI
36. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Sale Lake City, USA, June 18-23, 2018; IEEE, 2018; pp. 8759-68. DOI
37. Zhu, C.; He, Y.; Savvides, M. Feature selective anchor-free module for single-shot object detection. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, USA, June 15-20, 2019; IEEE, 2019; pp. 840-9. DOI
38. Mitiche, I.; Morison, G.; Nesbitt, A.; Hughes-Narborough, M.; Stewart, B. G.; Boreham, P. Classification of partial discharge signals by combining adaptive local iterative filtering and entropy features. *Sensors.* **2018**, *18*, 406. DOI
39. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShuffleNet: an extremely efficient convolutional neural network for mobile devices. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, USA, June 18-23, 2018; IEEE, 2018; pp. 6848-56. DOI
40. Chen, J.; Kao, S. H.; He, H.; et al. Run, don't walk: chasing higher FLOPS for faster neural networks. *arXiv* **2023**, arXiv:2303.03667. Available online: https://doi.org/10.48550/arXiv.2303.03667 (accessed 14 Jan 2025).
41. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-net: efficient channel attention for deep convolutional neural networks. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, USA, June 13-19, 2020; IEEE, 2020; pp. 11534-42. DOI
42. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, USA, June 18-23, 2018; IEEE, 2018; pp. 7132-41. DOI
43. Gao, D.; Chen, T.; Miao, L. Improved road object detection algorithm for YOLOv8n. *Comput. Eng. Appl.* **2024** *60*, 186-97. Available online: https://link.cnki.net/urlid/11.2127.TP.20240528.1330.002 (accessed 14 Jan 2025).
44. Ma, S.; Xu, Y. MPDIoU: a loss for efficient and accurate bounding box regression. *arXiv* **2023**, arXiv:2307.07662. Available online: https://doi.org/10.48550/arXiv.2307.07662 (accessed 14 Jan 2025).