IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 5 May 2023, accepted 17 May 2023, date of publication 19 May 2023, date of current version 1 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3277954

RESEARCH ARTICLE

Foreign Objects Identification of Transmission Line Based on Improved YOLOv7

CHENHUI YU[1], YAKUI LIU[1,2], (Member, IEEE), WANRU ZHANG[1], XUE ZHANG[1], YUHAN ZHANG[1], AND XING JIANG[1]

[1]School of Mechanical and Automotive Engineering, Qingdao University of Technology, Qingdao 266033, China
[2]State Key Laboratory of Electrical Insulation and Power Equipment, Xi'an Jiaotong University, Xi'an 710049, China

Corresponding author: Yakui Liu (liuyakui@qut.edu.cn)

This work was supported in part by the Natural Science Foundation of Shandong Province under Grant ZR2021QE289, in part by the State Key Laboratory of Electrical Insulation and Power Equipment under Grant EIPE22201, and in part by the National Innovation and Entrepreneurship Training Program for College Students in Shandong Province under Grant 202210429107.

ABSTRACT As the grid coverage rises, foreign objects invade more and more frequently, causing grid failures to rise every year. To address this issue, this paper proposes a deep learning-based transmission line unmanned inspection of foreign objects recognition algorithm. The algorithm is based on YOLOv7 (You Only Look Once) algorithm, combining with hyperparameter optimization based on genetic algorithm (GA) and space-to-depth (SPD) convolution to complete the foreign object recognition of transmission line Unmanned Aerial Vehicle (UAV) images. The proposed method can promptly determine and locate these targets' presence in aerial images. Finally, this paper compares the improved YOLOv7 algorithm with other YOLO series algorithms (Faster-rcnn, Centernet, and other target detection models). The comparison results show that the method has the highest Mean Average Precision (mAP) of 92.2% and the Frames Per Second (FPS) of 19 is second only to Centernet. Compared with the unimproved YOLOv7, the average accuracy in the recognition of tower cranes has increased by 11.9%, which is the most obvious improvement in accuracy compared with other detection targets. Meanwhile, the hyperparameter optimization based on genetic algorithm speeds up the convergence of the model.

INDEX TERMS Transmission line, foreign objects, YOLOv7, SPD convolution.

I. INTRODUCTION

In order to meet the increasing electricity demand, most countries are vigorously promoting the construction of high-voltage transmission lines. So, more and more transmission lines are passing through densely populated towns and harsh mountainous areas, which increases the probability of damage to transmission lines while achieving freedom of electricity for residents. To ensure the stability of the power supply, the relevant departments will conduct manual inspections to check whether the grid accessories are replaced [1] and whether there is potential external damage caused by foreign intrusion [2]. With the increase of grid coverage, manual inspection has become a costly and inefficient problem. To solve the above problems, 5G technology [3], robotics [4], transient analysis [5], and deep learning [6] have been applied to the maintenance of electrical equipment. In this paper, an unmanned inspection method of transmission lines based on deep learning has been proposed. The key to this method is the training and deploying of target detection models. In recent years, many scholars have found that the deep learning-based transmission line patrol method has the characteristics of low cost and high accuracy, which is better than the traditional manual patrol method.

The commonly used target detection algorithms in the unmanned inspection of transmission lines based on deep learning are YOLO [7] and Convolutional Neural Networks (CNN) [8]. Literature [9] proposes a Cascade RCNN model for detecting the presence of defects in seismic hammers, which is designed using a deep feature extraction network instead of Resnet and a bi-directional feature fusion network instead of Feature Pyramid Networks (FPN) to improve the

The associate editor coordinating the review of this manuscript and approving it for publication was Giambattista Gruosso.

model's feature extraction and feature fusion capabilities. A YOLOv5 model for insulator defect detection is proposed in the literature [10]. The model is designed with a dynamic weight assignment idea combined with the characteristics of the sample data set to improve the learning ability for complex samples. Literature [11] proposes a multi-task convolutional neural network for detecting transmission line fittings for aerial images with background complexity and variability. The model combines a Region Proposal Network (RPN) with a multi-scale training strategy to improve the mAP on the test dataset. A YOLOv3 dense network model for insulator identification is proposed in the literature [12], and the detection accuracy of the model in a complex background environment is improved by incorporating a multi-feature fusion module and a multi-feature mapping module. Literature [13] proposes a YOLOv3 detection model combined with FPN networks, which reduces the insulator leakage rate while preventing overfitting through improved information utilization and model pruning. Literature [14] proposes an improved Faster-rcnn algorithm for defect detection in transmission line fittings, in which the model recognition accuracy and inference speed are improved by adjusting the convolutional kernel size of CNN and data expansion. Inspired by the human visual system, a Dual-Domain Network (DDnet) model incorporating an Receptive Field Block (RFB) module is proposed in literature [15], which improves the model's perceptual field and small target recognition accuracy by introducing an attention mechanism, and its performance is better than the traditional RetinaNet model. In the above work, primarily deep learning is used for defect detection of transmission line accessories, in addition to transmission line prevention of external damage is also an essential work of unmanned patrol.

Most external damage to transmission lines is caused by foreign object intrusion. An improved YOLOX algorithm is proposed in literature [16] for identifying foreign objects on transmission line corridors. The algorithm first uses a pyramid pooling structure to implement a multi-scale learning approach while introducing an attention mechanism to achieve focused learning of the target. Finally adjusts the loss function and introduces Generalized Intersection over Union (GIOU) loss to make the improved YOLOX outperform Single Shot MultiBox Detector (SSD), Faster R-CNN, YOLOv5, and other algorithms. Literature [17] proposes a transmission line foreign object detection method by combining convolutional neural networks with random forests. The method uses a convolutional neural network to extract target features and a random forest algorithm for target classification. The above changes prove that the approach outperforms individual network model structures such as GoogLeNet, Residual Networks (ResNet), and AlexNet. Literature [18] proposes an Overhead Transmission Line (OTL) classifier that first classifies whether an image is a foreign object using a binary classifier and then uses Faster-rcnn for specific identification of the foreign object. Literature [19] presents an improved bird's nest recognition algorithm for YOLOv3, centered on improving feature fusion. A fusion of high-level semantic features with low-level semantic features is achieved by adding an attentional feature fusion network, which makes the improved YOLOv3 superior to the traditional YOLOv3. Literature [20] focuses on deploying the YOLOv5 algorithm and applying YOLOv5-s to the DJI M300 RTK UAV to solve the problems of poor real-time and low accuracy of the current UAV patrol. Literature [21] applies Faster-rcnn to an unmanned inspection of bird nests on transmission lines, in which k-mean clustering is used to obtain anchor frames with the inclusion of focal loss functions, etc. so that the model achieves high recognition accuracy. In addition to detecting objects such as bird's nests, kites, and plastics, foreign object detection on transmission lines also requires identifying the presence of large engineering vehicles around transmission lines because engineering vehicles can touch and touch the lines during construction. Hence, the literature focuses on the recognition model of YOLOv3 on engineering vehicles [22]. Firstly, Darknet-53 is replaced by Mobilenetv2, and depth-separable convolution is substituted for standard convolution, which reduces the model learning parameters without degrading the model performance and introduces the idea of Fully Convolutional One-Stage (FCOS) to reduce the complexity of the network. The improved YOLOv3 simplifies the network and has a faster inference speed without any loss of accuracy. Target localization is usually used to determine the location of targets in unmanned patrol work. However, the literature has also achieved remarkable results using the semantic segmentation algorithm Mask-rcnn for identifying foreign objects on transmission lines [23].

Many scholars have shown strong interest in deep learning in recent years, and more deep learning algorithms have been proposed, so the target detection algorithms are updated rapidly. A more mature seventh version of the YOLO series has emerged, and the literature has applied YOLOv7 to the farming industry for animal population counting [24]. Literature [25] improves YOLOv7 by adding an attention mechanism to its backbone or head to extract critical features. The algorithm can effectively identify driver behavior and reduce traffic accidents. Literature [26] incorporates the Convolutional Block Attention Module (CBAM) attention mechanism in YOLOv7 and successfully applies it to citrus recognition for an automated picking technique. Literature [27] has significantly improved YOLOv7 by first adding the coordinate attention mechanism module to the backbone network, then using SCYLLA-IoU (SIOU) with the focal loss function to accelerate the model convergence, and finally improving the Non Maximum Suppression (NMS) with the clustered anchor frames to train the dataset.

According to previous studies, the unmanned inspection of transmission lines has attracted the attention of a large number of scholars. And foreign object invasion is also an essential factor in the occurrence of grid failure. In addition, there is a significant gap in the detection of foreign object intrusion with the target detection algorithm being updated so rapidly. Therefore, this paper proposes a YOLOv7 transmission line foreign objects recognition model incorporating SPD convolution and the idea of using genetic algorithm to optimize the model hyperparameters. Where SPD convolution can improve the detection accuracy of the model for low-resolution images as well as small targets, and ultimately improve the overall model recognition accuracy. Meanwhile the use of hyperparameter optimization can accelerate the convergence of the model and shorten the model training time.

From the above literature, an unmanned inspection of transmission lines is mainly based on detecting the presence of faults in accessories. Moreover, there needs to be more research on foreign object invasion, especially in the early warning of large engineering vehicles at present, there needs to be more research. The YOLO algorithm commonly used in the unmanned inspection of transmission lines based on deep learning is mainly based on v3 and v5, and YOLO has now emerged as a more mature v7. Hence, the research on applying YOLOv7 to the unmanned inspection of transmission lines is currently less. The main contributions are summarized as follows,

(1) A deep learning-based algorithm for foreign object recognition on transmission lines is proposed. The algorithm can realize the judgment of whether there is foreign object invasion around the transmission line and the real-time monitoring of engineering vehicles by UAV.

(2) An improved YOLOv7 target detection algorithm is proposed. A new CNN convolution module is added to YOLOv7 to make YOLOv7 better at the detection of small targets. Also, during model training, the hyperparameters are optimized using a GA to speed up the convergence of the model.

(3) The comparison with similar algorithms shows that the improved YOLOv7 outperforms the unimproved YOLOv7, YOLOv5, and Faster-rcnn.

The presented paper is organized as follows. Section II introduces the principles of YOLOv7 algorithm and SPD convolution, as well as the composition of the loss function in YOLOv7. Section III mainly introduces the production of data sets, optimization of hyperparameters and evaluation criteria of model identification results. Section IV focuses on the experimental results and algorithm comparison. Section V is the conclusion.

II. ALGORITHM PRINCIPLE

A. PRINCIPLE OF YOLOv7

YOLOv7 is divided into three parts: Backbone, Neck, and Head. The images to be detected first enter the Backbone part and then undergo three standard convolutional processes, including convolution, batch normalization, and SiLU activation. The MP module is a two-channel structure that performs maximum pooling and convolution on the input feature images and adjusts the number of channels in the whole network structure. The SPPCSPC module is a spatial pyramid pooling structure that can avoid the image distortion problem when scaling the input images, reduce the network structure for repeated extraction of the same features, and save computational costs. The detection image will enter the Neck section for feature fusion after the feature extraction in the Backbone section, which is an FPN network structure. The feature images of different scales will be fused in the bottom-up feature fusion to achieve a multi-scale target detection effect [28]. The entire YOLOv7 network structure is shown in Figure 1.

YOLOv7 is designed concerning the currently popular label-matching strategy, combining the positive and negative sample assignment strategies of YOLOX and YOLOv5. In YOLOv5, nine different sizes of anchors are used to predict the positions of ground truth, so a ground truth often corresponds to multiple anchors in the training session of the model. In order to better distinguish between positive and negative samples, YOLOv5 calculates the ground truth's width and height ratio to each anchor frame. Finally, it records the anchors with the smallest ratio. If the ratio is smaller than the set threshold, the anchor box is defined as a positive sample [29]. YOLOX uses a label-matching strategy called SimOTA. In the field of target detection, the target and background can be considered as the supply side, and the prediction frame can be considered as the demand side. So SimOTA, which solves the optimal transmission problem, can divide positive and negative samples. First, each prediction point is traversed, the presence of the target is determined, and candidate frames are generated. Then the Intersection over Union (IOU) loss and classification loss between candidate frames and ground truth are calculated. For the convenience of calculation, each actual frame is computed with only k candidate frames, and then the computed IOUs are sorted from largest to smallest and rounded cumulatively. The number of positive samples is determined according to the value after rounding, and the candidate frames used as positive samples are determined according to the IOU values [30]. In summary, YOLOv7's label-matching strategy can be summarized as follows,

(1) Determine the positive samples using YOLOv5's positive and negative sample matching strategy.

(2) Calculate the loss of each positive sample with the ground truth.

(3) Determine the number of positive samples k to be assigned.

(4) Determine the minimum k positive samples according to the loss size.

(5) Remove the case that a sample is matched to more than one ground truth.

B. PRINCIPLE OF SPD CONVOLUTION

SPD convolution is a convolutional block for learning low-resolution images with small targets, and the whole process is shown in Figure 2. The whole process of SPD

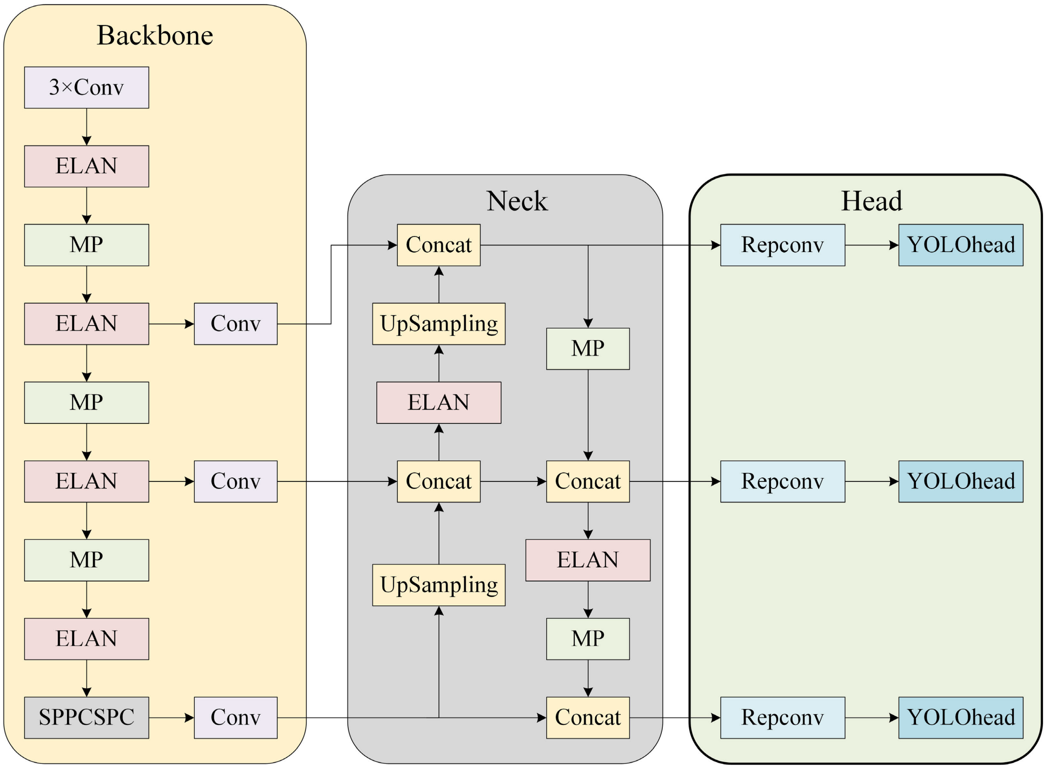


FIGURE 1. YOLOv7 network structure.

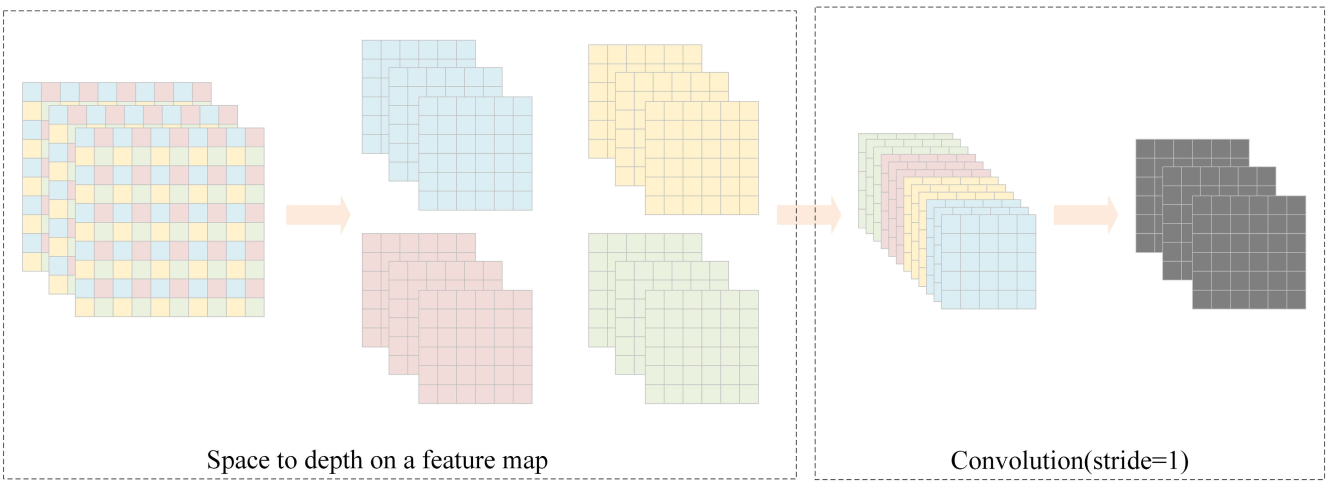


FIGURE 2. SPD convolution.

convolution is divided into two steps: pre-processing of the input image and standard convolution. Figure 2 shows a three-channel image, and the input image is sliced first. After slicing, four sets of sub-feature images are obtained.

The number of channels of the sub-feature images is the same as the input image, and the width and height can be adjusted by parameters that are usually half of the input. The obtained sub-feature images are concatenated along the channel dimension as the input to the standard convolution, which is used with a step size of 1 because of better preservation of all discriminative feature information.

In order to add SPD convolution to YOLOv7, it is decided to modify the ELAN module of YOLOv7. The ELAN module consists of several convolutions, and its structure has reached a saturation state. It is not easy to make significant changes [31], so the final output part of the ELAN module is modified in this paper. At the end of the ELAN module, there will be an operation to concatenate multiple convolution results. Then the concatenated results will be sent to the convolution layer for final processing. Therefore, in this paper, the last standard convolution in the ELAN module is replaced by the SPD convolution, and the modified ELAN module is shown in Figure 3(b).

C. LOSS FUNCTION OF YOLOV7

The loss function L of YOLOv7 consists of three parts, which are Complete Intersection over Union (CIOU) loss function L_{CIOU}, classification loss function L_{cls} and confidence loss function L_{conf}, and the calculation formula can be expressed as (1).

$$L = L_{\mathrm{CLOU}} + L_{\mathrm{cls}} + L_{\mathrm{conf}} \tag{1}$$

The specific formula for L_{CIOU} is as follows:

$$L_{\mathrm{CLOU}} = 1 - \mathrm{IOU} + \frac{\rho^2}{m^2} + \alpha \cdot \nu \tag{2}$$

$$\alpha = \frac{\nu}{1 - \mathrm{IOU} + \nu} \tag{3}$$

$$\nu = \frac{4}{\pi^2}(\arctan\frac{w}{h} - \arctan\frac{\bar{w}}{\bar{h}}) \tag{4}$$

where IOU represents the cross-merge ratio, ρ^2 represents the Euclidean distance between the center point of the predicted box and ground truth, m represents the diagonal distance of the smallest closure area that contains both the predicted box and the ground truth. In (4), (w, h) represents the width and height of ground truth, $(\bar{w}, \bar{h})$ represents the width and height of predicted box. The CIOU loss considers the aspect ratio based on the DIOU loss, which allows the shape of the prediction box to be closer to the ground truth [32].

The formula for calculating the classification loss is shown in (5), as shown at the bottom of the next page. s^2 and B are fixed parameters of 49 and 9, respectively. 1_{ij}^{obj} is a zero-one variable, if the i-th frame falls into the j-th grid, then 1_{ij}^{obj} is 1 otherwise it is 0. c is the classification of the detected target, $p_i(c)$ is the true classification of the target, $\hat{p}_i(c)$ is the predict classification of the target. The formula for calculating the confidence loss is shown in (6), as shown at the bottom of the next page. λ_{coord} is usually taken as 5, (x_i, y_i, r_i) are the center coordinates and radius of the true bounding circle, $(\hat{x}_i, \hat{y}_i, \hat{z}_i)$ are the center coordinates and radius of the predicted bounding circle [33].

III. IDENTIFYING METHOD

A. DATA SETS

In order to better apply the model to the inspection of UAVs, this paper mainly uses UAV aerial images as the dataset of the target detection model. The targets in the dataset that need to be learned by the model are divided into two categories: bird nests and large engineering vehicles. The construction of engineering vehicles is the leading cause of external damage to transmission lines, and bird's nests are the leading representatives of foreign object invasion on transmission lines. Engineering vehicles can be subdivided into cranes, excavators, bulldozers, tower cranes, trucks, and bird nests for detecting six targets. Due to the lack of publicly available transmission line aerial photography datasets, the data images involved in this paper are obtained from the UAV aerial image database. Some of the aerial images in the dataset are shown in Figure 4.

As Fig. 4, illustrates the aerial photography dataset has the problems of a complex environment and severe target occlusion, which requires the high performance of the target detection model to enable it to accurately identify targets in the complex environment. In this paper, the website Make Sense is used to complete the annotation of the dataset and get the coco format tags of the aerial photography dataset. Also, the dataset is divided according to the ratio of 9:1, with 7923 images in the training set and 880 images in the test set. Due to the large size of the tower crane and its more severe obscuration in the aerial photography dataset, the number of training samples for the tower crane was increased in the production of the dataset. Furthermore, the number of samples for each target is shown in Table 1.

B. HYPERPARAMETER OPTIMIZATION

Large models often require pre-set hyperparameters during training, and the setting of hyperparameters often affects the convergence speed and overall performance of the model. There are 9 hyperparameters in the YOLO algorithm, such as learning rate, weight decay coefficient, and momentum, which often need to be obtained by optimization. Therefore, in this paper, the hyperparameters are set as the independent variables, the minimization model loss is set as the objective function, and the genetic algorithm is selected to find the optimization of the hyperparameters. The final optimal hyperparameters are presented in Table 2. Lr0 represents the initial learning rate, and this hyperparameter is used for the update of the model parameters. lrf represents the cyclic learning rate, and this hyperparameter allows the learning rate to increase or decrease regularly within an interval. This parameter can be set to avoid skipping the optimum due to too large learning rate and falling into local optimum due to too small learning rate. Momentum can make the optimization search process moderate and jump out of the local optimum. weight_decay represents the weight decay coefficient, which serves to prevent over-fitting of the model. In the early stage of model training, i.e., the pre-learning stage, the model parameters are more random, thus larger hyperparameters can lead to instability of the model. So, a specific set of hyperparameters is needed in the pre-learning phase. warmup_epochs represents the number of preheat learning, warmup_momentum represents preheat learning momentum, and warmup_bias_lr represents the preheat initial learning rate. iou_t represents the threshold of IoU and anchor_t represents the aspect ratio of anchor.

In order to determine the optimal hyperparameters, the genetic algorithm (GA) is chosen here to complete the

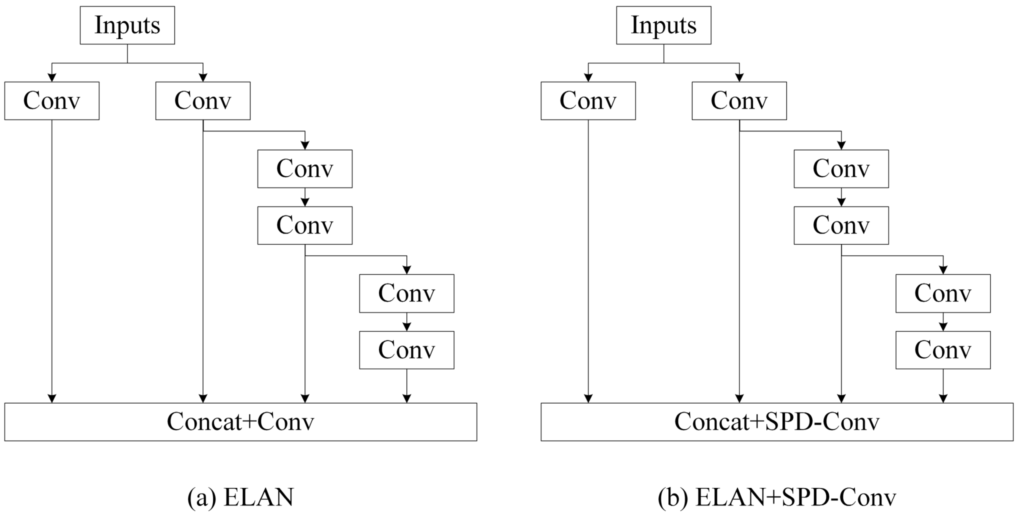


FIGURE 3. ELAN module before and after improvement.

FIGURE 4. UAV aerial photography dataset.

optimization of the hyperparameters. The specific optimization results are shown in Table 2 [34].

C. EVALUATION INDICATORS

Common evaluation metrics used in target detection are precision, recall, and mAP. It can be seen from the definition that precision is inversely proportional to recall so each data set can determine a precision-recall curve based on different IOU thresholds [35]. Taking Figure 5 as an example, the area enclosed by this PR curve and the coordinate axis is the mAP. For a more rigorous representation of the mAP values, the mAP @ 0.5 is usually used to represent the mAP at an IOU threshold of 0.5 [36].

In addition to the accuracy of the model, the recognition speed of the model is also an important indicator. The recognition speed of a model can be expressed as FPS (frames per

$$L_{\text{cls}} = \sum_{i=0}^{s^2} \sum_{j=0}^{B} 1_{ij}^{obj} \sum_{c \in classes} [-p_i(c) \log \hat{p}_i(c) - (1 - p_i(c)) \log(1 - \hat{p}_i(c))] \tag{5}$$

$$L_{\text{conf}} = \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^{B} 1_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^{B} 1_{ij}^{obj} (r_i - \hat{r}_i)^2 \tag{6}$$

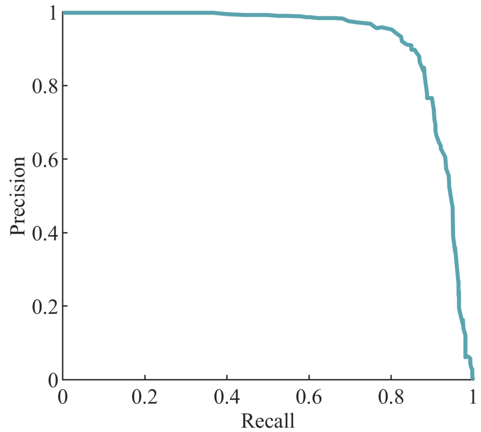


FIGURE 5. PR curve.

second), which is the number of images that the model can process per second [37]. FPS can be computed as follows:

$$FPS = \frac{Num}{Time} \tag{7}$$

where Num represents the number of frames and Time represents the time, which is usually taken as 1 second. This experiment calculates the FPS by recording the time consumed in model recognition [38].

D. EXPERIMENTAL ENVIRONMENT

In order to verify the validity of the model proposed in this paper, this section conducted experiments with the following computer configurations and hyperparameters (shown in Table 3).

E. EXPERIMENTAL PROCESS

First, this paper collects transmission line UAV aerial images and make a dataset, then split the dataset into training and test sets by 9:1. In the second step, the improved YOLOv7 model is used for training, and the training results and weight files are saved after the training. The third step uses the obtained weight files to recognize the aerial images of UAVs obtained in the field. The fourth step compares the recognition results with the unimproved YOLOv7 to determine whether the improved approach is practical. The fifth step compares the improved YOLOv7 algorithm with YOLOv5, Faster-rcnn, and Centernet to determine the best transmission line foreign object recognition algorithm model based on the recognition results [39]. The experimental process is shown in Figure 6.

IV. RESULTS AND COMPARISON

The transmission line aerial photography dataset is put into the improved YOLOv7 (with SPD convolution added) model for training and the loss curves at the end of training are shown in Figure 7. The total loss function in YOLOv7 consists of CIOU loss, classification loss and confidence loss, so the loss curve in Figure 7 is essentially a superposition of three curves with different properties. The CIOU loss reflect

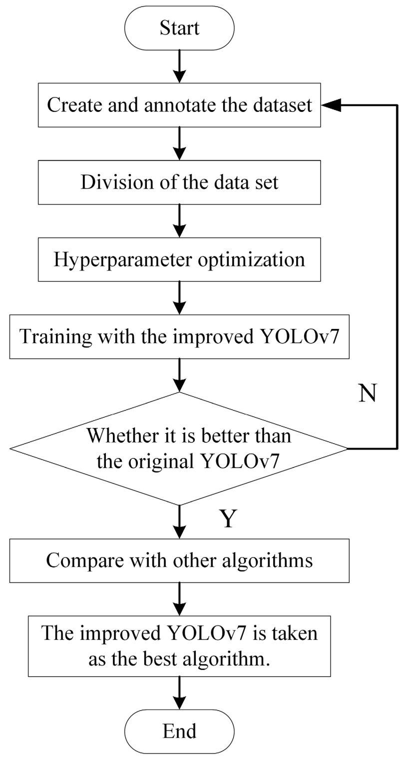


FIGURE 6. Experimental process flow chart.

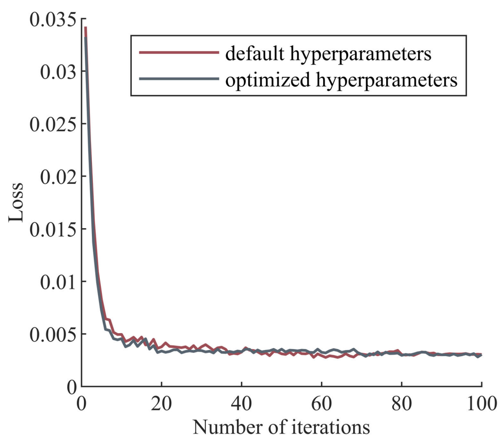


FIGURE 7. Loss function curve.

the error between the predicted and actual boxes. Classification loss is a cross-entropy loss, which reflects the degree of false detection of the network [40]. Confidence loss reflect the probability of whether a target exists in the region of interest to the network.

Figure 7 shows the loss function curves before and after hyperparameter optimization. As can be seen from the figure, the two curves basically overlap after 70 iterations, i.e., converge to the same loss value. The loss values reflect the recognition accuracy of the model, so the optimization of hyperparameters does not improve the model recognition performance significantly. However, after hyperparameter opti-

TABLE 1. Number of samples.

Categories	Crane	Excavator	Truck	Bulldozer	Tower crane	Nest
Quantity	1127	1503	990	1728	2346	1109

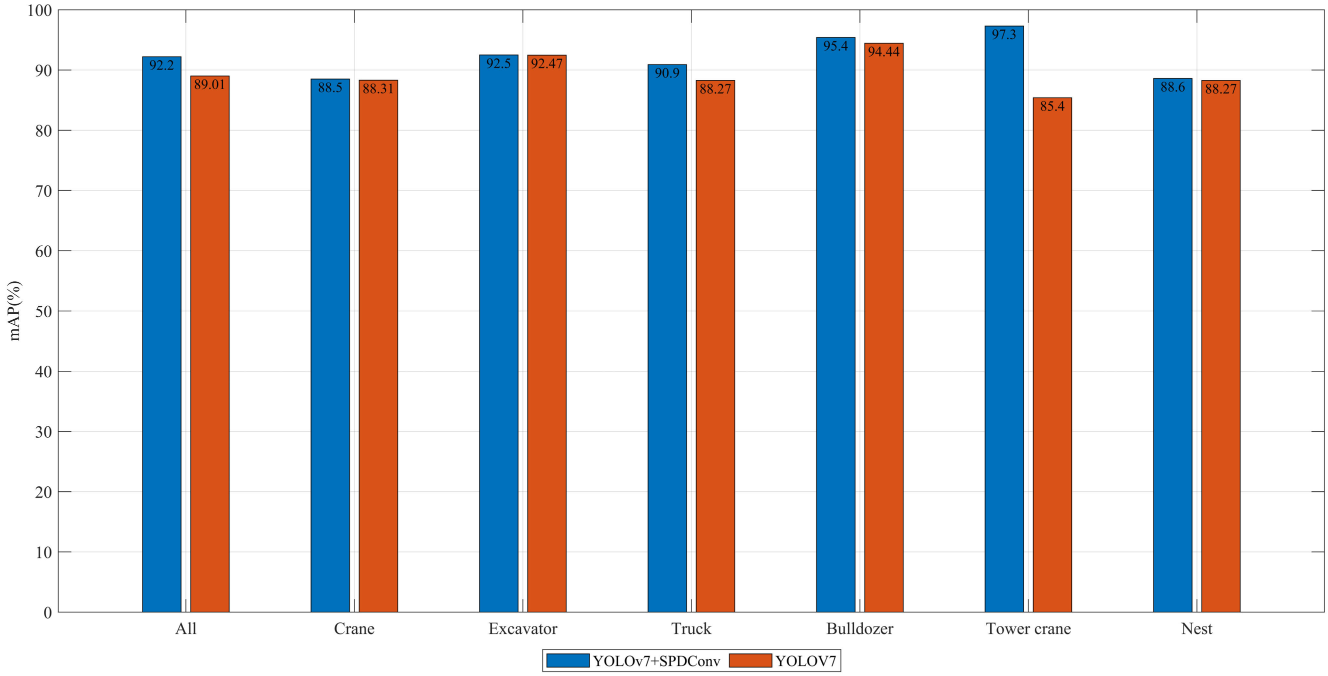


FIGURE 8. Comparison of mAP.

TABLE 2. Hyperparameters after genetic algorithm optimization.

Hyperparameters	Default	Optimization
Lr0	0.01	0.00341
lrf	0.1	0.17006
Momentum	0.937	0.74451
weight_decay	0.0005	0.00018
warmup_epochs	3	3
warmup_momentum	0.8	0.65951
warmup_bias_lr	0.1	0.18819
iou_t	0.2	0.2
anchor_t	4.0	2.7858

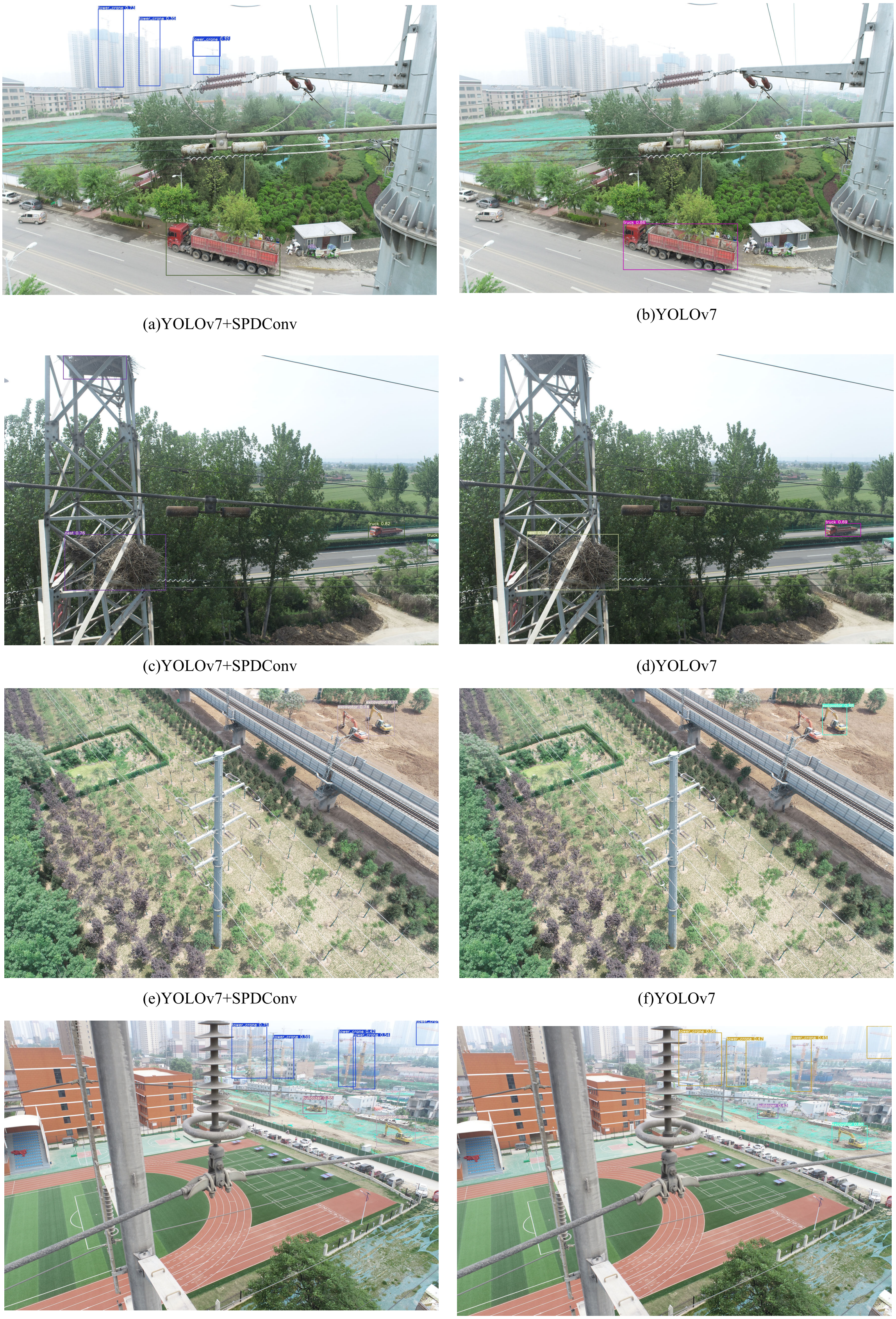
TABLE 3. Experimental environment configuration.

Parameter	Configuration
CPU	AMD Ryzen 7 5800 8-Core Processor
GPU	NVIDIA GeForce RTX 3070 Ti
CUDA	11.3
Pytorch	1.10
Epochs	100
Workers	2
Batch size	4
Img-size	[640, 640]
Optimizer	Adam

mization, the model reaches convergence in about 20 iterations, and the unoptimized model reaches convergence in about 40 iterations. Therefore, optimization of hyperparameters can improve the convergence speed of the model and speed up the training.

YOLOv7 will calculate mAP on the validation set near the end of training. The mAP corresponding to each of the foreign objects detected in this paper is shown in Figure 8. The addition of SPD convolution has improved the correct rate of target detection, as can be seen in Figure 8. The average accuracy of the tower crane improved from 85.4% to 97.3%, a total improvement of 11.9%. The average accuracy of the next truck and bulldozer also improved slightly. The average accuracy of the crane, excavator, and bird's nest changed very little and did not improve significantly. Overall, the average accuracy of the improved YOLOv7 is about 3.19% higher than that of the unmodified one, which proves the effectiveness of the improvement idea in this paper.

Figure 9 compares the recognition effect between the improved YOLOv7 and the original YOLOv7. From (a)(b) of Fig. 9, it can be seen that all the trucks in the immediate area are recognized correctly. However, the improved YOLOv7 recognizes the cranes in the outlying area, indicating that the algorithm has improved the recognition effect of the cranes. From (c)(d) of Fig. 9, it can be seen that the improved YOLOv7 identifies the bird's nest with the truck that only



(a)YOLOv7+SPDConv (b)YOLOv7

(c)YOLOv7+SPDConv (d)YOLOv7

(e)YOLOv7+SPDConv (f)YOLOv7

(g)YOLOv7+SPDConv (h)YOLOv7

FIGURE 9. Comparison of recognition results.

TABLE 4. Comparison of the MAP AND FPS of different algorithms.

Algorithms	mAP	FPS
YOLOv7+SPDConv	92.20%	19
YOLOv7+CBAM+SPPFCSPC	91.10%	15
YOLOv5+BiFPN	84.60%	16
YOLOv4+Densenet	82.47%	14
Faster-rcnn+Resnet	84.67%	6
Centernet	83.92%	29

appears partially in the picture, indicating that the algorithm has improved the recognition of the heavily obscured target. From (e)(f) of Fig. 9, it can be seen that the unimproved YOLOv7 shows the missed detection phenomenon, which indicates that the algorithm improvement reduces the missed detection rate. As seen in (g)(h) of Fig. 9, excavators and bulldozers are correctly identified in the near distance, but the improved YOLOv7 can identify more heavily obscured tower cranes. From the above, YOLOv7, with the addition of SPD convolution, has better recognition results than the original YOLOv7.

To further demonstrate the effectiveness of the algorithm proposed in this paper, the improved YOLOv7 algorithm is compared with YOLOv7+CBAM+SPPFCSPC [41], YOLOv5+CA [42], YOLOv5+BiFPN [43], YOLOv4+Densenet [44], Faster-rcnn+Resnet [45], and Centernet [46] in this paper. The average accuracy of each algorithm and the recognition speed are shown in Table 4. From the table, it can be seen that the average accuracy of the proposed algorithm is higher than other target detection algorithms in terms of recognition accuracy. In terms of recognition speed, Centernet is much faster than other algorithms with 29 frames per second, Faster-rcnnt lags behind other algorithms with 6 frames per second, and the algorithms of YOLO series do not differ much in terms of recognition speed, all of them are around 15 frames per second. The algorithm proposed in this paper can reach 19 frames per second in recognition speed, which is slightly better than other YOLO algorithms. Combining the recognition accuracy and recognition speed can be concluded that the algorithm proposed in this paper is better than other types of algorithms, thus verifying the feasibility of the method in this paper.

Literature [47] proposes a Faster-rcnn based transmission line bird nest identification method, which can identify fewer transmission line foreign objects compared to this paper. Literature [48] proposes a YOLOv5 method for transmission line component detection, while this paper applies the latest YOLO version to the identification of transmission line hazards, which is equivalent to supplementing and improving this literature. While literature [49] applies FCOS to transmission line inspection, this paper's YOLOv7 draws on the idea of FCOS to improve the model recognition accuracy by means of multi-scale target detection. In this paper, SPD convolution is added to YOLOv7, while the hyperparameters are optimized using a genetic algorithm. Overall, the proposed method achieves good results in the identification of foreign objects in transmission lines.

V. CONCLUSION

It is essential to carry out unmanned inspections of transmission lines to prevent external damage and foreign object invasion. In this paper, a transmission line foreign object recognition algorithm is proposed and that can identify large engineering vehicles and bird nests around transmission lines. In order to improve the recognition effect of this model, the ELAN module in YOLOv7 is improved and the model hyperparameters are optimized in this paper. The general convolution is replaced by SPD convolution while retaining the stability of its module performance to improve the model's recognition of low-resolution targets and small targets. And the optimization of the model hyperparameters using genetic algorithms improves the convergence of the model and accelerates the model training. Based on the experimental results, this paper draws the following conclusions.

(1) The YOLOv7 model with SPD convolution added is 3.19% higher than the unmodified YOLOv7 model in recognition accuracy. In the recognition accuracy of the tower crane, the improved YOLOv7 is 11.9% higher than the original YOLOv7, and the improvement effect is most apparent compared with other targets. The recognition results indicate that YOLOv7 improves the recognition accuracy of obscured and overlapping objects through the improvement of the module, especially in the recognition of tower cranes. For targets other than tower cranes, the improved YOLOv7 is slightly higher in accuracy than the unmodified YOLOv7. The model with default hyperparameters converges in about 40 iterations, and the model converges in about 20 iterations after hyperparameter optimization. The convergence speed of YOLOv7 with hyperparameter optimization is significantly improved.

(2) The algorithm proposed in this paper is compared with YOLOv7, which adds an attention mechanism, YOLOv5, which replaces the feature fusion network, YOLOv4, which combines with Densenet, Fast-rcnn, which uses Resnet as a feature extraction network; and Centernet, which is based on anchor free detection. The results show that the mean average precision of this paper's algorithm is 92.20%, which is higher than the other five algorithms. In terms of recognition speed, this paper's algorithm has an FPS of 19, which is second only to Centernet.

The presented paper has achieved specific results in the unmanned inspection of transmission lines based on deep learning, but the research is not perfect and need further explore. For example, it is difficult to achieve real-time inspection work of the whole line just by relying on UAV. The range of the UAV is also considered when this target detection method is combined with the UAV. If the range is not stable, real-time monitoring will not be possible. In future research, the YOLOv7 model structure will be better optimized, and better results will be achieved in transmission line unmanned patrol.

REFERENCES

[1] W. Xu, X. Zhong, M. Luo, L. Weng, and G. Zhou, "End-to-end insulator string defect detection in a complex background based on a deep learning model," *Frontiers Energy Res.*, vol. 10, pp. 1–11, Jul. 2022, doi: 10.3389/fenrg.2022.928162.

[2] J. Zhu, Y. Guo, F. Yue, H. Yuan, A. Yang, X. Wang, and M. Rong, "A deep learning method to detect foreign objects for inspecting power transmission lines," *IEEE Access*, vol. 8, pp. 94065–94075, 2020, doi: 10.1109/ACCESS.2020.2995608.

[3] L. Guo, C. Ye, Y. Ding, and P. Wang, "Allocation of centrally switched fault current limiters enabled by 5G in transmission system," *IEEE Trans. Power Del.*, vol. 36, no. 5, pp. 3231–3241, Oct. 2021, doi: 10.1109/TPWRD.2020.3037193.

[4] X. Hou, L. Zhang, Y. Su, G. Gao, Y. Liu, Z. Na, Q. Xu, T. Ding, L. Xiao, L. Li, and T. Chen, "A space crawling robotic bio-paw (SCRBP) enabled by triboelectric sensors for surface identification," *Nano Energy*, vol. 105, Jan. 2023, Art. no. 108013, doi: 10.1016/j.nanoen.2022.108013.

[5] C. Guo, C. Ye, Y. Ding, and P. Wang, "A multi-state model for transmission system resilience enhancement against short-circuit faults caused by extreme weather events," *IEEE Trans. Power Del.*, vol. 36, no. 4, pp. 2374–2385, Aug. 2021, doi: 10.1109/TPWRD.2020.3043938.

[6] P. Luo, B. Wang, H. Wang, F. Ma, H. Ma, and L. Wang, "An ultra-small bolt defect detection method for transmission line inspection," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023, doi: 10.1109/TIM.2023.3241994.

[7] T. Renwei, Z. Zhongjie, B. Yongqiang, G. Ming, and G. Zhifeng, "Key parts of transmission line detection using improved YOLO v3," *Int. Arab J. Inf. Technol.*, vol. 18, no. 6, pp. 747–754, 2021, doi: 10.34028/iajit/18/6/1.

[8] L. Yang, J. Fan, S. Xu, E. Li, and Y. Liu, "Vision-based power line segmentation with an attention fusion network," *IEEE Sensors J.*, vol. 22, no. 8, pp. 8196–8205, Apr. 2022, doi: 10.1109/JSEN.2022.3157336.

[9] F. Zhou, G. Wen, G. Qian, Y. Ma, H. Pan, J. Liu, and J. Li, "A high-efficiency deep-learning-based antivibration hammer defect detection model for energy-efficient transmission line inspection systems," *Int. J. Antennas Propag.*, vol. 2022, pp. 1–12, Sep. 2022, doi: 10.1155/2022/3867581.

[10] Y. Li, G. Zou, H. Zou, C. Zhou, and S. An, "Insulators and defect detection based on the improved focal loss function," *Appl. Sci.*, vol. 12, no. 20, p. 10529, Oct. 2022, doi: 10.3390/app122010529.

[11] H. Zhang, L. Wu, Y. Chen, R. Chen, S. Kong, Y. Wang, J. Hu, and J. Wu, "Attention-guided multitask convolutional neural network for power line parts detection," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–13, 2022, doi: 10.1109/TIM.2022.3162615.

[12] C. Liu, Y. Wu, J. Liu, and Z. Sun, "Improved YOLOv3 network for insulator detection in aerial images with diverse background interference," *Electronics*, vol. 10, no. 7, pp. 1–20, 2021, doi: 10.3390/electronics10070771.

[13] X. Zhang, Y. Zhang, J. Liu, C. Zhang, X. Xue, H. Zhang, and W. Zhang, "InsuDet: A fault detection method for insulators of overhead transmission lines using convolutional neural networks," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021, doi: 10.1109/TIM.2021.3120796.

[14] X. Zheng, R. Jia, Aisikaer, L. Gong, G. Zhang, and J. Dang, "Component identification and defect detection in transmission lines based on deep learning," *J. Intell. Fuzzy Syst.*, vol. 40, no. 2, pp. 3147–3158, Feb. 2021, doi: 10.3233/JIFS-189353.

[15] Y. Gong, W. Zhou, K. Wang, J. Wang, R. Wang, H. Deng, and G. Liu, "Defect detection of small cotter pins in electric power transmission system from UAV images using deep learning techniques," *Electr. Eng.*, vol. 105, no. 2, pp. 1251–1266, Apr. 2023, doi: 10.1007/s00202-022-01729-8.

[16] M. Wu, L. Guo, R. Chen, W. Du, J. Wang, M. Liu, X. Kong, and J. Tang, "Improved YOLOX foreign object detection algorithm for transmission lines," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–10, Oct. 2022, doi: 10.1155/2022/5835693.

[17] Y. Yu, Z. Qiu, H. Liao, Z. Wei, X. Zhu, and Z. Zhou, "A method based on multi-network feature fusion and random forest for foreign objects detection on transmission lines," *Appl. Sci.*, vol. 12, no. 10, p. 4982, May 2022, doi: 10.3390/app12104982.

[18] F. Zhang, Y. Fan, T. Cai, W. Liu, Z. Hu, N. Wang, and M. Wu, "OTL-classifier: Towards imaging processing for future unmanned overhead transmission line maintenance," *Electronics*, vol. 8, no. 11, p. 1270, Nov. 2019, doi: 10.3390/electronics8111270.

[19] J. Zhang, Q. Qi, H. Zhang, Q. Du, Z. Guo, and Y. Tian, "Detection of bird's nest on transmission lines from aerial images based on deep learning model," *Int. J. Innov. Comput. Inf. Control*, vol. 18, no. 6, pp. 1755–1768, 2022, doi: 10.24507/ijicic.18.06.1755.

[20] H. Li, Y. Dong, Y. Liu, and J. Ai, "Design and implementation of UAVs for bird's nest inspection on transmission lines based on deep learning," *Drones*, vol. 6, no. 9, p. 252, Sep. 2022, doi: 10.3390/drones6090252.

[21] J. Li, D. Yan, K. Luan, Z. Li, and H. Liang, "Deep learning-based bird's nest detection on transmission lines using UAV imagery," *Appl. Sci.*, vol. 10, no. 18, p. 6147, Sep. 2020, doi: 10.3390/app10186147.

[22] H. Li, L. Liu, J. Du, F. Jiang, F. Guo, Q. Hu, and L. Fan, "An improved YOLOv3 for foreign objects detection of transmission lines," *IEEE Access*, vol. 10, pp. 45620–45628, 2022, doi: 10.1109/ACCESS.2022.3170696.

[23] W. Chen, Y. Li, and C. Li, "A visual detection method for foreign objects in power lines based on mask R-CNN," *Int. J. Ambient Comput. Intell.*, vol. 11, no. 1, pp. 34–47, Jan. 2020, doi: 10.4018/IJACI.2020010102.

[24] K. Jiang, T. Xie, R. Yan, X. Wen, D. Li, H. Jiang, N. Jiang, L. Feng, X. Duan, and J. Wang, "An attention mechanism-improved YOLOv7 object detection algorithm for hemp duck count estimation," *Agriculture*, vol. 12, no. 10, p. 1659, Oct. 2022, doi: 10.3390/agriculture12101659.

[25] S. Liu, Y. Wang, Q. Yu, H. Liu, and Z. Peng, "CEAM-YOLOv7: Improved YOLOv7 based on channel expansion and attention mechanism for driver distraction behavior detection," *IEEE Access*, vol. 10, pp. 129116–129124, 2022, doi: 10.1109/ACCESS.2022.3228331.

[26] J. Chen, H. Liu, Y. Zhang, D. Zhang, H. Ouyang, and X. Chen, "A multiscale lightweight and efficient model based on YOLOv7: Applied to citrus orchard," *Plants*, vol. 11, no. 23, p. 3260, Nov. 2022, doi: 10.3390/plants11233260.

[27] J. Zheng, H. Wu, H. Zhang, Z. Wang, and W. Xu, "Insulator-defect detection algorithm based on improved YOLOv7," *Sensors*, vol. 22, no. 22, pp. 1–23, 2022, doi: 10.3390/s22228801.

[28] K. Patel, C. Bhatt, and P. L. Mazzeo, "Improved ship detection algorithm from satellite images using YOLOv7 and graph neural network," *Algorithms*, vol. 15, no. 12, p. 473, Dec. 2022, doi: 10.3390/a15120473.

[29] J. Liu, W. Qiao, and Z. Xiong, "OAB-YOLOv5: One-anchor-based YOLOv5 for rotated object detection in remote sensing images," *J. Sensors*, vol. 2022, pp. 1–11, Dec. 2022, doi: 10.1155/2022/8515510.

[30] G. Wang, Z. Liu, H. Sun, C. Zhu, and Z. Yang, "Yolox-BTFPN: An anchor-free conveyor belt damage detector with a biased feature extraction network," *Measurement*, vol. 200, Aug. 2022, Art. no. 111675, doi: 10.1016/j.measurement.2022.111675.

[31] Z. Yang, C. Ni, L. Li, W. Luo, and Y. Qin, "Three-stage pavement crack localization and segmentation algorithm based on digital image processing and deep learning techniques," *Sensors*, vol. 22, no. 21, p. 8459, Nov. 2022, doi: 10.3390/s22218459.

[32] C. Liu, Y. Wu, J. Liu, Z. Sun, and H. Xu, "Insulator faults detection in aerial images from high-voltage transmission lines based on deep learning model," *Appl. Sci.*, vol. 11, no. 10, p. 4647, May 2021, doi: 10.3390/app11104647.

[33] R. Gai, N. Chen, and H. Yuan, "A detection algorithm for cherry fruits based on the improved YOLO-v4 model," *Neural Comput. Appl.*, vol. 35, no. 19, pp. 13895–13906, Jul. 2023, doi: 10.1007/s00521-021-06029-z.

[34] T. Zeng, S. Li, Q. Song, F. Zhong, and X. Wei, "Lightweight tomato real-time detection method based on improved YOLO and mobile deployment," *Comput. Electron. Agricult.*, vol. 205, Feb. 2023, Art. no. 107625, doi: 10.1016/j.compag.2023.107625.

[35] Z. Yang, C. Zhao, H. Maeda, and Y. Sekimoto, "Development of a large-scale roadside facility detection model based on the mapillary dataset," *Sensors*, vol. 22, no. 24, pp. 1–12, 2022, doi: 10.3390/s22249992.

[36] M. Hussain, H. Al-Aqrabi, M. Munawar, R. Hill, and T. Alsboui, "Domain feature mapping with YOLOv7 for automated edge-based pallet racking inspections," *Sensors*, vol. 22, no. 18, p. 6927, Sep. 2022, doi: 10.3390/s22186927.

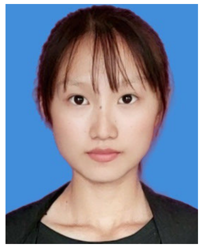
[37] A. Zainab and D. Syed, "Deployment of deep learning models on resource-deficient devices for object detection," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 73–78, doi: 10.1109/ICIoT48696.2020.9089651.

[38] L. Zhang, M. Wang, K. Liu, M. Xiao, Z. Wen, and J. Man, "An automatic fault detection method of freight train images based on BD-YOLO," *IEEE Access*, vol. 10, pp. 39613–39626, 2022, doi: 10.1109/ACCESS.2022.3165835.

[39] Z. Chen, R. Wu, Y. Lin, C. Li, S. Chen, Z. Yuan, S. Chen, and X. Zou, "Plant disease recognition model based on improved YOLOv5," *Agronomy*, vol. 12, no. 2, p. 365, Jan. 2022, doi: 10.3390/agronomy12020365.
[40] A. Malta, M. Mendes, and T. Farinha, "Augmented reality maintenance assistant using YOLOv5," *Appl. Sci.*, vol. 11, no. 11, pp. 1–14, 2021, doi: 10.3390/app11114758.
[41] J. Yan, Z. Zhou, D. Zhou, B. Su, Z. Xuanyuan, J. Tang, Y. Lai, J. Chen, and W. Liang, "Underwater object detection algorithm based on attention mechanism and cross-stage partial fast spatial pyramidal pooling," *Frontiers Mar. Sci.*, vol. 9, pp. 1–15, Nov. 2022, doi: 10.3389/fmars.2022.1056300.
[42] X. Xu, X. Tian, and F. Yang, "A retrieval and ranking method of mathematical documents based on CA-YOLOv5 and HFS," *Math. Biosci. Eng.*, vol. 19, no. 5, pp. 4976–4990, 2022, doi: 10.3934/mbe.2022233.
[43] F. Hu, P. Qian, Y. Jiang, and J. Yao, "An improved waste detection and classification model based on YOLOV5," in *Proc. Int. Conf. Intell. Comput.*, in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 13395, 2022, pp. 741–754, doi: 10.1007/978-3-031-13832-4_61.
[44] A. M. Roy, R. Bose, and J. Bhaduri, "A fast accurate fine-grain object detection model based on YOLOv4 deep neural network," *Neural Comput. Appl.*, vol. 34, no. 5, pp. 3895–3921, Mar. 2022, doi: 10.1007/s00521-021-06651-x.
[45] Q. Ma, G. Du, Z. Yu, H. Yuan, and X. Wei, "Classification of damage types in liquid-filled buried pipes based on deep learning," *Meas. Sci. Technol.*, vol. 34, no. 2, Feb. 2023, Art. no. 025010, doi: 10.1088/1361-6501/ac9b7b.
[46] J. Zhou, Z. Chen, and X. Huang, "Weakly perceived object detection based on an improved CenterNet," *Math. Biosci. Eng.*, vol. 19, no. 12, pp. 12833–12851, 2022, doi: 10.3934/mbe.2022599.
[47] F. Li, J. Xin, T. Chen, L. Xin, Z. Wei, Y. Li, Y. Zhang, H. Jin, Y. Tu, X. Zhou, and H. Liao, "An automatic detection method of bird's nest on transmission line tower based on Faster_RCNN," *IEEE Access*, vol. 8, pp. 164214–164221, 2020, doi: 10.1109/ACCESS.2020.3022419.
[48] W. Bao, X. Du, N. Wang, M. Yuan, and X. Yang, "A defect detection method based on BC-YOLO for transmission line components in UAV remote sensing images," *Remote Sens.*, vol. 14, no. 20, p. 5176, Oct. 2022, doi: 10.3390/rs14205176.
[49] L. Zhao, C. Liu, Z. Zhang, and H. Qu, "Transmission line object detection method based on label adaptive allocation," *Mathematics*, vol. 10, no. 12, p. 2150, Jun. 2022, doi: 10.3390/math10122150.

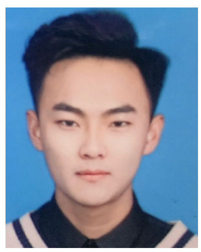
CHENHUI YU was born in Weifang, Shandong, China, in 2002. He is currently pursuing the bachelor's degree with the Qingdao University of Technology. During his bachelor's studies, he has been involved in research related to deep learning-based transmission line inspection.

YAKUI LIU (Member, IEEE) was born in Jining, Shandong, China, in 1990. He received the B.S. degree in electrical engineering from Linyi University, Linyi, China, in 2013, the M.S. degree in smart grid engineering from the Qilu University of Technology, Jinan, China, in 2016, and the Ph.D. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2020. He is currently an Associate Professor with the School of Mechanical and Automatic Engineering, Qingdao University of Technology, Qingdao, China. His research interests include fault diagnosis and artificial intelligence.



WANRU ZHANG was born in Xintai, Tai'an, China, in 2002. She is currently pursuing the bachelor's degree with the Qingdao University of Technology. During her bachelor's studies, she has been involved in research related to deep learning-based transmission line inspection.

XUE ZHANG was born in Liaocheng, Shandong, China, in 2002. She is currently pursuing the bachelor's degree with the Qingdao University of Technology. During her bachelor's studies, she has been involved in research related to deep learning-based transmission line inspection.



YUHAN ZHANG was born in Zibo, Shandong, China, in 2002. He is currently pursuing the bachelor's degree with the Qingdao University of Technology. During his bachelor's studies, he has been involved in research related to deep learning-based transmission line inspection.

XING JIANG was born in Yantai, Shandong, China, in 2002. He is currently pursuing the bachelor's degree with the Qingdao University of Technology. During his bachelor's studies, he has been involved in research related to deep learning-based transmission line inspection.