This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3104731, IEEE Access

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Detection and Location of Safety Protective Wear in Power Substation Operation Using Wear-enhanced YOLOv3 Algorithm

BAINING ZHAO[1], HAIJUAN LAN[2], ZHEWEN NIU[1], HUILING ZHU[1], TONG QIAN[1] AND WENHU TANG.[1],(Senior Member, IEEE)
[1]School of Electric Power Engineering, South China University of Technology, Guangzhou 510641, China
[2]Schneider Electric (Xiamen) Switchgear co. Ltd, Xiamen 361000, China

Corresponding author: Haijuan Lan (e-mail: haijuan_lan@163.com).

This work was supported by the National Natural Science Foundation of China (No.51977082).

ABSTRACT Wearing personal safety protective equipment (PSPE) plays a key role in reducing electrical injuries to electrical workers. However, due to the lack of safety awareness, operators often do not wear PSPE when carrying out inspection or maintenance projects in substations, which is the main reason for personal injury accidents. Therefore, it is necessary to detect the wearing of PSPE in real-time through a video surveillance system. In this paper, a wear-enhanced YOLOv3 method for real-time detection of PSPE wear of substation operators is proposed. In order to improve the detection accuracy, the gamma correction is applied as the preprocessing method to highlight the details of the operators. Besides, K-means++ algorithm is introduced to get the most suitable prior bounding box size to improve the detection speed. Based on the proposed method, it can quickly and effectively detect whether the substation operators are wearing safety helmets and insulating gloves and boots correctly. Finally, extensive experiments are carried out using a dataset of real substation monitoring images to illustrate the effectiveness of the proposed method for real-time PSPE wear detection.

INDEX TERMS objects detection, deep learning, YOLOv3, power substation, safety helmet

I. INTRODUCTION

Power substations play a key role in the voltage conversion, power concentration and distribution in power systems [1]. The continuous expansion of power systems means the increasing number and scale of substations. The long-term safe and stable operation of substation is necessary for the safe and stable operation of power systems [2].

At present, despite the continuous development of intelligent power grid, the operation of the substation still needs regular inspection and maintenance [3]. In order to ensure the safe and stable operation of substation, it is necessary for operators to inspect and maintain the substation regularly [4]. However, in the actual inspection and maintenance process of fields, due to lack of sufficient attention and lack of safety awareness, operators tend to ignore the operation rules and regulations, leading to various violations [5]. For example, not wearing safety helmets, not wearing insulated gloves and boots, crossing safety barriers, etc. These violations are main reasons for power grid accidents. As a conclusion, it is necessary to monitor and control these common violations in real time through a video surveillance system [6]. Failure to wear personal safety protection equipment (PSPE) e.g., helmet, insulating gloves and boots, in a dangerous area of a substation is a typical violation, which will cause great hidden dangers to operators and electrical equipment. Therefore, the use of object detection methods to recognize the wearing of PSPE of substation operators has significant engineering practical value in substation operations.

Currently, mainstream safety wear detection methods are basically divided into two types: 1) traditional methods based on image processing; 2) object detection methods based on deep learning. The core of the traditional method is to extract skin color, head, and face information through image processing technology [7]–[9]. For instance, Waranusast et al.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3104731, IEEE Access

[10] developed an automatically detect system for safety helmet based on K-Nearest-Neighbor (KNN). Li et al. [11] exploited the Hough circular transformation to determine the shape of safety helmet and use the extracted Histogram of Oriented Gradients (HOG) features to train a support vector machine (SVM), which could accurately detect safety helmet. However, traditional detection methods for helmets often do not consider the impact of the complex substation operating environment [12]. The detection result is susceptible to light, small and medium objects, which leads to a high false alarm rate.

In recent years, deep learning has made a breakthrough in image classification and object detection, and becomes the most effective automatic feature learning method [13]. Deep learning refers to such a kind of algorithm, which feeds a certain amount of input and output data to the neural network that consists of multiple layers. Through repeated training, the network can learn the mapping relationship between the current input and output data [14]. The development of deep learning provides a new idea for image object detection tasks in power systems [15]–[17]. As deep learning-based methods to detect objects, Faster-RCNN [18], [19], Single Shot Multi-Box Detector (SSD) [20], deconvolutional single shot detector (DSSD) [21], RetinaNet [22] and You Only Look Once (YOLO) [23], [24] have shown their advantages in object detection tasks for power systems. In [25], an improved Faster-RCNN model was proposed to detect the coordinates, orientation angle, and class type of individual equipment parts. In [26], a safety helmet detection method for substation workers was proposed based on Faster-RCNN and obtained a detection accuracy as 90%. In [27], online hard example mining was combined with a multi-part detection method to identify whether a worker is wearing a safety helmet. Bounding-box regression and transfer learning was used to improve the convolutional neural network-based face detection for safety helmet detection in [28].

Among the detection methods discussed above, the accuracy of YOLOv3 [29] is slightly better than that of SSD and slightly inferior to that of Faster-RCNN. However, the speed of YOLOv3 is at least twice as fast as SSD and Faster-RCNN. It shows that the helmet wearing detection algorithm based on YOLOv3 increased the feature map scale, optimized the prior dimensional algorithm of specific helmet dataset to accurately detect whether the helmet is worn by the standard in [30]. In [31], a safety helmet wearing detection method based on the YOLOv3 algorithm was proposed, which met the real-time performance of the detection task. It means that YOLOv3 is a promising tool for detecting and locating PSPE wear during the operation of power substation.

Most above-mentioned methods in power systems were designed for the detection of safety helmets and few papers focused on complete PSPE detection. In [32], three deep learning models built on the YOLO architecture to verify PSPE compliance of workers. However, the mean average precision is less than 75% and the detection speed can be further improved. Moreover, the real-time PSPE detection tasks in substations have their own specific characteristics: 1) The background of the substation is complex and the cameras are distributed far away, resulting in a low resolution and intensity contrast of collected video images; 2) The small size of the PSPE in the substation scene is less distinguishable from the background; 3) Real-time monitoring requires a high processing speed for the detection model. Although the analysis of the aforementioned methods indicates a satisfactory performance of objects detection methods based on deep learning, there is still a lack of investigation of AI-based techniques for real-time PSPE detection in the field of online surveillance for substations. Therefore, such simple applications of the deep learning method can not meet the requirements, which requires fast and accurate detection of PSPE for substation workers.

To address the above mentioned problems, this paper proposes a wear-enhanced YOLOv3 method for real-time detection of PSPE wear of substation operators. The gamma correction is applied as the preprocessing method to highlight the details of the operators and improve the detection accuracy. Besides, in order to improve the detection speed, K-means++ algorithm is introduced to get the most suitable prior bounding box size. Based on the proposed method, it is able to quickly and effectively realize whether the substation operators are wearing safety helmets and insulating gloves and boots correctly. Finally, extensive experiments are carried out using a dataset of real substation monitoring images to illustrate the effectiveness of the proposed method for real-time PSPE wear detection purposes. The main contributions of this paper are illustrated as follows.

1) The detection of PSPE including safety helmet, insulating gloves, and insulating boots in substation is considered simultaneously using a data-driven method. To the authors' knowledge, this is the first research identifying multi-class PSPE of power substations.
2) A new object detection method namely wear-enhanced YOLOv3 is proposed. The images and videos are preprocessed using gamma correction and the prior box is selected by K-means++ according to the PSPE characteristics.
3) Different methods are used to compared with the proposed method in detection performance. Extreme conditions of PSPE are considered to verify the effectiveness of the proposed wear-enhanced YOLOv3.

The rest of this paper is organized as follows. Section II is the description of the proposed real-time PSPE wear detection method for substation operators. The results of experiments and the discussion are presented in Section III. Finally, Section IV summarizes the conclusion of this paper and the future research plan is given as well.

II. DETECTION OF PSPE WEAR BASED ON WEAR-ENHANCED YOLOV3

To address the complex background and changeable angle of substation monitoring images, this paper proposes a new comprehensive detection and location model of substation operators' PSPE wear. Different from the single objective model, the proposed method can comprehensively detect the wearing condition of safety helmet, insulating boots and gloves. The proposed method consists of four stages: preprocessing, data enhancement, wear-enhanced YOLOv3 (WEYOLOv3) model training, and transfer learning, as shown in Fig. 1.

A. PREPROCESSING

First, in order to filter the irrelevant information in the image, make the neural network more focused on the extraction of useful information, and accelerate the training speed of the network, the image set of PSPE wear is preprocessed.

In the field of computer graphics, the conversion curve between the screen output voltage and the corresponding brightness is called gamma curve. The gamma correction is to edit the gamma curve of the image, redact the nonlinear tone of the image, detect the dark part and light part of the image signal, and increase the proportion of the two, so as to improve the image contrast effect. The basic flow of the algorithm is as follows.

1) Normalization: converts the pixel value to the real number between 0 and 1. In other words, the algorithm is as follows: $(I + 0.5)/256$, where I is the original pixel value.
2) Precompensation: according to the formula, the corresponding value of the normalized pixel data with $1/gamma$ as index is obtained.
3) Denormalization: the real value after precompensation is reversed to the integer value between 0 and 255.

B. DATA ENHANCEMENT

Because the amount of data about substation operation is small, this paper utilizes the online data enhancement technology. It can enhance the generalization ability of the network and avoid the problem of overfitting in the learning process of the network. Three kinds of data enhancement methods are used in the WEYOLOv3.

1) Random scaling: reduce the images of substation operation randomly, and fill the spare part with gray.
2) Random translation and flipping: randomly move the images in the left and right directions, and randomly flip the images in the left and right directions.
3) HSV adjustment: convert the image to HSV space, and randomly adjust the exposure, saturation and brightness.

C. WEAR-ENHANCED YOLOV3

1) Detection and location network based on YOLOv3 algorithm

YOLOv3 algorithm belongs to the single-stage object detection algorithm. It outputs the position of the object boundary box through a neural network to realize the object detection and location. The basic structure of the network is shown in Fig. 1, which mainly includes two parts: 1) feature extraction: using Darknet-53 basic network to extract the features of the input image. Based on the idea of residual network ResNet, the network sets up a fast link, which consists of two convolution layers and a fast link to form a residual component, so as to avoid the problems of gradient disappearance and gradient explosion, and ensure the learning performance of deep neural network [19]. 2) Detection: with the deepening of the network, part of the image feature information is lost, which is not conducive to the detection of small objects. For this reason, the method of multi-scale fusion is adopted in YOLOv3, which reduces the loss of information by concatenating the feature maps of different layers. And multi-scale detection is used to improve the accuracy of small objects. When detecting, the output feature map will be divided into s × s grids, and each grid predicts B bounding boxes and C category probabilities, as shown in Fig. 2. The output information of each bounding box includes t_x, t_y, t_w and t_h, which denotes the location information of the bounding box, the probability of the target object appearing in the prediction object box (prediction box), the confidence c of the accuracy of the prediction box and its category respectively. Finally, non maximum suppression (NMS) is used to select the bounding box with the highest probability, that is, the most representative of the object's character, to locate the object.

2) Prior box selection

In order to reduce the training difficulty, each grid in the output feature map is configured with prior boxes of the same size, and the prediction box will be adjusted based on the prior box. Original YOLOv3 uses K-means clustering algorithm to get the most suitable prior box size. K-means algorithm is simple and practical because the initial K clustering points are randomly selected. However, the classification results are easily affected by the initial points. In order to reduce the impact of the random selection of the initial point, this paper will use K-means++ algorithm to improve the selection of prior box size. It can analyze the height and width of the helmet, insulating boots and gloves in the work safety wearing images.

When the distance between the grid and the origin of the feature graph is (C_x, C_y), and the width and height of the prior box are p_w and p_h respectively, the position of the prediction box will be adjusted by (1).

$$\begin{aligned} b_{\mathrm{x}} &= \sigma(t_{\mathrm{x}}) + C_{\mathrm{x}}, \\ b_{\mathrm{y}} &= \sigma(t_{\mathrm{y}}) + C_{\mathrm{y}}, \\ b_{\mathrm{w}} &= P_{\mathrm{w}} e^{t_{\mathrm{w}}}, \\ b_{\mathrm{h}} &= P_{\mathrm{h}} e^{t_{\mathrm{h}}}, \end{aligned} \tag{1}$$

where b_{x} and b_{y} are the center coordinates of the prediction frame; b_{w} and b_{h} are the width and height of the prediction frame; and $\sigma(*)$ are the output of sigmoid function which limits the prediction offset value in the range of 0 to 1. The screening index of prediction box is that the intersection of prediction box and real boundary box is larger than IOU

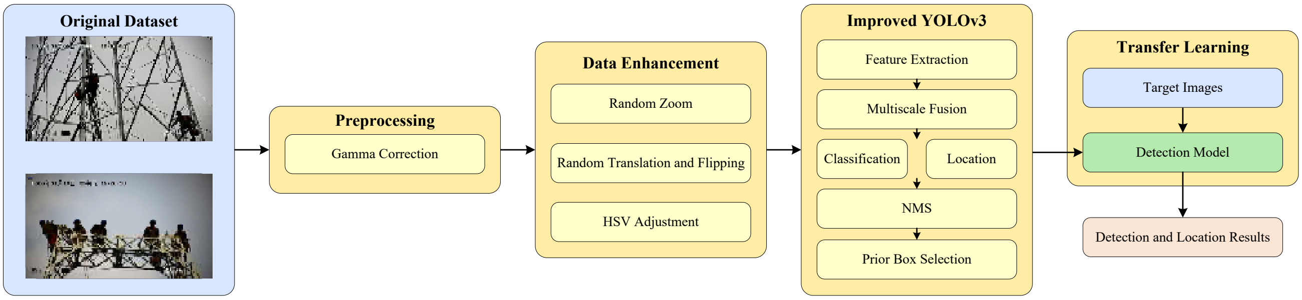


FIGURE 1. Proposed wear-enhanced YOLOv3 for PSPE wear detection.

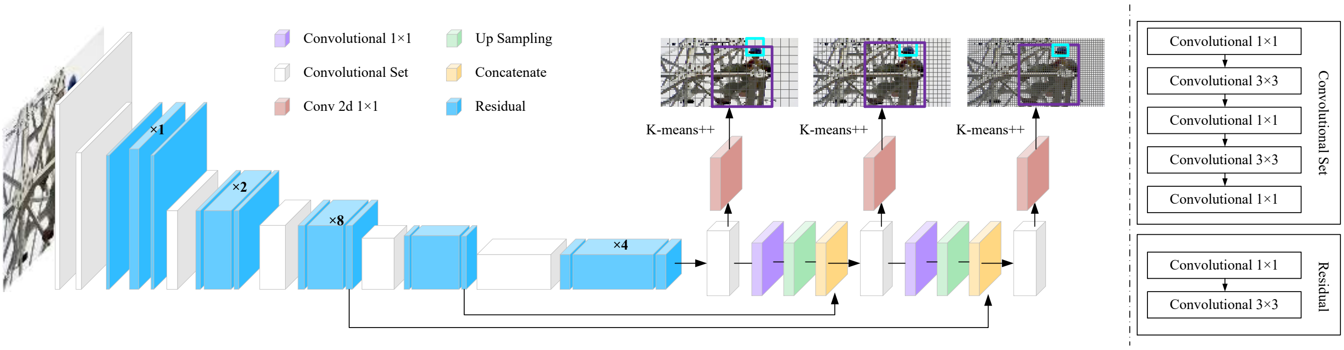


FIGURE 2. Identificaiton and location network based on YOLOv3

(intersection over union), which has nothing to do with the size of prediction box. Therefore, the distance measure uses the IOU distance, and the calculation formula is shown as follows.

$$d = 1 - IOU = 1 - \frac{A \cap B}{A \cup B} \tag{2}$$

where A and B are two bounding boxes, d is the IOU distance between A and B, and IOU is the intersection and union ratio between A and B. The larger the IOU is, the closer the bounding boxes of A and B are.

K-means++ considers that the further the point is from the current cluster center, the higher the probability of being selected as the cluster center, which effectively degrades the classification error caused by the selection of the initial cluster point and improves the detection speed. The specific calculation process is as follows.

1) The size of each PSPE image to be tested is taken as the analysis sample. Set $Y_1, Y_2, ..., Y_k$ as k cluster points.
2) Randomly select a sample as the first clustering point Y_1.
3) The IOU distance $d(x_i, Y_k)$ between each sample x_i and the nearest clustering point is calculated. And the probability that the point is selected as the next clustering point is derived according to (3).

$$P = d(x_i, Y_k)^2 \Big/ \sum d(x_i, Y_k)^2 \tag{3}$$

Finally, the next cluster center is selected according to the roulette method.

4) Repeat last step until the number of selected cluster points is K.
5) For each sample, the IOU distance from it to K clustering points is calculated. Then it is assigned to the cluster with the smallest IOU distance.
6) Modify the cluster points to the median of the IOU distance in the cluster, and repeat the last step until the cluster points are no longer changed. Then the K cluster points determined are the representative values of the size of the wearable equipment to be detected obtained by K-means++ algorithm.

3) Optimization Objective

The loss function of YOLOv3 can be divided into three parts: coordinate loss L_{loc}, confidence loss L_{conf} and category loss L_{cls}, as shown in (4) [18].

$$f_{\text{loss}} = \frac{1}{n}(L_{\text{loc}} + L_{\text{conf}} + L_{\text{cls}}), \tag{4}$$

where n is the total number of processed images. L_{loc}, L_{conf}, and L_{cls} are calculated as follows.

$$L_{\text{loc}} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathrm{I}_{ij}^{\text{obj}} (2 - wh) \times (L_{\text{xy}} + 0.5 L_{\text{wh}}) \tag{5}$$

$$L_{\text{conf}} = \sum_{i=0}^{S^2}\sum_{j=0}^{B} \mathrm{I}_{ij}^{\text{obj}}[-\hat{c}_i^j \log c_i^j + (1-\hat{c}_i^j)\log(1-c_i^j)] + \sum_{i=0}^{S^2}\sum_{j=0}^{B} \mathrm{I}_{ij}^{\text{noobj}}[-\hat{c}_i^j \log c_i^j + (1-\hat{c}_i^j)\log(1-c_i^j)] \tag{6}$$

$$L_{\text{cls}} = \sum_{i=0}^{S^2}\sum_{j=0}^{B} \mathrm{I}_{ij}^{\text{obj}}[-\hat{p}_i^j \log p_i^j + (1-\hat{p}_i^j)\log(1-p_i^j)] \tag{7}$$

where L_{xy} and L_{wh} are coordinate loss and width height loss, respectively. w is the width of the normalized prediction box, h is the height of the normalized prediction box. $\mathrm{I}_{ij}^{\text{obj}}$ is 1 when the jth prior box in the ith grid is responsible for predicting the object, otherwise it is 0. $\mathrm{I}_{ij}^{\text{noobj}}$ is 1 when the jth prior box in the ith grid is not responsible for predicting the object, otherwise it is 0. c_i^j is the confidence prediction value of the prediction box corresponding to the jth prior box in the ith grid, and $\hat{c}_i^j$ is the real confidence value of the prediction box corresponding to the jth prior box in the ith grid. When the jth prior box in the ith grid is responsible for predicting the object, the real confidence value of the corresponding prediction box is 1, otherwise it is 0. p_i^j is the probability prediction value of the prediction box corresponding to the jth prior box in the ith grid, and $\hat{p}_i^j$ is the probability true value of the prediction box corresponding to the j-th prior box in the i-th grid.

Among the IOUs of all prior frames and their real object frames (real frames) in the same grid, the prior frame will be responsible for predicting the object when the IOU between a prior frame and the real frame is the largest. As the prior frame is small, the offset degree has a great impact on the accuracy. Thus, $2-wh$ is used to balance the influence of the prediction frame size in the calculation of coordinate loss. When wh is small, i.e., the prediction frame is small, $2-wh$ is large, which improves the offset loss of a smaller prediction frame. When wh is large, i.e., the prediction frame is large, $2-wh$ is smalle, which reduces the offset loss of a larger prediction frame.

D. TRANSFER LEARNING

The effect of deep neural network depends on the number of data samples. Only on the basis of a large number of training data samples can a neural network with good performance be fitted. If the training data sample is small, the trained neural network is easy to overfit and the generalization ability is poor.

The dataset established in this paper is relatively small. If the network is directly trained by dataset of PSPE wear, not only the training speed is slow, but also the accuracy is poor. In order to accelerate the network training, enhance the network generalization ability and improve the accuracy, Pascal VOC dataset is used to pretrain the wear-enhanced YOLOv3. Pascal VOC dataset has 27450 images, covering 20 kinds of objects such as people, animals, furniture and vehicles. The pretraining model trained by Pascal VOC dataset has strong generalization ability. It can reflect the low-level features of shallow learning image, including image edge, color information and so on. The deep layer will learn that the high-level semantic features of images will have different performance with different images. Therefore, for the detection of PSPE wear in substation, this paper will improve the performance of the proposed method by transfer learning. It is divided into the two following stages.

1) Transplant the network parameters of each layer of the pre-training model to the WEYOLOv3 network based on PSPE wear. The parameters are not adjusted before freezing the selected layers. In order to accelerate the training speed of the WEYOLOv3 network, the parameters of the back layer are trained with a larger learning rate.
2) Unfreeze all the parameters of the WEYOLOv3 network and verify the loss function by monitoring. It is able to dynamically reduce the learning rate and fine-tune all the parameters of the network.

III. EXPERIMENTAL STUDY

A. DATASET

There is no public dataset of PSPE available, in this experiment the data samples are collected from the operation sites of several substations in Qingyuan City, Guangdong Province from 2019 to 2020. The dataset contains 500 images, including operators and three kinds of PSPE i.e., safety helmets, insulating boots and insulating gloves. Fig. 3 shows the typical samples of our dataset. It can be seen from Fig. 3 that some interference information inevitably appears in the substation operation monitoring image, such as the surrounding guardrail, transformer and so on. When the camera angle changes, the power equipment may partially block the operator, resulting in fuzzy boundary box labeling, affecting the learning of boundary box. Moreover, the relationship between PSPE and the staff is difficult to determine, that is, the staff may only carry the PSEP but not wear properly. At the same time, the image has the disadvantages of complex background, weak contrast and unclear image, which makes it difficult to detect the wearing of PSPE accurately. In this dataset, about 80% of the images are randomly selected as the training set, and the remaining 20% as the network test set.

In this paper, three indexes for evaluating object detection models are illustrated. The first index is Average Precision (AP), which is used to measure the detection results for every class. The second index is mean average precision (mAP), which is the mean of APs for all classes. The third index is frames per second (FPS), which indicates the number of images detected per second. It reflects the running speed of deep learning model. By these three indexes, the effect of following detection and location models can be evaluated.

B. COMPARISON BETWEEN ENHANCED AND ORIGINAL PRIOR BOXES

The original YOLOv3 uses K-means algorithm, while the wear-enhanced YOLOv3 uses K-means++ algorithm. In the

(a) (b) (c) (d) (e) (f) (g) (h)

FIGURE 3. Typical samples of substation operation monitoring image dataset.

same dataset, two experiments of enhanced and original YOLOv3 were carried out. The same experiment was repeated ten times to reduce the error. The experimental results are shown in Table 1. Although the recognition precision of operators and helmets decreased slightly, the recognition precision of insulating gloves and boots increased significantly. Overall, the average precision increased by 8.04%. Meanwhile, the time of image recognition and location is almost the same, but the recognition precision of the enhanced prior boxes is higher than that of the original prior boxes. The WEYOLOv3 algorithm can better select the size of the prior box, so as to identify the small size of insulating gloves and boots. The proposed model is more suitable for the detection and location of substation safety wear.

C. COMPARISON BETWEEN PREPROCESSED AND ORIGINAL IMAGES

After gamma correction, the operators and their wearing equipment in the image are more easily recognized. First, in substation operation, workers often operate in all kinds of weather. In addition, different light and angles also make it difficult to distinguish the safety helmet, insulating gloves and insulating boots of operators. Secondly, the background of substation is complex. Different from other single background images, substation operation images often contain lightning arresters, circuit breakers and other substation equipment, which is quite confusing. Based on this, the image consists of substation operators is significantly improved after gamma correction. Fig. 4 shows the preprocessing results based on gamma correction. As shown in Fig. 4, under the backlight condition, the gamma corrected images are obviously easier to be identified and detected the wearing of PEPE. It can be seen that gamma correction preprocessing method can highlight the details of operators, enhance the ability of detail visibility, and reduce the influence of complex background to a certain extent. The images obtained by the above methods are more obvious, and can be better identified and located. Through image preprocessing, the above two problems are effectively solved, and the application of YOLOv3 algorithm in the detection and positioning of substation operators' safety wearing is further improved.

In fact, the image preprocessing method will enhance the detection ability of the model by enhancing the contrast and feature details of the image. As shown in Fig. 5, the detection and positioning accuracy of operators, safety helmets, insulating gloves and insulating boots are all improved. The improvement of average precision fully shows the superiority of pretreatment, as presented in Table 2. Pretreatment makes the object detection algorithm more suitable for substation safety operation monitoring.

D. COMPARISON BETWEEN PROPOSED METHOD AND OTHER IMAGE DETECTION METHODS

In order to evaluate the effectiveness of the proposed method on the images of PSPE wear, the object recognition and location method of PSPE wear using WEYOLOv3 is compared with the Faster R-CNN, SSD, DSSD, and the RetinaNet algorithm. The parameters of comparative algorithms are similar to the proposed WEYOLOv3, and the original dataset is obtained under the pre-training weight.

It can be seen from Table 3 that the mAP of Faster R-CNN, SSD, and DSSD is much lower than that of the model pro-

TABLE 1. Results of detection models using enhanced and original prior boxes.

Model	Human	Safety Helmet	Insulating Gloves	Insulating Boots	mAP	FPS
Original Prior Boxes	83.12%	75.05%	59.68 %	59.18 %	69.26%	20.5
Enhanced Prior Boxes	76.13%	74.72%	72.17 %	86.18 %	77.30%	20.1

TABLE 2. Results of detection models coping with preprocessed and original images.

Model	Human	Safety Helmet	Insulating Gloves	Insulating Boots	mAP	FPS
No preprocessing	76.13%	74.72%	72.17 %	86.18 %	77.30%	20.1
Preprocessing	78.79%	78.12%	73.40 %	88.07 %	79.58%	19.5

(a) original image (b) preprocessed image

(c) original image (d) preprocessed image

FIGURE 4. PSPE images: (a)(c) original images; (b) (d) preprocessed images.

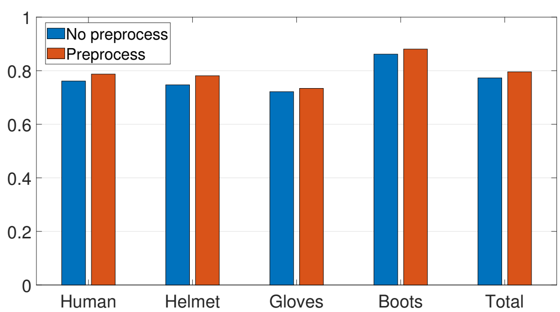


FIGURE 5. Average precision of no preprocess and preprocess.

posed in this paper. Although AP of human detection using SSD and DSSD is slightly higher than that of WEYOLOv3, WEYOLOv3 achieves a much higher detection precision for helmet, gloves, and boots. The detection precision of the RetinaNet is close to the proposed WEYOLOv3. However,

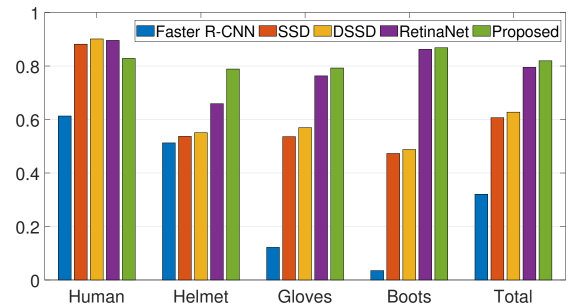


FIGURE 6. Average precision of Faster R-CNN, SSD, DSSD, RetinaNet and proposed WEYOLOv3.

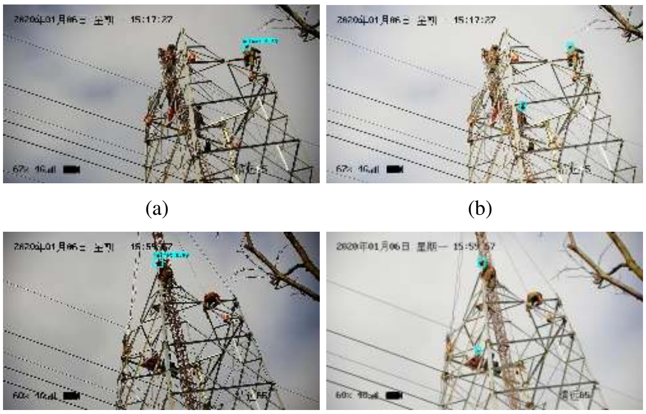
the detection speeds of the Faster R-CNN, SSD, DSSD, and RetinaNet are lower than the proposed WEYOLOv3. The FPS of the other comparative methods is even half of that of the WEYOLOv3, as shown in Table 3. The comparison results are shown in Fig. 6. Through regularization, prior box size selection, multi-scale detection and other optimization, the accuracy of substation operators' safe wear model based on the YOLOv3 algorithm is improved. Moreover, the network becomes more complex, deeper and more extensive. By analyzing the characteristics of substation operators' safety protective wear image, the image is preprocessed, and the network parameters are optimized accordingly. Then the transfer learning method is used to take the substation operators' protective wear image detection based on WEYOLOv3 in this paper, which can ensure the fast speed of detection and positioning, and greatly improve the detection precision and processing efficiency. In this way, the proposed model is more suitable than Faster R-CNN, SSD, DSSD, and RetinaNet to achieve the accurate detection and location of PSPE wear. It meets the requirement of daily real-time detection and provides effective support for further standardized operation.

E. DETECTION RESULTS UNDER EXTREME CONDITIONS

In order to verify the reliability of the proposed PSPE wear detection method for substation workers under extreme shooting conditions. This experiment evaluates the detection capabilities of the proposed model in three complex scenarios.

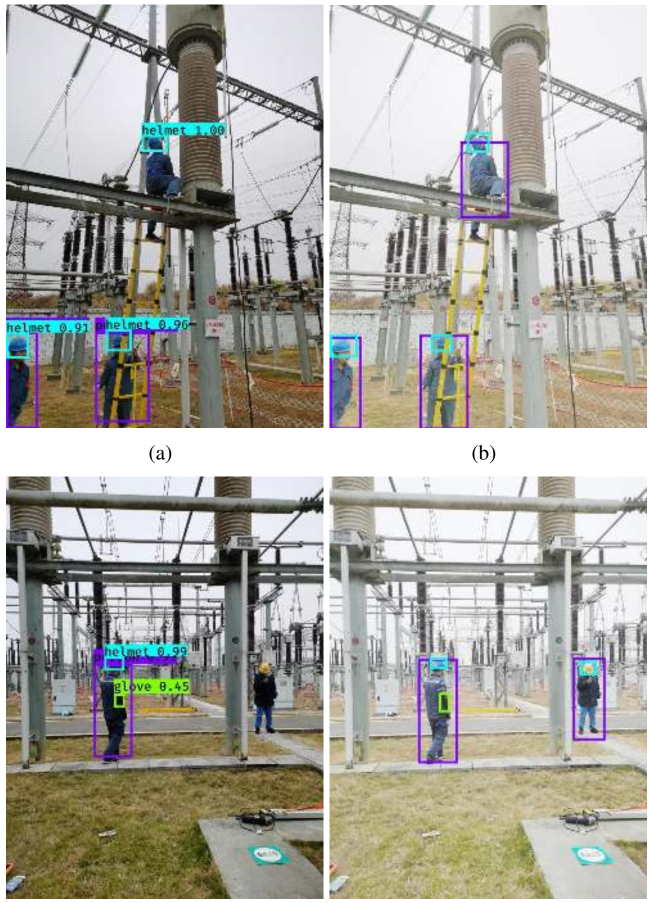
TABLE 3. Results of different detection models.

Model	Human	Safety Helmet	Insulating Gloves	Insulating Boots	mAP	FPS
Faster R-CNN [18]	61.31%	51.25 %	12.18 %	3.51 %	32.06 %	10.1
SSD [20]	88.16%	53.69%	53.60%	47.29%	60.68%	13.7
DSSD [21]	90.12%	55.07%	56.97%	48.78%	62.74%	12.3
RetinaNet [22]	89.58%	65.92%	76.31%	86.24%	79.51%	12.0
Proposed	82.84%	78.87%	79.25 %	86.83 %	81.95%	19.4



(a) (b) (c) (d)

FIGURE 7. First case: backlight shooting in (a) (c) original images; (b) (d) preprocessed images.



(a) (b) (c) (d)

FIGURE 8. Second case: extrem complex background in (a) (c) original images; (b) (d) preprocessed images.

The first case is the identification and positioning of the safe wear of the workers working at height under the backlight, as shown in Fig. 7. On the one hand, the backlight condition will lead to the lack of contrast in the wear details of operators. It's hard to tell what they're wear. On the other hand, the operator is smaller in the picture. Therefore, new requirements are put forward for prior frames. YOLOv3 algorithm is difficult to recognize in this case, only one helmet can be recognized. In contrast, the proposed method can detect two safety helmets at this angle. The accuracy of recognition and location is significantly improved. This shows the superiority of this method.

The second situation is that when the background is relatively complex, the recognition is easily interfered and missed, as shown in Fig. 8. There are a lot of columnar equipment in the substation, such as switch, arrester and its pillar. From the perspectives of the identification box, it is similar to the shape of the operator, so it is prone to interference. As shown in Fig. 8 (a) (c), there are cases of missing identification of substation operators. However, the proposed method based on WEYOLOv3 is improved by a prior box. K-means algorithm is replaced by Kmeans++ algorithm. Therefore, the selection of a prior box is more in line with the identification and positioning of substation operators, as shown in Fig. 8 (b) (d).

The third case is the problem that recognition is easy to be missed when the wearable image equipment is partially covered or overlapped, as shown in Fig. 9. In the field operation, the camera angle can not always be in front of the operator. When the operator's side is photographed, the safety helmet, insulating gloves and insulating boots are partially covered or overlapped. In Fig. 9 (a), only part of the helmet can be identified and located by YOLOv3. The leakage identification of insulating gloves and boots often appears in the original YOLOv3 algorithm, as shown in Fig. 9 (c) (e). However, the proposed method based on WEYOLOv3 has higher ability to identify the details of grid safety wearable equipment. Whether the helmet is partially covered, or the overlapped insulating gloves and boots can be stably identified and positioned, as shown in Fig. 9 (b) (d) (f).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3104731, IEEE Access

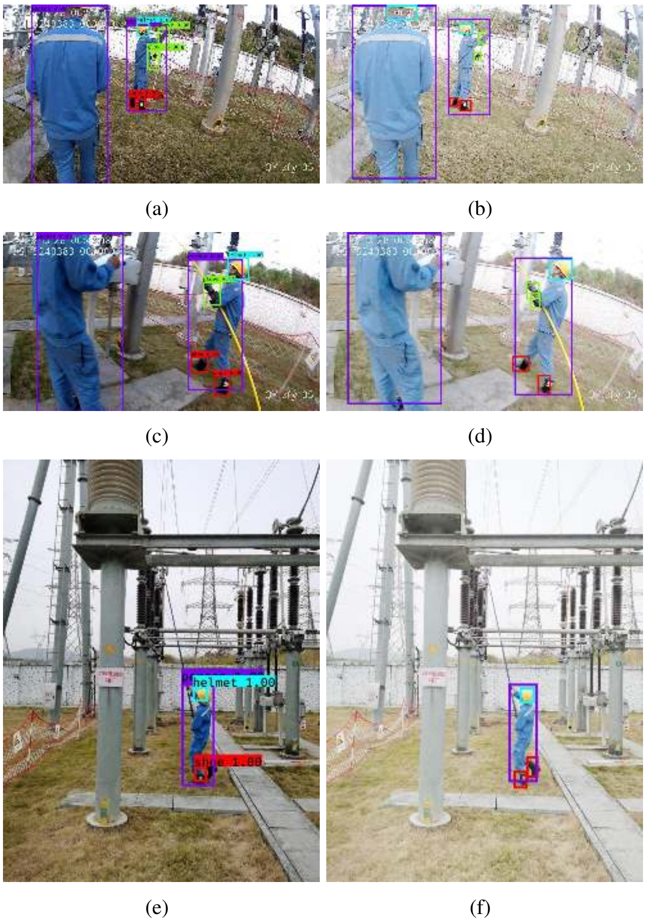


FIGURE 9. Third case: PSPE covered or overlapped in (a) (c) (e) original images; (b) (d) (f) preprocessed images.

IV. CONCLUSION

The detection and positioning of safety wear of substation operators is an important step to realize the safety operation of substation's automatic monitoring and a critical part of intelligent operation of substation. In order to solve the complex interference of background, camera angle and occlusion overlap in the detection and location of PSPE wear, this paper proposes a detection and location method of PSPE wear in substation based on wear-enhanced YOLOv3. Firstly, gamma correction is used to enhance the contrast of the original image, reduce the influence of complex background and improve the image quality. Then, the small sample of field safety wear is enhanced by data enhancement approaches. Furthermore, the scale parameters of YOLOv3 are simplified, and the prior frame size parameters are conducted by K-means++ algorithm. Finally, transfer learning is used to train the network, which realizes the accurate detection of safety wear images in substation operation. The experimental results show that the average precision of this method is 81.95%. Compared with other algorithms, the proposed method has very good detection and location effect. The safety monitoring of substation operation provides the basis for the follow-up safety specification and responsibility determination.

In the future work, the characteristics of PSPE images will be further investigated and the detection and location model will be further improved.

REFERENCES

[1] W. Cai, J. Le, C. Jin, and K. Liu, "Real-time image-identification-based anti-manmade misoperation system for substations," *IEEE Transactions on Power Delivery*, vol. 27, no. 4, pp. 1748–1754, Oct. 2012.

[2] B. A. S. Oliveira, A. P. D. F. Neto, R. M. A. Fernandino, R. F. Carvalho, A. L. Fernandes, and F. G. Guimarães, "Automated monitoring of construction sites of electric power substations using deep learning," *IEEE Access*, vol. 9, pp. 19 195–19 207, 2021.

[3] O. Gómez, C. Portilla, and M. A. Rios, "Reliability analysis of substation monitoring systems based on branch pmus," *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 962–969, Mar. 2015.

[4] Z. Baohua, L. Jing, and S. Yin, "Study of general monitoring system for unattended substation," *Electric Power Science and Engineering*, vol. 25, no. 9, pp. 76–P78, 2009.

[5] P. H. Meira de Andrade, J. M. M. Villanueva, and H. D. d. Macedo Braz, "An outliers processing module based on artificial intelligence for substations metering system," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3400–3409, Sept. 2020.

[6] J. Zhong, W. Li, R. Billinton, and J. Yu, "Incorporating a condition monitoring based aging failure model of a circuit breaker in substation reliability assessment," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3407–3415, Nov. 2015.

[7] M. R. Bastos and M. S. L. Machado, "Visual, real-time monitoring system for remote operation of electrical substations," in *2010 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America (T D-LA)*, Nov. 2010, pp. 417–421.

[8] M. J. Bharata Reddy, B. K. Chandra, and D. K. Mohanta, "A dost based approach for the condition monitoring of 11 kv distribution line insulators," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 18, no. 2, pp. 588–595, Apr. 2011.

[9] R. R. V. e. Silva, K. R. T. Aires, and R. d. M. S. Veras, "Helmet detection on motorcyclists using image descriptors and classifiers," in *2014 27th SIBGRAPI Conference on Graphics, Patterns and Images*, Aug. 2014, pp. 141–148.

[10] R. Waranusast, N. Bundon, V. Timtong, C. Tangnoi, and P. Pattanathaburt, "Machine vision techniques for motorcycle safety helmet detection," in *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*, Nov. 2013, pp. 35–40.

[11] J. Li, H. Liu, T. Wang, M. Jiang, S. Wang, K. Li, and X. Zhao, "Safety helmet wearing detection based on image processing and machine learning," in *2017 Ninth International Conference on Advanced Computational Intelligence (ICACI)*, Feb. 2017, pp. 201–205.

[12] B. Yogameena, K. Menaka, and S. S. Perumaal, "Deep learning-based helmet wear analysis of a motorcycle rider for intelligent surveillance system," *IET Intelligent Transport Systems*, vol. 13, no. 7, pp. 1190–1198, 2019.

[13] 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. 94 065–94 075, 2020.

[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[15] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 142–158, Jan. 2016.

[16] W. Wang, H. Yin, C. Chen, A. Till, W. Yao, X. Deng, and Y. Liu, "Frequency disturbance event detection based on synchrophasors and deep learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3593–3605, 2020.

[17] H. Lin, J. D. Deng, D. Albers, and F. W. Siebert, "Helmet use detection of tracked motorcycles using cnn-based multi-task learning," *IEEE Access*, vol. 8, pp. 162 073–162 084, 2020.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, Sept. 2015.

[19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[20] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*. Springer, 2016, pp. 21–37.

[21] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, "Dssd: Deconvolutional single shot detector," *arXiv preprint arXiv:1701.06659*, 2017.

[22] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.

[23] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 779–788.

[24] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 6517–6525.

[25] X. Gong, Q. Yao, M. Wang, and Y. Lin, "A deep learning approach for oriented electrical equipment detection in thermal images," *IEEE Access*, vol. 6, pp. 41 590–41 597, 2018.

[26] S. Chen, W. Tang, T. Ji, H. Zhu, Y. Ouyang, and W. Wang, "Detection of safety helmet wearing based on improved faster r-cnn," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–7.

[27] N. Li, X. Lv, S. Xu, Y. Wang, Y. Wang, and Y. Gu, "Incorporate online hard example mining and multi-part combination into automatic safety helmet wearing detection," *IEEE Access*, 2020.

[28] J. Shen, X. Xiong, Y. Li, W. He, P. Li, and X. Zheng, "Detecting safety helmet wearing on construction sites with bounding-box regression and deep transfer learning," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 2, pp. 180–196, 2021.

[29] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv*, 2018.

[30] L. Huang, Q. Fu, M. He, D. Jiang, and Z. Hao, "Detection algorithm of safety helmet wearing based on deep learning," *Concurrency and Computation: Practice and Experience*, p. e6234, 2021.

[31] H. Shi, X. Chen, and Y. Yang, "Safety helmet wearing detection method of improved yolov3," *Computer Engineering and Applications*, vol. 55, no. 11, pp. 213–220, 2019.

[32] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Automation in Construction*, vol. 112, p. 103085, 2020.

BAINING ZHAO received the B.Sc. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2019. He is currently pursuing the M.Sc. degree with the same university. His research interests include deep learning, fault diagnosis, and wind energy.

HAIJUAN LAN received the B.S. degree in Electrical Engineering from the South China University of Technology , Guangzhou, China, in 2016 and the M.S. degree in Electrical Engineering from the University of Queensland,Brisbane, Australia, in 2018. She have been working in Schneider Electric (Xiamen) Switchgear Company since 2018. Her research interests include HVDC and renewable energy.

ZHEWEN NIU received his B.Sc. degree in Electrical Engineering from Taiyuan University of Technology, Taiyuan, China, in 2016. He is currently pursuing the Ph.D. degree in Electrical Engineering at South China University of Technology, Guangzhou, China. Now, he is also with the Department of Electrical and Computer Engineering, University of Alberta, Alberta, Canada, as a visiting scholar under the support of China Scholarship Council (CSC). His research interests include deep learning and data mining in power systems.

HUILING ZHU received the B.E. degree in electrical engineering and automation from South China University of Technology in 2018, where she is currently pursuing the master's degree in electrical engineering. Her main reserach interests include image processing and the application of artificial intelligence technology in power system.

TONG QIAN (M'21) obtained a B.E. and Ph.D. degrees in Electrical Engineering from South China University of Technology, Guangzhou, China, in 2015 and 2020, respectively. He is currently working in the same area in SCUT as a postdoctoral researcher. His research interests include distributed cooperation of networked multi-agent system and the flexible operation of interdependent infrastructures in smart cities.

WENHU TANG (M'05-SM'13) received the B.Sc. and M.Sc. degrees in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 1996 and 2000, respectively, and the Ph.D. degree in electrical engineering from The University of Liverpool, Liverpool, U.K., in 2004. He was a Postdoctoral Research Associate and subsequently a Lecturer at The University of Liverpool between 2004 and 2013. He is currently a Distinguished Professor and the Dean of the School of Electric Power Engineering, South China University of Technology, Guangzhou, China. He also is a Fellow of IET. He has authored and coauthored 150 research papers, including 60 international journal papers and 1 Springer monograph. His research interests include power systems risk assessment, renewable energy integration in power grids, condition monitoring and fault diagnosis for power apparatus, multiple-criteria evaluation and intelligent decision support systems.