IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 6, 2021, accepted July 23, 2021, date of publication August 13, 2021, date of current version September 17, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3104731

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], (Member, IEEE), 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 Company Ltd., Xiamen 361000, China

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

This work was supported by the National Natural Science Foundation of China under Grant 51977082.

ABSTRACT Wearing personal safety protective equipment (PSPE) plays a key role in reducing electrical injuries to electrical workers. However, substation employees often ignore this regulation due to lack of safety awareness and discomfortable feeling of wearing PSPE. Therefore, it is necessary to develop a detection algorithm for PSPE and workers to build real-time video surveillance systems in power substations. In this paper, a wear-enhanced YOLOv3 method for real-time detection of PSPE and substation workers is proposed. The gamma correction is first applied as the preprocessing method to highlight the details of the operators and data augmentation is performed. Next, K-means++ algorithm replaces K-means in wear-enhanced YOLOv3 method to derive the most suitable prior bounding box size and lift the detection speed. Then, the proposed method can be quickly and effectively trained based on transfer learning. Finally, extensive experiments are carried out on a dataset of images about usage of safety helmets, insulating gloves and boots. Using the proposed method, the mean average precision is improved by over 2% and the frames per second is the highest compared with other typical object detection methods, which illustrates the effectiveness of the wear-enhanced YOLOv3 method for PSPE and workers 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 electric systems [1]. The continuous expansion of power systems leads to the increasing number of large-scale substations, whose safe and stable operation is important to power transmission [2]. Automatic substation maintenance attracts a great amount of research interest in the development of intelligent power systems [3]. Monitoring appropriate personal safety protective equipment (PSPE) on the job site becomes one of the critical problems to reduce workplace accidents and worker injuries [4]. Due to lack of sufficient attention and lack of safety awareness, substation workers often violate the operation regulations in actual fields, leading to various accidents [5]. The most frequent violation events of PSPE requirement include not wearing safety helmets, not wearing insulated gloves and boots, crossing safety barriers, etc. Thus, it is necessary to enforce and monitor the appropriate usage of PSPE in real time through a video surveillance system [6]. The use of object detection methods for recognizing the PSPE and workers is valuable to prevention of fatalities.

Current methods of automatic detection of PSPE can be basically categorized into two types: traditional image processing-based and deep learning-based. Traditional methods utilize image processing technology to extract skin color, head, and face information of workers [7]–[9]. For instance, Waranusast *et al.* [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

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

for helmets often fail to consider the impact of the complex substation operating environment [12]. The detection results are susceptible to environmental light, small medium objects, etc., which brings about frequent false alarms.

In recent years, deep learning techniques have been widely used in object detection owing to the ability to learn effective automatic feature [13]. Deep neural networks that consists of multiple layers are fed a certain amount of input and output data directly. Through repeated training, the networks 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 PSPE identificatioin of industry applications. 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. Besides, 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 was able to increase the feature map scale and optimize 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. In conclusion, YOLOv3 is a promising tool for detecting and locating PSPE appliance in power substation.

Most above-mentioned methods in power systems were designed for the detection of safety helmets, but few papers focused on multi-class 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 in the specific case of power substations. The real-time PSPE detection tasks in substations have three specific characteristics. Firstly, the backgrounds of the substation videos are usually complex and the cameras are distributed far away from the substation employees, resulting in a low resolution and intensity contrast of collected video images. Secondly, the small sizes of the PSPE make it less distinguishable from the backgrounds. Last but not least, real-time monitoring requires a high processing speed for the detection model. Although the existing methods indicated a promising performance of PSPE detection methods based on deep learning, there is still a lack of investigation of deep learning-based techniques for real-time PSPE detection in the field of online surveillance for substations. Therefore, such simple applications of the deep learning method cannot meet the requirements and a fast and accurate detection method of PSPE for substation workers is required.

To address the above-mentioned problems, this paper proposes a wear-enhanced YOLOv3 (WEYOLOv3) method for real-time detection of PSPE and workers in substations. The gamma correction is applied as the preprocessing method to highlight the details of the workers and improve the detection accuracy. Moreover, K-means++ algorithm is introduced to determine the most suitable bounding box priors to improve the detection precision and speed. Based on the proposed method, it is able to monitor whether the substation workers are wearing safety helmets, insulating gloves and boots correctly. Finally, a number of experiments are carried out on a dataset of real substation monitoring images to illustrate the effectiveness of the proposed method for real-time PSPE detection. 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 bounding box priors are acquired by K-means++ according to the PSPE characteristics.
3) Different methods are compared with the proposed method in detection performance. Serious 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 approach for PSPE and workers detection from visual data. The results of experiments and the discussion are presented in Section III. Finally, Section IV summarizes the conclusion of this paper and gives the future research plan.

II. DETECTION OF PSPE AND WORKERS BASED ON WEAR-ENHANCED YOLOv3

To address the complex backgrounds and changeable angles of substation monitoring images, this paper proposes a new comprehensive detection and location model for PSPE and workers in substations. Different from the existing models merely focusing on safety helmet identification, the proposed method can comprehensively detect the substation workers, safety helmets, insulating boots and gloves. The proposed method consists of four stages: preprocessing, data augmentation, WEYOLOv3 model training, and transfer learning, as shown in Fig. 1.

A. PREPROCESSING

The monitoring images or videos are preprocessed through gamma correction to reduce irrelevant information and accelerate the training process of WEYOLOv3.

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 and redact the nonlinear tone of the image. It is realized by detecting the dark and light part of the image signals and changing their proportion. Then images with a better contrast effect can be derived. The specific steps of the gamma correction are as follows.

1) Normalization: every pixel value is converted to a range between 0 and 1. In other words, it is calculated by: $I_N = (I + 0.5)/256$, where I is the original pixel value and I_N is the normalized pixel value.
2) Pre-compensation: the corresponding value I'_N is obtained by: $I'_N = I_N^{\frac{1}{\gamma}}$, where γ is the parameter of the gamma correction.
3) Denormalization: the normalized value I'_N after pre-compensation is converted to the original range between 0 and 255.

B. DATA AUGMENTATION

Because the data of PSPE and workers in substations are usually limited, the WEYOLOv3 utilizes the three data augmentation techniques. Data augmentation can enhance the generalization ability of the WEYOLOv3 and avoid the problem of overfitting in the learning process. The first approach is random erasing. It works by randomly selecting a specific size of patch of an image and masking it with either 0 s, 255 s, mean pixel values, or random values. Next, random translation and rotation are used to increase the number of images. Shifting images left, right, up, or down can be a very useful transformation and original images are rotated left or right on an axis between 1 and 359 degrees. Third, monitoring images are converted to HSV space and their exposure, saturation and brightness are randomly adjusted to generate auxiliary images.

C. IMPROVED YOLOv3

1) DETECTION AND LOCATION NETWORK BASED ON YOLOv3

YOLOv3 is a single-stage object detection algorithm. After improved, it can predict bounding boxes for PSPE and workers in substations. The basic structure of the network is shown in Fig. 1.

To extract the features of the input images, YOLOv3 utilizes Darknet-53, which is able to solve the problems of gradient disappearance or explosion [19]. YOLOv3 predicts boxes at three different scales, which reduces the loss of information when the layer corresponding to the feature map is deep. Multiscale feature maps help to improve the PSPE location of diverse size. After feature extraction, the output feature map will be divided into s × s grids, and each grid cell predicts bounding boxes and probabilities of classes, as shown in Fig. 2. Each bounding box includes 4 coordinates t_x, t_y, t_w and t_h, coupled with an objectness prediction representing its confidence score. Finally, non-maximum suppression (NMS) is used to select the similar bounding box with the highest probability and the PSPE classification and location in videos or images are derived.

2) DETERMINATION OF BOUNDING BOX PRIORS

In order to reduce the training difficulty, bounding box priors are used to adjust the predictive bounding boxes in the PSPE location. Original YOLOv3 used K-means clustering to acquire the most suitable bounding box priors because K-means clustering is practical. However, the initial K clustering points are randomly selected, which highly affects the final clustering results. Thus, the WEYOLOv3 applies K-means++ clustering to improve the determination of prior box size. It can effectively analyze the height and width of the helmets, insulating boots and gloves in the monitoring images.

If the cell is offset from the top left corner of the image by (C_x, C_y), and the width and height of the bounding box prior are p_w and p_h respectively, then the position of the predictive 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 bounding box. The width and height of bounding box are denoted as b_{w} and b_{h}, respectively. $\sigma(*)$ represents the output of sigmoid function, which limits the prediction offset value in the range of 0 to 1. The distance measured in clustering utilizes the IOU (intersection-over-union) distance, which can be calculated 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. The larger the IOU is, the closer the bounding boxes of A and B are.

By comparison, WEYOLOv3 utilizes K-means++ clustering, where the probability that a point is likely regarded as a clustering point is high when the point is far from the current clustering point. It effectively reduces the impact caused by the random selection of the initial clustering points and improves the detection speed. The specific calculation process is as follows.

1) The labeled information of bounding boxes in the training set are considered. Let $Y_1, Y_2, \ldots, Y_k$ represent k clustering points.

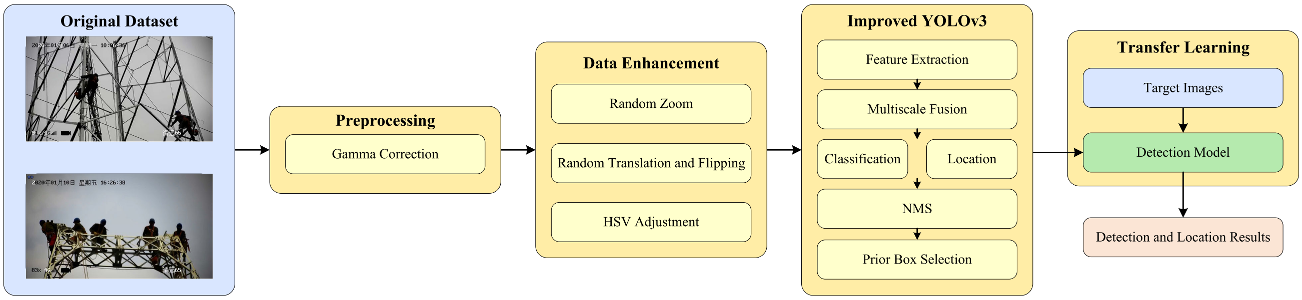


FIGURE 1. **Proposed wear-enhanced YOLOv3 for PSPE and workers detection.**

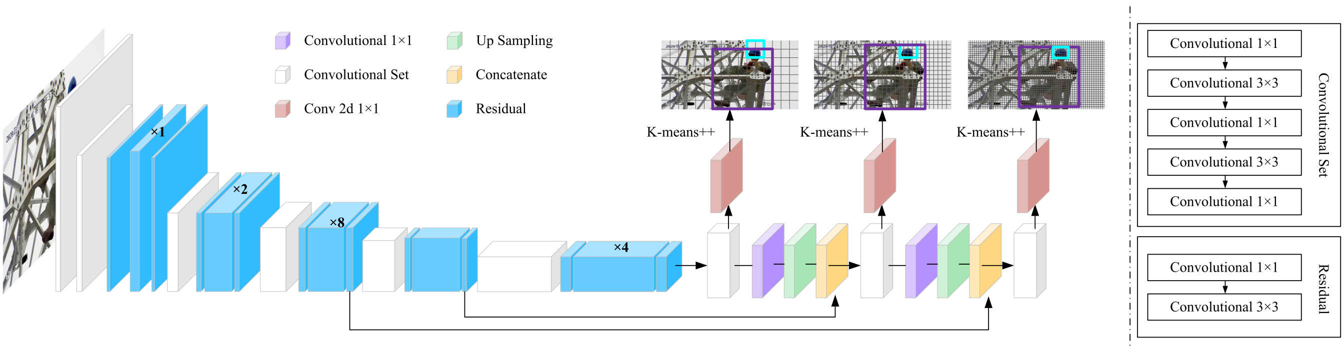


FIGURE 2. **Identificaiton and location network based on YOLOv3.**

2) A bounding box is randomly determined as the first clustering point Y_1.
3) The IOU distance $d(x_i, Y_k)$ between each sample x_i and its nearest clustering point is calculated. The probability that the point is selected as the next clustering point is derived according to (3).

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

The next clustering point is determined according to the roulette method.
4) The last step 3 is repeated until the number of clustering points is K.
5) For each bounding box, the IOU distance from it to the determined K clustering points is calculated. Then it is assigned to the cluster with the smallest IOU distance.
6) The cluster points are modified according to the median number of the IOU distances in the cluster. Then 5 is repeated until the clustering points are no longer changed. Finally, the K clustering points are the representative values of the size of the PSPE and workers.

3) OPTIMIZATION OBJECTIVE

The loss function of WEYOLOv3 can be divided into three parts: location loss L_{loc}, confidence loss L_{conf} and classification loss L_{cls}, as shown in (4).

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

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

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

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

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

where L_{xy} and L_{wh} are relative to coordinate, width, and height of bounding boxes, respectively. w is the width and h is the height of the normalized predictive bounding boxes. $\mathrm{I}_{ij}^{\text{obj}}$ is 1 when the jth prior box in the ith grid cell is responsible for predicting the object, otherwise it is 0. $\mathrm{I}_{ij}^{\text{noobj}}$ is 1 when the jth bounding box prior in the ith grid cell is not responsible for predicting the object, otherwise it is 0. c_i^j represents the confidence score of the predictions corresponding to the jth bounding box prior in the ith grid cell, and $\hat{c}_i^j$ is the real corresponding value. When the jth bounding box prior in the ith grid cell is responsible for predicting the PSPE or workers, the real confidence of the corresponding bounding


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

FIGURE 3. Typical samples in substation monitoring image dataset.

TABLE 1. Detection results using improved and original bounding box priors.

Model	Worker	Safety Helmet	Insulating Gloves	Insulating Boots	mAP	FPS
Original YOLOv3	83.12%	75.05%	59.68 %	59.18 %	69.26%	20.5
WEYOLOv3	76.13%	74.72%	72.17 %	86.18 %	77.30%	20.1

box is 1, otherwise it is 0. p_i^j is the classification prediction value of the prediction corresponding to the jth bounding box prior in the ith grid cell, while $\hat{p}_i^j$ is the corresponding labeled value.

D. TRANSFER LEARNING

The Darknet-53 requires a great number of training data for training. However, the data of monitoring images in substations is inadequate. Direct training WEYOLOv3 by original dataset leads to huge computation consumption and poor performance. To overcome these challenges, WEYOLOv3 is pre-train on Pascal VOC dataset. Pascal VOC dataset includes 27450 images, covering 20 kinds of objects such as people, animals, furniture and vehicles. Transfer learning enable WEYOLOv3 to possess stronger generalization ability and learn basic features and mid-level features (e.g., image edge, color information, object shape) that is useful in the detection of PSPE and workers. The process can be divided into two following stages. First, the network parameters of each layer in the pre-trained model are transferred to the WEYOLOv3 network. Then, the parameters in former layers are fixed and that in last several layers are trained with a relatively large learning rate to accelerate the training speed of WEYOLOv3. Finally, the learning rate is dynamically reduced and all the parameters are unfreezed and fine-tuned.

III. EXPERIMENTAL STUDY

A. DATASET

There is no public dataset of PSPE available and the data in this experimental study are collected from several substations of Qingyuan Power Supply Bureau in Guangdong Province, China from 2019 to 2020. The dataset contains 522 images, including workers and three kinds of PSPE, i.e., safety helmets, insulating boots and insulating gloves. Fig. 3 shows typical samples of our dataset. It can be seen from Fig. 3 that much noise inevitably appears in the substation monitoring images, such as the surrounding guardrail, transformer and so on. With the camera angle changes, the power equipment may partially block the workers, resulting in fuzzy boundary box labeling and affecting the learning of boundary box. At the same time, the images have characteristics of complex background, weak contrast and unclear image, which makes it difficult to detect the PSPE and workers accurately. In the experimental study, about 80% of the images are randomly selected as the training set and the remaining 20% are used in the test set.

In this paper, three indexes for evaluating detection performance 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

TABLE 2. Detection results about preprocessed and original monitoring images in substations.

Model	Worker	Safety Helmet	Insulating Gloves	Insulating Boots	mAP	FPS
No preprocessed	76.13%	74.72%	72.17 %	86.18 %	77.30%	20.1
Preprocessed	78.79%	78.12%	73.40 %	88.07 %	79.58%	19.5

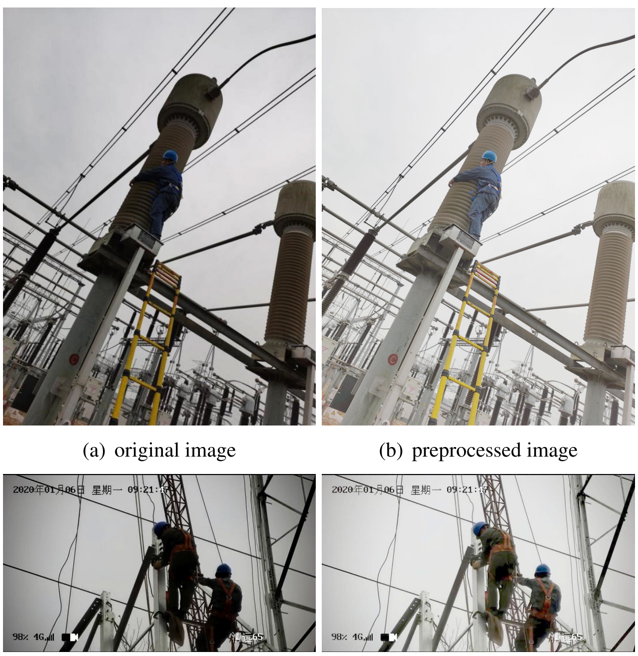
TABLE 3. Comparative results using different detection models.

Model	Worker	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 WEYOLOv3	82.84%	78.87%	79.25 %	86.83 %	81.95%	19.4

frames per second (FPS), indicating the number of images detected per second. It reflects the running speed of detection models. Through these three indexes, the effect of following detection and location models can be analyzed.

B. COMPARISON BETWEEN IMPROVED AND ORIGINAL BOUNDING BOX PRIORS

The original YOLOv3 utilizes K-means clustering, while the WEYOLOv3 uses K-means++ clustering. Each experiment was repeated ten times and the experimental results are shown in Table 1. Although the AP of workers and safety helmets falls slightly, that of insulating gloves and boots rises significantly. Overall, the mAP sees a significantly rise by 8.04%. Meanwhile, the FPS of original YOLOv3 is almost identical to that of WEYOLOv3, but the recognition and location of the improved bounding box priors are accurate than before. WEYOLOv3 can properly determine the bounding box priors and detect the small object i.e., insulating gloves and boots. The proposed model is more suitable for the classification and location of PSPE and workers in substations.



(a) original image (b) preprocessed image

(c) original image (d) preprocessed image

FIGURE 4. PSPE and worker in (a)(c) original images and (b) (d) preprocessed images.

C. COMPARISON BETWEEN PREPROCESSED AND ORIGINAL IMAGES

Several characteristics of monitoring images and videos increase difficulties of PSPE and workers detection. On the one hand, different weather conditions affect the monitoring videos and images greatly. In addition, different angles also make it difficult to detect the safety helmet, insulating gloves and insulating boots. On the other hand, the background of substation significantly influence the location of PSPE. Different from images with simple backgrounds, monitoring images in substations often contain lots of kinds of equipment like lightning arresters, circuit breakers and so on. After gamma correction, the workers and PSPE in the monitoring images are more easily recognized. Fig. 4 shows the preprocessed results based on gamma correction. For example, objects in the images after gamma correction are obviously easier to be classified and located under the backlight condition. Gamma correction can highlight the details of PSPE and reduce the influence of complex background. Through image preprocess, the above-mentioned two problems are effectively solved and WEYOLOv3 can achieve the detection of PSPE and workers.

In fact, the image preprocess approach can improve the detection ability of the model by enhancing the contrast and feature details of the images. As shown in Fig. 5, the AP of operators, safety helmets, insulating gloves and insulating boots all increases. The improvement of detection performance fully shows the superiority of preprocess, as presented in Table 2. Gamma correction enables WEYOLOv3 more suitable for classification and location of PSPE and workers.

D. COMPARISON AMONG PROPOSED METHOD AND OTHER IMAGE DETECTION METHODS

In order to evaluate the effectiveness of the proposed method coping with substation monitoring images, the detection results of PSPE and workers using WEYOLOv3 is compared with the Faster R-CNN, SSD, DSSD, and RetinaNet.

It can be seen from Table 3 that the mAP of Faster R-CNN, SSD, and DSSD is lower than that of the model proposed in this paper. Although AP of worker detection using SSD and DSSD is slightly higher than that of WEYOLOv3, the proposed WEYOLOv3 achieves a much higher detection precision for helmet, insulating gloves, and insulating boots. The detection precision of the RetinaNet is slightly lower than that of the proposed WEYOLOv3. However, the FPS of the Faster R-CNN, SSD, DSSD, and RetinaNet is significantly lower than that of WEYOLOv3. The detection speed of the proposed method is at least 50% faster than that of other comparative methods, as shown in Table 3 and Fig. 6. Through preprocess, improved bounding box priors, and effective transfer learning, the detection precision of PSPE and workers based on YOLOv3 is improved. The image is preprocessed by analyzing the characteristics of monitoring images and videos. K-means++ replaces K-means clustering and enhance the detection ability of PSPE, as shown in Table 3. Then the transfer learning is used to accelerate the training process of WEYOLOv3 and greatly improve the detection precision. In this way, the proposed model is more powerful than Faster R-CNN, SSD, DSSD, and RetinaNet to achieve precise classification and location of PSPE and workers. Besides, the detection speed of WEYOLOv3 is fast enough to deal with daily real-time detection task and the proposed method overcomes other four image detection networks.

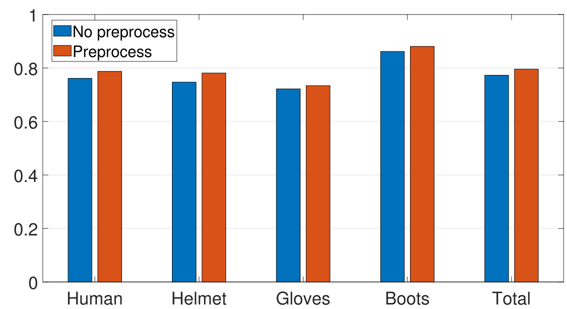


FIGURE 5. Average precision of original and preprocessed monitoring images.

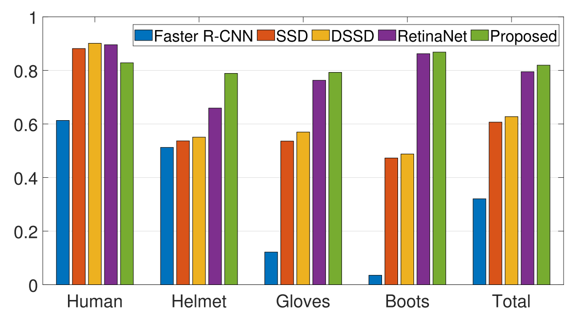


FIGURE 6. Average precision of Faster R-CNN, SSD, DSSD, RetinaNet and proposed WEYOLOv3.

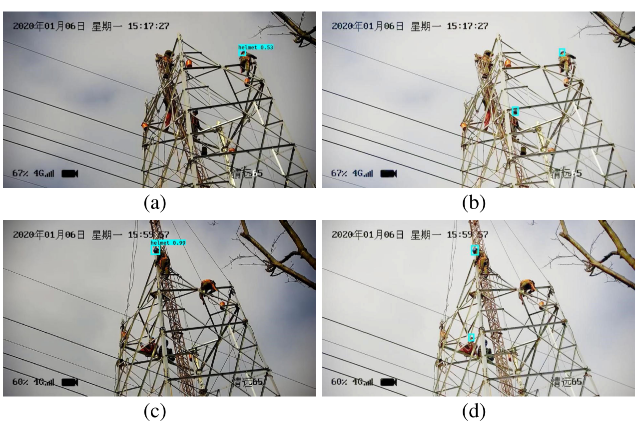


FIGURE 7. First case: (a) (c) original images and (b) (d) preprocessed images under backlight in substations.

E. DETECTION RESULTS UNDER EXTREME CONDITIONS

In order to verify the reliability of the proposed detection method for PSPE and workers in substations, experiments are conducted to evaluate the detection capabilities of the proposed WEYOLOv3 in three scenarios under extreme conditions.

The first case is the identification and location of the PSPE and the employees working on high poles 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

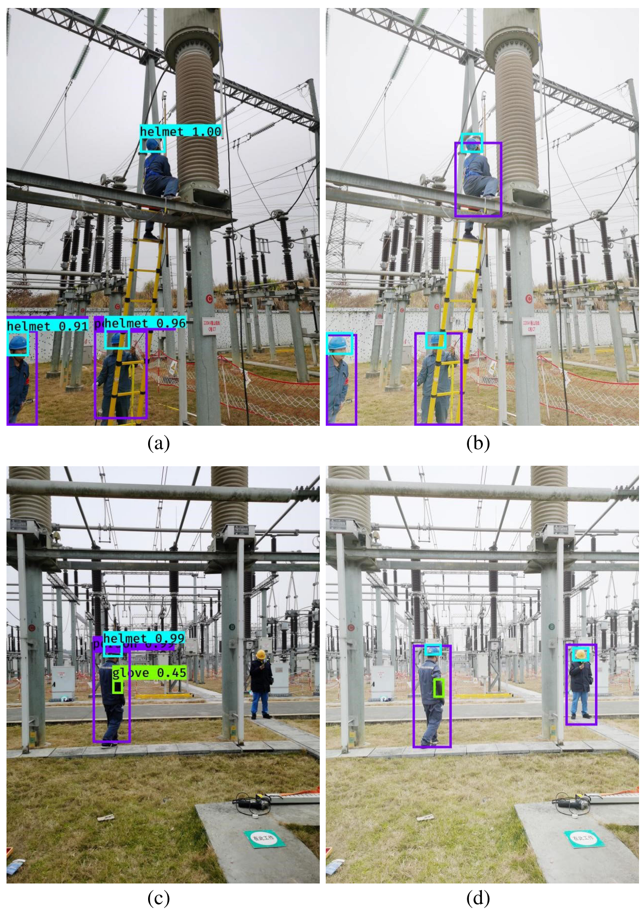


FIGURE 8. Second case: detection against complex backgrounds using (a) (c) original YOLOv3; (b) (d) proposed WEYOLOv3.

and the PSPE workers are wearing is hard to be distinguished. On the other hand, the workers may sometimes be smaller in the picture. Therefore, the detection in monitoring videos and images requires more effective bounding box priors. It is clear that YOLOv3 is difficult to perform detection in this case and only one helmet can be recognized. In contrast, the proposed method can detect two safety helmets at the same angle. The detection precision is significantly improved, which shows the superiority of WEYOLOv3.

The second situation is that the detection is easily interfered and missed when the backgrounds are relatively complex, as shown in Fig. 8. There are a lot of columnar equipment in the substations, such as switches, arresters and its pillars, whose shape and color are diverse. As shown in Fig. 8 (a) (c), there are cases of missing identification of substation employees. However, the proposed WEYOLOv3 improves the determination of bounding box priors. Therefore, the bounding box priors are more in accord with the identification and location of PSPE and workers against the complex backgrounds, as shown in Fig. 8 (b) (d).

The third case is that PSPE and workers partially appearing in images or overlapped by other objects are difficult to be identified and located, as shown in Fig. 9. In substations, the camera angle can hardly be kept in front of the workers. When workers keep moving, the safety helmet, insulating gloves and insulating boots usually need be detected according to incomplete information. In Fig. 9 (a), only part of the safety helmet appear and the helmet is ignored by YOLOv3. The important omissions of detecting insulating gloves and boots often appears using the original YOLOv3, as shown in Fig. 9 (c) (e). Comparatively, the proposed WEYOLOv3 is able to identify the details of PSPE and workers in substations. Thus, the safety helmet, insulating gloves and boots can be successfully detected even if only part of information is given, as shown in Fig. 9 (b) (d) (f).

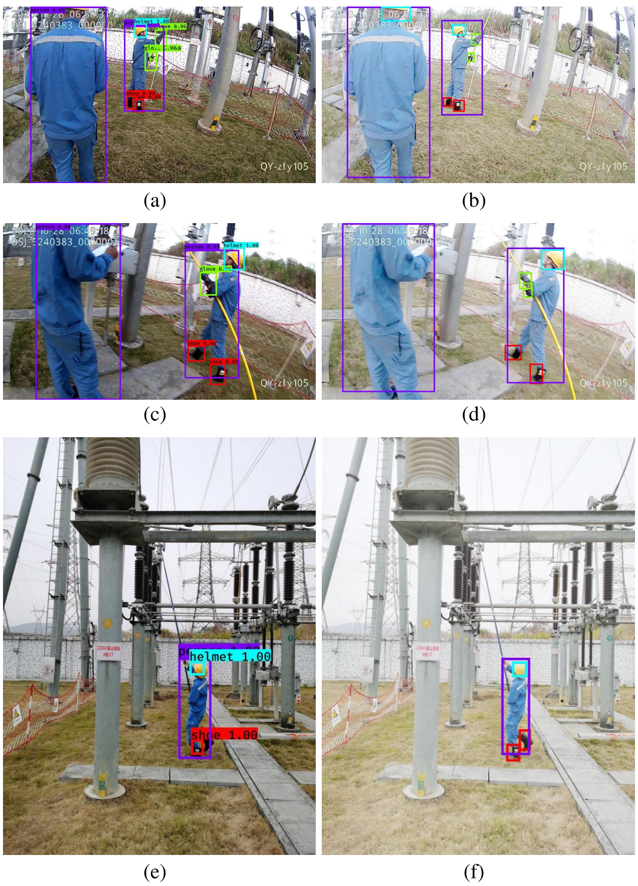


FIGURE 9. Third case: detection on PSPE and workers partially appearing in images or overlapped by other objects using (a) (c) (e) original YOLOv3; (b) (d) (f) proposed WEYOLOv3.

IV. CONCLUSION

The detection of PSPE and workers in substations is an important step to realize the automatic monitoring. In order to solve the problem of backlight, interference of complex backgrounds, and diverse shape and size of PSPE and workers, this paper proposes a novel detection model namely wear-enhanced YOLOv3 for PSPE and workers detection in substations. Firstly, gamma correction is used to enhance the contrast of monitoring images, reducing the impact of backlight and complex backgrounds. Then, more artificial images of PSPE and workers is derived by data augmentation approaches. Furthermore, the bounding box priors are determined by K-means++ clustering rather than K-means in original YOLOv3. Finally, transfer learning is used to train the network, which realizes the accurate detection of safety wear images in substation operation.

Based on the experimental results, the following conclusions can be drawn.

1) The usage of gamma correction improves the mAP by 2.28%. K-means++ clustering in WEYOLOv3 is able to produce more effective bounding box priors than K-means clustering in original YOLOv3, with 8.04% mAP improvement.
2) the WEYOLOv3 not only outperforms the widely-used image detection techniques in detection precision but also possesses a at least 50% faster FPS.
3) Extreme conditions including backlight, complex backgrounds, and fragmentary object information are considered, which further illustrates the superiority of the proposed WEYOLOv3.

In the future work, the detection precision and speed will be further improved by further investigating the characteristics of PSPE images and videos.

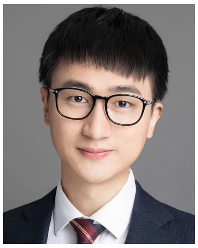
REFERENCES

[1] W. Cai, J. Le, C. Jin, and K. Liu, "Real-time image-identification-based anti-manmade misoperation system for substations," *IEEE Trans. Power Del.*, vol. 27, no. 4, pp. 1748–1754, Oct. 2012.
[2] B. A. S. Oliveira, A. P. De Faria Neto, R. M. A. Fernandino, R. F. Carvalho, A. L. Fernandes, and F. G. Guimaraes, "Automated monitoring of construction sites of electric power substations using deep learning," *IEEE Access*, vol. 9, pp. 19195–19207, 2021.
[3] O. Gómez, C. Portilla, and M. A. Ríos, "Reliability analysis of substation monitoring systems based on branch PMUs," *IEEE Trans. Power Syst.*, 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," *Electr. Power Sci. Eng.*, vol. 25, no. 9, pp. 76–78, Sep. 2009.
[5] P. H. M. de Andrade, J. M. M. Villanueva, and H. D. D. M. Braz, "An outliers processing module based on artificial intelligence for substations metering system," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3400–3409, Sep. 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 Trans. Power Syst.*, 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 *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo. (T&D-LA)*, Nov. 2010, pp. 417–421.
[8] M. J. B. Reddy, B. K. Chandra, and D. Mohanta, "A DOST based approach for the condition monitoring of 11 kV distribution line insulators," *IEEE Trans. Dielectr. Electr. Insul.*, 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 *Proc. 27th SIBGRAPI Conf. Graph., Patterns 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 *Proc. 28th Int. Conf. Image Vis. Comput. New Zealand (IVCNZ)*, 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 *Proc. 9th Int. Conf. Adv. Comput. Intell. (ICACI)*, Feb. 2017, pp. 201–205.
[12] B. Yogameena, K. Menaka, and S. Saravana Perumaal, "Deep learning-based helmet wear analysis of a motorcycle rider for intelligent surveillance system," *IET Intell. Transp. Syst.*, vol. 13, no. 7, pp. 1190–1198, Apr. 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. 94065–94075, 2020.
[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, 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 Trans. Pattern Anal. Mach. Intell.*, 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 Trans. Smart Grid*, vol. 11, no. 4, pp. 3593–3605, Jul. 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. 162073–162084, 2020.
[18] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.
[19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
[20] W. Liu *et al.*, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, Amsterdam, The Netherlands, Oct. 2016, pp. 21–37.
[21] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, "DSSD : Deconvolutional single shot detector," 2017, *arXiv:1701.06659*. [Online]. Available: http://arxiv.org/abs/1701.06659
[22] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2980–2988.
[23] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788.
[24] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 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. 41590–41597, 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 *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 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*, early access, Dec. 16, 2020, doi: 10.1109/ACCESS.2020.3045155.
[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," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 36, no. 2, pp. 180–196, Jun. 2020.
[29] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*. [Online]. Available: http://arxiv.org/abs/1804.02767
[30] L. Huang, Q. Fu, M. He, D. Jiang, and Z. Hao, "Detection algorithm of safety helmet wearing based on deep learning," *Concurrency Comput., Pract. Exper.*, vol. 33, no. 13, p. e6234, Feb. 2021.
[31] H. Shi, X. Chen, and Y. Yang, "Safety helmet wearing detection method of improved yolov3," *Comput. Eng. Appl.*, vol. 55, no. 11, pp. 213–220, Sep. 2020.
[32] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Autom. Construct.*, vol. 112, Jan. 2020, Art. no. 103085.

BAINING ZHAO received the B.Sc. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2019, where he is currently pursuing the M.Sc. degree. 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, QLD, Australia, in 2018. Since 2018, she has been working with Schneider Electric (Xiamen) Switchgear Company Ltd. Her research interests include HVDC and renewable energy.



ZHEWEN NIU received the 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 with the South China University of Technology, Guangzhou, China. He is also a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada, 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 the South China University of Technology, in 2018, where she is currently pursuing the master's degree in electrical engineering. Her main research interests include image processing and the application of artificial intelligence technology in power systems.

TONG QIAN (Member, IEEE) received the B.E. and Ph.D. degrees in electrical engineering from the South China University of Technology (SCUT), Guangzhou, China, in 2015 and 2020, respectively. He is currently working as a Postdoctoral Researcher in electrical engineering with SCUT. His research interests include distributed cooperation of networked multi-agent systems and the flexible operation of interdependent infrastructures in smart cities.

WENHU TANG (Senior Member, IEEE) 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, from 2004 to 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 has authored or coauthored 150 research articles, including 60 international journal articles and one 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. He is a fellow of IET.