Received: 4 October 2021 | Revised: 6 December 2021 | Accepted: 26 December 2021

IET Generation, Transmission & Distribution

DOI: 10.1049/gtd2.12387

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

Deep learning-based substation remote construction management and AI automatic violation detection system

Kai Yan[1,2,*] | Quanjing Li[1,2,*] | Hao Li[1,2] | Haifeng Wang[1,2] | Yuping Fang[2,4] | Lin Xing[3,2] | Yang Yang[1,2] | Haicheng Bai[1,2] | Chengjiang Zhou[1,2]

[1] School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China

[2] The Laboratory of Pattern Recognition and Artificial Intelligence, Yunnan Normal University, Kunming, China

[3] School of Physics and Electronic Information, Yunnan Normal University, Kunming, China

[4] School of Vocational and Technical Education, Yunnan Normal University, Kunming, China

Correspondence

Yuping Fang, The Laboratory of Pattern Recognition and Artificial Intelligence, Yunnan Normal University, Kunming 650500, China, and School of Vocational and Technical Education, Yunnan Normal University, Kunming 650500, China.
Email: 757004044@qq.com

Lin Xing, School of Physics and Electronic Information, Yunnan Normal University, Kunming 650500, China, and The Laboratory of Pattern Recognition and Artificial Intelligence, Yunnan Normal University, Kunming 650500, China.
Email: lin.xing@ynnu.edu.cn

Yang Yang, School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China, and The Laboratory of Pattern Recognition and Artificial Intelligence, Yunnan Normal University, Kunming 650500, China.
Email: yyang_ynu@163.com

[*]Kai Yan and Quanjing Li contributed equally to this work and should be considered co-first authors.

Funding information

National Natural Science Foundation of China, Grant/Award Number: [41971392]; Yunnan Province Ten-thousand Talents Program

Abstract

Workplace video surveillance and timely response to operational violations are critical to avoid operator injuries at power construction sites. Here, a system that combines remote substation construction management and artificial intelligence object detection techniques to intellectualize the power construction management process and identify violations during construction in real time is proposed. To improve the detection accuracy, a data augmentation method, including three operations: (1) object segmentation and background fusion; (2) partial erasing; and (3) other basic transformations, is also proposed. Six variants of the You Only Look Once (YOLO) model are trained as detectors for comparative experiments on a dataset collected at the practical construction site. The experimental results show that the detection precision and recall of the YOLOv5-s model are 0.852 and 0.922, with high accuracy and low miss rate, which meet the requirements of robustness and accuracy in detecting realistic power construction violations.

1 | INTRODUCTION

Power construction work is an important link in substation power transmission and distribution. It is essential for the safe and reliable operation of the substation electricity grid. Substation construction management system has been widely applied as the automation level of substations has improved. In power construction, a substation construction management system is used to manage work orders and monitor construction progress. Efficient work order management can substantially increase power construction efficiency, while real-time operation monitoring may effectively warn of safety incidents. As a result, an efficient and real-time substation constriction management system is particularly vital in power construction.

With the rapid development of 4G and 5G networks, remote construction management has become more convenient. As the result, many studies have focused on the design of substation construction management systems. These systems primarily use the Browser/Server (B/S) architecture to construct the substation management systems, digitize manual work order issuance, optimize the issuance process, and monitor the entire construction site. However, they do not implement remote verification and digital signature on construction sites, and their video surveillance system is still working as a traditional video monitoring model, with only basic functions such as video storage and playback, and no capacity to efficiently analyze the obtained video data. Furthermore, the monitoring of violations caused by substation construction operators is still based on artificial inspections, which does not fully utilize the potential of intelligent video surveillance technology and makes it difficult to identify operation violations with good efficiency, high accuracy, and strong timeliness.

The purpose of this paper is to propose an AI-based smarter substation remote construction management and monitoring system. A system is proposed for substation remote work order management and construction violation detection. In surveillance videos, a convolutional neural network (CNN) is used to detect irregularities. It can detect not only single object violations but also more complex multi-object violations. The detection network is trained end-to-end with a large training dataset that includes both practical images captured by realistic power construction sites and synthetic images generated by data augmentation.

The main contributions of this paper are as follows.

1) A substation construction management system based on AI technology, called Substation Remote Licensing Intelligent Safety Construction Management and Control System, is proposed to accomplish efficient work order management and real-time violations detection in surveillance.
2) A data augmentation method is proposed to augment the quantity and diversity of the dataset, whose original images are collected in realistic power construction sites, and exploit the augmented dataset for CNN training.

The remainder of this paper is organized as follows: The related work is introduced in Section 2. Section 3 presents the system framework. In Section 4, the detection methods for various objects are illustrated. In Section 5, our methods of data augmentation are explained in detail. The experiments conducted to evaluate our method are described in Section 6. Finally, the conclusions are presented in Section 7.

2 | RELATED WORK

In the past few years, many studies have focused on intelligence substation. Guo et al. [1] elaborated on the specific architecture and networking of intelligent substations, proposed the current problems in intelligent substations, and put forward a new system integration architecture design concept for the development trend of intelligent and integrated substation equipment. Zhang et al. [2] introduced the four modules of the upgraded intelligent distribution station monitoring system and then elaborated on the development of the intelligent distribution station inspection control system. Long [3] introduce an intelligent inspection system based on cell phone terminals, which is applied to flexible management and monitoring of substation equipment. In substation intelligence security, various sensors are commonly employed to detect abnormal conditions. Liang et al. [4] developed a system based on big data and artificial intelligence technology for intelligent security supervision system. Personnel authentication, intelligent warning, risk analysis, and all-round protection can be achieved through smart wearable sensors worn by workers.

The continuous development of 4G and 5G networks, as well as the widespread usage of video surveillance technologies, have made it possible to monitor substations in real time. Hu [5] analyse the defects in the application process of the existing video surveillance system for electric power enterprises. Then on this basis, the core technologies that will be applied in the new substation security video surveillance system are introduced. Lu et al. [6] propose an intelligent monitoring solution for substations in a big data environment. The solution is based on big data theory and intelligent analysis algorithms, which can assist monitoring personnel to fully understand the meaning of alarm signals and reduce the burden of substation duty personnel. However,current substation video monitoring systems are still primarily working as traditional video surveillance, with only simple functions, such as video capture, storage, and playback, which lacks the capacity to effectively analyse video data. Besides, the monitoring of substation workers is still based on artificial inspect and patrol, which does not fully use the potential of intelligent video surveillance technology.

With the continuous development of deep learning technology, CNN have become increasingly powerful. CNN such as region with CNN feature (RCNN) [7], You Only Look Once (YOLO) [8–10], and Single Shot Multibox Detector (SSD) [11] perform high accuracy, efficiency, and timeliness in object detection. To overcome the limitations of human inspection, automated detection techniques using CNN need to be developed to assist or replace human decisions. Ren et al. [12] discuss the concept of intelligent substation system monitoring based on deep learning as well as analyse the advantages and disadvantages of monitoring using traditional methods and deep learning. Li et al. [13] apply artificial intelligence technology to substation inspection to achieve automatic confirmation and alarm of abnormal conditions such as foreign objects, smoke, and flame in the station area through intelligent identification monitoring. Wang et al. [14] propose a region-based full convolutional network (R-FCN) multi-class detector to achieve multi-target detection and localization, and design a virtual electronic fence to prevent workers from entering hazardous areas and other non-working areas. Wang et al. [15] propose a convolutional neural network-based intelligent analysis of power field monitoring video data party, the use of video cloud storage platform for the original monitoring video storage and data cleaning, and classification analysis and labeling of hidden behaviour, to achieve intelligent early warning analysis of video big data. Tang et al. [16] introduced Gaussian mixed background algorithm on the basis of the traditional three-frame differential method to identify the active targets in the monitoring video more accurately and effectively, and provide effective support for unmanned monitoring of intelligent substations.

The majority of existing research in the area of safety construction monitoring are focused on helmet detection. Bo et al. [17] proposed a helmet detection method based on YOLOv3 with good accuracy performance. The helmet method proposed by Wu et al. [18] also use YOLOv3 as their Backbone network. Gu et al. [19] proposed a helmet detection method using YOLOv4. Then, Zhou et al. [20] proposed a helmet detection method based on YOLOv5 that achieved good performance in both detection accuracy and real time. Other than YOLO, Long et al. [21], Wu et al. [22], and Li et al. [23] proposed helmet detection method based on SSD, respectively, and also achieved good detection results. Furthermore, CNN have been also used in some studies to detect workers' safety vests [24–26], safety belts [27, 28], and insulators [29, 30] at power substations in surveillance video. Although the above studies have shown promising results in the applying of AI technology to the construction surveillance, however, all of the above studies focus on the detection of a single object, lacking simultaneous detection of multiple objects, moreover, the majority of their detection models are independent of the management system and cannot be integrated into one system to solve real-world problems more conveniently.

3 | SYSTEM FRAMEWORK

Our proposed Substation Remote Licensing Intelligent Safety Construction Management and Control System is composed of three main components: multi-video stream capture devices, AI detection server, and remote construction management platform. Our system, which is based on a traditional power substation work order management system, implements the remote intelligent work order issuance procedure and integrates AI technology to automatically identify violations in substation construction. The framework of our proposed system is shown in Figure 1.

Our system overcomes the drawbacks of traditional manual safety monitoring methods in the power construction process, such as low efficiency of work order issuance in substation management, small monitoring range, poor real-time performance, and a high missed detection rate. The original construction site surveillance video will be automatically saved as work order operation records for later viewing. When violations are detected during the construction process, an alarm is broadcast to the entire site via external speakers to remind workers of safety, and a remote call is initiated through the system to the principal of the site to remind workers of the safety regulations once more. At the same time, violations will be automatically

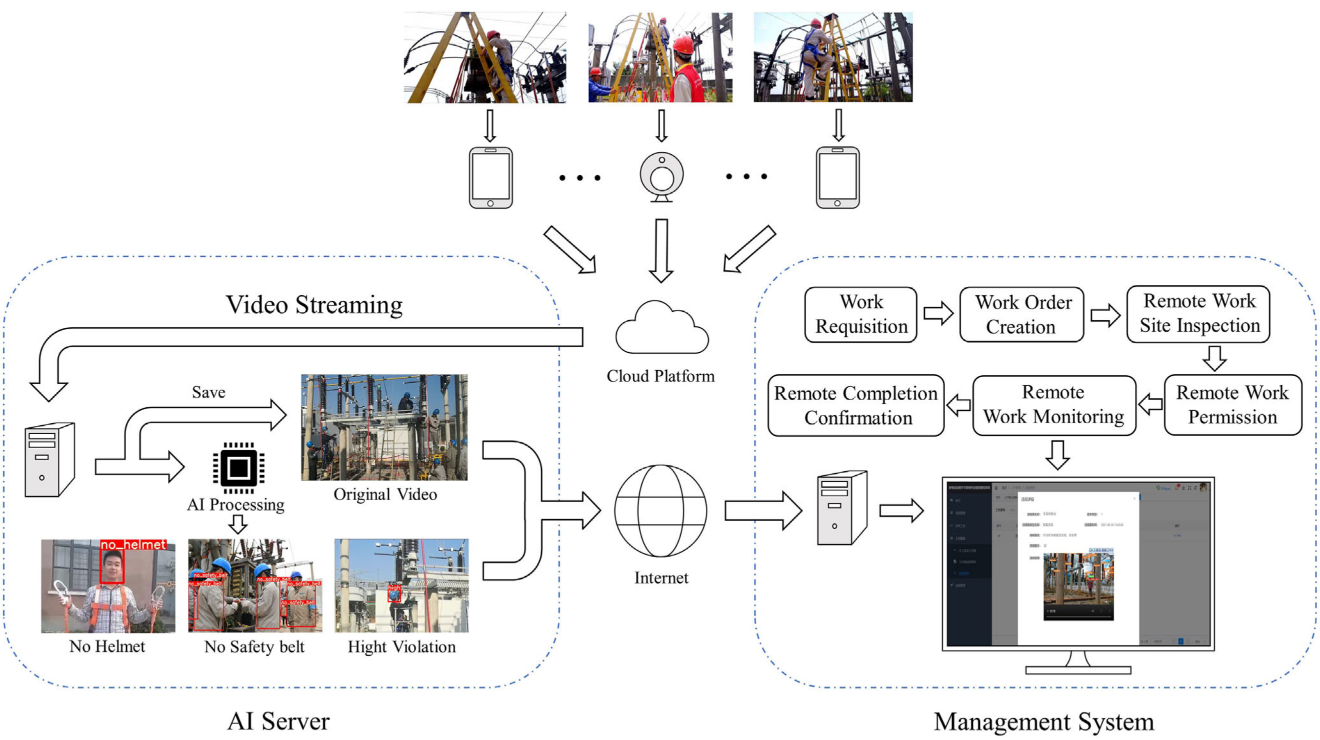

FIGURE 1 Framework of the Substation Remote Licensing Intelligent Safety Construction Management and Control System

classified and saved as violation video clips during the construction process to facilitate accountability and safety education for violators after the project is completed.

3.1 | Multi-video stream capture

Realistic power construction sites generally cover a wide range, making full coverage of the operation site monitoring difficult to obtain with a single video capture device. To overcome the limitations of a single video capture device, we capture video streams on power construction sites by using multiple video capture devices such as mobile cameras, iPads, and pan–tilt–zoom (PTZ) cameras. These multiple video stream capture devices are used to monitor and record real-time personnel wearing and operations on the construction sites from a variety of perspectives. The video stream is captured and uploaded to a cloud platform using Real-Time Streaming Protocol (RTSP) via a 4G or 5G network, then transmitted to an AI server through the cloud platform.

The whole power construction process is monitored by video capture devices fixed at various locations on the construction site. All video capture devices connect to the substation work order intelligent management system via a 4G or 5G network and collaborate with the Substation Remote Licensing Intelligent Safety Construction Management and Control System center to capture real-time video streams and monitor the entire construction process.

3.2 | AI detection server

In view of the poor timeliness and high false detection rate of traditional manual inspection and video monitoring to detect violations on construction sites, we use AI technology with high precision, good timeliness, and low cost to complete the detection of violations. In AI violation detection, the object detection algorithm in deep learning is used to constantly detect violations in video streams from operation sites. After the AI server receives the video streams transmitted from the cloud platform, the AI algorithm deployed on the server detects the wearing of operators and the distance between operators and power devices in the obtained video stream of a construction site. When a violation is detected, the video will be automatically classified marked and saved as a violation video clip.

While detecting violations, the AI server also continuously transmits the violation video clips which have been detected and saved, and unprocessed original surveillance videos to the remote construction management platform via the Internet. The platform then saves the original surveillance videos as operation records, and stores different types of operation violation video clips according to violation tags for classified management.

3.3 | Remote construction management platform

A work order is an important certificate for power substation construction. A power construction task must wait for the work order to be generated and reviewed before beginning construction. Work order management processes such as remote work order generation, approval and signing, remote permission, remote monitoring, and remote acceptance are all possible with the substation remote licensing intelligent management system. To ensure that work is completed smoothly, the work order issuer, work order principal, and work order licensor all work together to manage the substation work order process. The specific implementation process of the substation remote licensing management system will be introduced in detail in the following section.

3.3.1 | Devices binding

Work order issuer logs into the work order intelligent management system after determining the power construction task, fills in the department and purpose of video capture devices to be used on the construction site and generates a unique quick response (QR) code. The work order principal then logs into the system through a mobile device and scans the QR code, fills in the relevant information on the operation site, and binds it with the video capture devices.

3.3.2 | Work order generation

The work order issuer creates a new work order and checks to see if the content is complete and the construction site information is correct. The work order issuer will electronically sign the work order after it has been approved. If a work order is incorrectly filled out, it must be resubmitted and reviewed after revision.

3.3.3 | Work order permission

The work order is in a state of pending permission before power construction starts. The work order licensor must initiate a remote video call to the work order principal at construction site by the system, for identity verification and safety inspection on the power operation site. When the verification is complete, the work order principal will fill in construction start time and other information on the work order, as well as sign an electronic signature to authorize the start of power construction.

3.3.4 | Remote construction monitoring

At the start of power construction work, the work order principal is responsible for fixing the multiple video capture devices around the operation site in order to monitor the entire work process in real time. After being processed by AI server, the video streams from the construction site will be saved on the server of the Substation Remote Licensing Intelligent Safety Construction Management and Control System.

3.3.5 | Remote construction acceptance

When the work is completed, the work order licensor initiates a remote video call to the work order principal through the system to verify whether the power construction task has been completed and the operation site has been cleaned up. After passing the acceptance, the work order principal is required to sign the completion electronic signature.

4 | DETECTION METHOD

4.1 | Detection architecture

The YOLO series of algorithms is the most popular object detection algorithm in practical detection applications as the detection accuracy and speed are well balanced. Furthermore, it has faster speed and better precision in multiple objects detection.

The focus of this paper is on detecting violations in power constructions. In substation construction safety detection, objects which need to be detected, such as safety helmets, safety belts, insulators, and ladders, are multiple. Because the entire system must transmit video streams and issue timely alerts after a violation is detected, the real-time requirement of the detection model is high. Furthermore, owing to the uniqueness of power construction safety detection, once an operation is started, the system must remain operational throughout the entire process. As a result, the stability of the detection model is also critical. The YOLOv5 model is chosen as the detection model in our proposed power construction safety monitoring system since that it performs well in terms of detection precision, real-time performance, stability, and ease of deployment.

In the following section, we will start with four main single detection objects: safety helmet, safety belt, insulator, ladder, and two multi-object relationship detections – height violation and no safety belt wearing violation, to individually introduce the detection methods we used.

4.2 | Single object detection

4.2.1 | Safety helmet detection

A safety helmet is an essential protective tool for substation operators. The operator must wear a safety helmet, according to the power construction safety regulations. Operators in power construction work on complex sites surrounded by various electric devices. Any negligence could result in a threatening head injury because of the head hitting the power device. Therefore, operators must wear a safety helmet throughout the construction process to protect their own lives.

- Difficulty: The large number of operators wearing safety helmets, as well as the complex background of construction sites, are two major challenges in safety helmets detection. Traditional manual inspections and video surveillance methods for detecting safety helmet wearing have poor timeliness and detection precision.
- Method: To improve detection precision, a safety helmet dataset with sufficient positive and negative samples (operators wearing and not wearing safety helmets) is used for detection model training to accurately distinguish both of them in the same scene. Images collected in realistic sites under various viewing perspectives, definitions, and lighting conditions are simulated using data augmentation. The details of data augmentation methods will be described in Section 5.3. Meanwhile, in order to eliminate the phenomenon that the safety helmet cannot be detected correctly due to the detection model's missed detection as well as short-term perspective change and occlusion of the operator's head in a few individual frames of the videos, the detection model's output will be sent to DeepSort tracking network. By correlating the safety helmet detection results of the front and back frames, it can eliminate the fragments with different detection results of the front and back frames, and decrease missed detection of the safety helmet wearing, thereby reducing the violation records storage redundancy.

4.2.2 | Safety belt detection

There are scenes of aerial operations in practical power construction that require the operator to climb to a high place via ladder for operation. When working at a height, the operator faces a serious risk of falling from the ladder due to the complexity of the construction site and improper operation, posing a serious threat to life safety. Therefore, wearing a safety belt correctly is a critical guarantee for the life safety of operators working at height.

- Difficulty: The safety belt, with a unique structure, whose groundtruth contains a wealth of background information, often result in missed detection. In some complex scenes, a detection model trained by the dataset with insufficient data diversity finds it difficult to detect safety belts precisely.
- Method: To overcome the difficulty of safety belts detection, similarly to the data augmentation method used in the safety helmets dataset, a dataset of safety belts with sufficient positive and negative samples (operators wearing and not wearing safety belts) is augmented to simulate images captured in realistic construction sites. Simultaneously, the results of the safety belt detection will be sent to the tracking network to eliminate missed and false detection results caused by perspective changes and complex scenes in a few frames of the videos.

4.2.3 | Insulator detection

A large number of high voltage power lines are erected in the air with the help of insulators at substation construction sites. To avoid electric shock, operators should keep their working distance from insulators as far as possible. For this reason, when constructing, the insulator needs to be detected to determine whether the operator is at a safe distance from it.

- Difficulty: The insulators in the backgrounds have a high rate of missed detection. Because of their small size, the feature of background insulators in the background is fuzzy. The detection network trained by both large insulators with distinct features in construction sites and the small insulators in the background proved poor detection precision.

- Method: To ensure sufficient training samples, a large number of insulator images in realistic substations with various backgrounds are collected. Furthermore, we use partial erasing to remove the insulators in the background of each image in the insulator dataset. The detection network, trained on the insulator dataset after partial erasing, can filter out the insulators in the background and effectively detect insulators in construction sites.

4.2.4 | Ladder detection

In some realistic power construction scenes, the operator must stand on a ladder to operate at a height. The ladder is detected to determine whether the operator is standing and climbing it.

- Difficulty: The occlusion of the ladder caused by the operator climbing on it, as well as the complex background often makes detection difficult. However, most of the existing ladder datasets lack an occlusion relationship and only have simple backgrounds, hence training with them is insufficient for ladder detection in complex power construction sites.
- Method: Object segmentation and background fusion are used to augment the dataset in addition to traditional data augmentation methods. Ladders and operators with climbing action are segmented from some images and combined to generate lots of images of ladders occluded by climbing operators. Then the patches are merged into some specific backgrounds to generate large numbers of ladder images which contain plentiful background and are occluded by operators. The details of objects segmentation and background fusion will be described in Section 5.1.

4.3 | Multi-object relationship detection

4.3.1 | Height violation detection

The construction sites are full of insulators. Because the close proximity of insulators to operators can cause electric shock, operators are required to maintain a safe distance from insulators. Any operation distance that exceeds the safety distance is deemed as a violation.

- Difficulty: There are two major difficulties in height violation detection: objects pairing and distance measurement. Owing to massive insulators in construction sites, it is difficult to select an insulator as a reference to measure the distance from it to the head of operators with a safety helmet. Meanwhile, how to measure the safety distance between insulators and safety helmets is also a difficult task.
- Method: In order to solve the difficulty of objects pairing, the bounding box parameter of the insulator and operator's head with a safety helmet is used to measure distance. If the center point abscissa of one insulator bounding box is located in the range of the bounding box abscissa of one head with a safety helmet, this insulator will be paired with the operator. If any not, the distance between the upper edge center point of one head bounding box with safety helmet and the lower edge center point of all insulator bounding boxes will be measured, and the nearest insulator will be paired to this head with safety helmet as the reference. In addition, all paired insulators and heads with safety helmets are excluded from the above process. Objects pairing will be complete after repeating the above process until all heads with safety helmets are paired. In terms of distance measurement, according to the requirements of electric power construction safety regulations and our statistics on practical construction data, the safety distance between insulator and operator's head with safety helmet is approximately 1.5 times the width of the head bounding box. Any operation distance less than this safety distance is considered a violation.

4.3.2 | No safety belt wearing violation detection

The aerial operation, such as operating on the ladder, is demanded to wear a safety belt to avoid a life-threatening incident caused by a sudden fall of the aerial operator, which is most often caused by the complexity of construction sites and improper operations.

- Difficulty: Distance measurement is the major difficulty in no safety belt wearing violation detection. The distance between bottom edge of the ladder and operator's safety belt is used to determine whether the operator is climbing at height. How to measure a safety distance to determine whether operator at a height is difficult.
- Method: To address the difficulty of distance measurement, the bounding box parameter of ladder and body with or without safety belt is used similarly. When the center point abscissa of one body bounding box is located in the range of the bounding box abscissa of one ladder, this ladder will be paired with this body. Electric power construction safety regulations and statistics based on realistic operation data suggest that the distance of the bounding box lower edge between ladder and body less than 1.2 times the width of the body bounding box is the safe operation height. Any operator who operates at a distance greater than this height without a safety belt will be regarded as a violation.

5 | DATA AUGMENTATION

The entire dataset is made up of four parts: The dataset of safety helmet and safety belt containing positive and negative samples are taken from the counterpart in VOC2007 dataset; A portion of the ladder dataset is crawled by Web; while the others are artificially generated. The last dataset are insulator images which are all taken on-site at a power substation. In data-driven deep

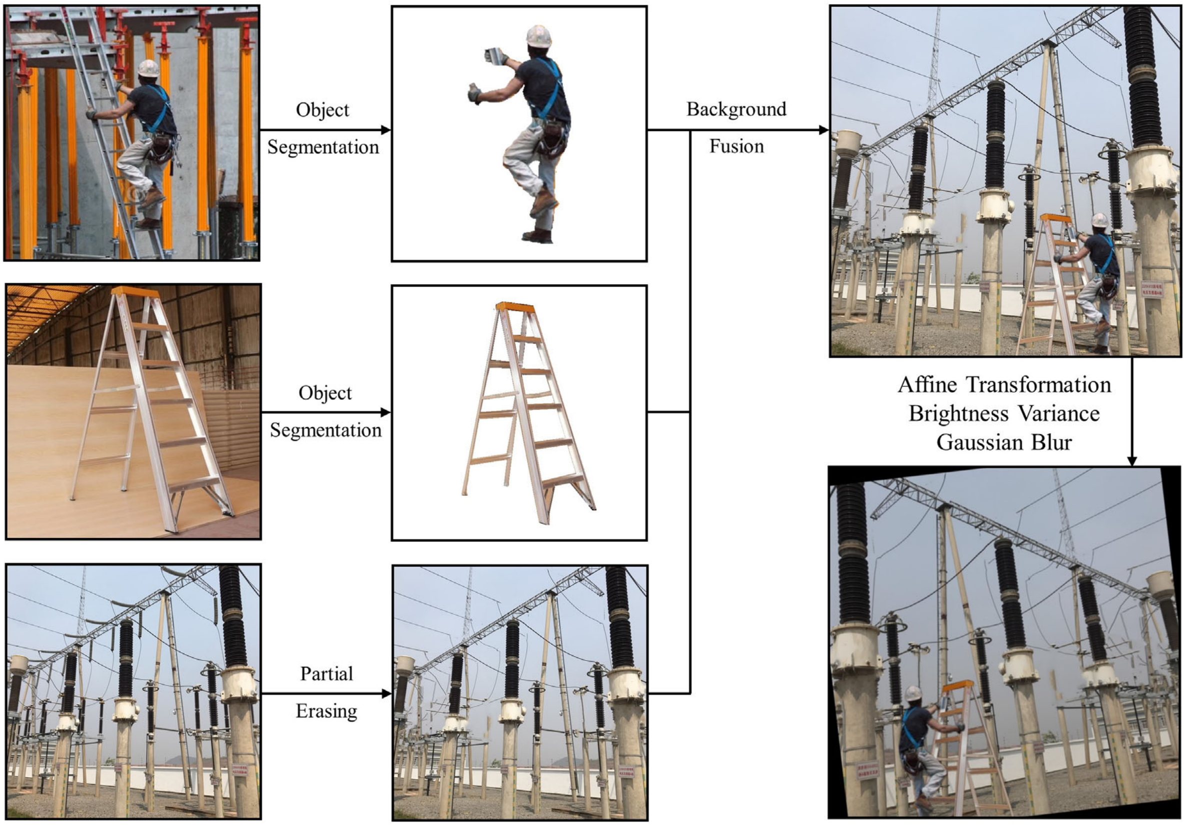

FIGURE 2 Data augmentation process

learning, in the case of a given network and limited prior knowledge, introducing more new data for training means the trained network is more robust. A common issue in object detection is that the training dataset is too small, which leads to poor detection capabilities of the trained network. Data augmentation is a strategy to increase the quantity and diversity of limited data. It extracts more useful information from limited data and generates value equivalent to more data. Therefore, data augmentation is necessary when the training sample size is limited. In our paper, the method of data augmentation is used to expand the dataset of some detection objects with insufficient training samples, and use the existing datasets to generate a new dataset. As shown in Figure 2, the data augmentation methods used in our paper mainly include five operations: (1) object segmentation and background fusion; (2) partial erasing; (3) affine transformation; (4) brightness transformation; (5) definition transformation. Each operation simulates various imaging environments in realistic surveillance video, such as viewing perspective changes, distance changes, background changes, definition changes, and illumination changes.

5.1 | Object segmentation and backgrounds fusion

The images used for training can be divided into two parts based on their role in training: (1) the foreground (detect targets such as safety helmets, safety belts, ladders, and insulators) and (2) the background (other image information). The background information in the ladders dataset is relatively simple when compared to the extensive background information in the images of safety helmets, safety belts, and insulators in the dataset. And in order to solve the problem that is not being detected in the practical operations due to the occlusion of the operators climbing on the ladder, the dataset images of the occlusion relationship formed by the person climbing on the ladder are also required. However, such images are insufficient in the ladder dataset. To solve the above problems, we add operators to ladders and generate sufficient training samples by putting them on background images.

The background images are taken in realistic construction sites and contain neither operators nor ladders. Ladders and operators with climbing actions are separated from some images using cutout software, and these patches images are merged into some images of people climbing ladders, then merged as a whole image into the background images without operators and ladders.

5.1.1 | Segmentation of climber and ladder

We select some typical images of people climbing ladders from the entire ladder dataset, segment the people inside, and then choose part of the ladder images from various perspectives that are suitable for fusion with segmented people in the dataset.

5.1.2 | Backgrounds fusion

Through segmentation, lots of image patches of climbers and ladders are obtained, then new sample images are generated by combining the image patches and fusing them in background images.

5.2 | Partial erasing

In realistic power construction, the devices are fixed to take videos of the nearby scene of electric power construction. In addition to the nearby detection objects in the view field, the distant backgrounds often contain insulators, which are one of the detection targets. The insulators in the background are small and blurry due to the far distance, causing losing their features in the videos. In the dataset that contains both the target insulators and the background insulators, only the target insulators are used as samples for training and the background insulators are ignored, then the trained detection model has a higher rate of missed detection; while when target insulators and background insulators are both used as the samples for training, the trained model has poor detection accuracy. Meanwhile, violations are detected in practical operation sites based on the positional relationship between the operator's safety helmet in the nearby operation site. The detection of background insulators will interfere with the judgment of violations at operating sites. Therefore, retouching software is used to partially erase the background insulators on each insulator image in the dataset to ensure the detection accuracy of the target insulators in the construction sites.

5.3 | Other methods of data augmentation

5.3.1 | Affine transformation

This paper use affine transformation to change the size, direction, and position of the ladders, insulators, operators wearing safety helmets and safety belts on the construction sites in the original images to simulate the on-site power construction scene captured by fixed cameras from different perspectives. The detection model's robustness can be improved by using these generated images for training.

5.3.2 | Brightness transformation

The brightness transformation is added to generate a dataset with brightness differences to simulate a realistic power construction scene and improve the detection model's robustness under different lighting conditions. Each image has a certain probability of randomly transforming its pixel intensity by 0.5 times to 1.5 times to adjust the image brightness.

5.3.3 | Definition transformation

Gaussian blur is also used in our data augmentation. We use the Gaussian kernel to convolute the image and transform images definition to get blurred images. Adding blurred images to the dataset can help the detection model better deal with the motion blur, out-of-focus blur, and other issues that arise during the detection process, thereby improving the robustness of the detection model.

FIGURE 3 Dataset of six classes' detection objects

6 | EXPERIMENTS

Here, we evaluate the detection performance of various detection methods through a series of experiments. First, we construct an experimental dataset and describe the experimental implementation details. Second, we discuss the advantages of our data augmentation method. Third, we compare the detection performances of various models and select the best one. Finally, detection results of images, which are collected in realistic power construction sites, are shown in detail.

6.1 | Experiment description

6.1.1 | Dataset construction

The dataset used in the experiment consists of six classes: (1) wearing a helmet; (2) no helmet wearing; (3) wearing a safety belt; (4) no safety belt wearing; (5) ladder; and (6) insulator. The former five classes of original images are either crawled from the Internet or taken in the power operation sites, while the last class of images is only taken on-site at the substation construction sites. The images of the above dataset are shown in Figure 3. A dataset with 14,840 images (with 30% data augmentation images), whose training dataset and validation dataset is at a ratio of 8:2, is obtained through image augmentation on

TABLE 1 Parameter setting of data augmentation

Parameter name	Value
rotation range	15°
scale range	0.8–1.2
horizontal flip probability	0.5
pixel intensity range	0.5–1.5
Gaussian Blur probability	0.5
Gaussian Kernel size	3

TABLE 2 Composition of the dataset

Class	Training	Validation	Total
Helmet(with & with out)	3632	908	4540
Safety_belt(with & with out)	4256	1064	5320
Insulator	480	120	600
Ladder	3504	876	4380
Total	11,872	2968	14,840

each class of images. The parameter settings of data augmentation are shown in Table 1. A testing dataset with 200 images, whose images captured by realistic power construction surveillance containingthe above six classes, is also collected. In addition, we adjust the number of images in each class so that the number of each class sample is basically the same. The detailed composition of the dataset is shown in Table 2.

6.1.2 | Experimental configuration

All of our experiments are based on the deep learning framework Pytorch. For model training, we use a computer with an Intel Core i7-10700 CPU, an NVIDIA RTX-3090 GPU with a memory of 24 GB, and 32 GB RAM. Our system is deployed on the Ubuntu operating system.

6.1.3 | Evaluation criteria

We use four widely used evaluation criteria to evaluate the performance of different detection models: Precision (P), Recall (R), F_1 Score (F_1), and mean Average Precision (mAP). The formula of P, R, and F_1 are defined as follows:

$$P = \frac{T_\mathrm{P}}{T_\mathrm{P} + F_\mathrm{P}} \tag{1}$$

$$R = \frac{T_\mathrm{P}}{T_\mathrm{P} + F_\mathrm{N}} \tag{2}$$

$$F_1 = \frac{2 * P * R}{P + R} \tag{3}$$

where F_P (false positive) and F_N (false negative) represent the number of incorrectly detected objects, and T_P (true positive) represent the number of correctly detected ones. $T_\mathrm{P} + F_\mathrm{P}$ is the total number of detected objects, and $T_\mathrm{P} + F_\mathrm{N}$ is the total number of actual objects. F_1 is the harmonic mean of P and R.

TABLE 3 Training parameter settings of detection models

Parameter name	Value
Epoch	300
Batch size	32
IoU threshold	0.2
Initial learning rate	0.001
Momentum	0.937

TABLE 4 Performance of dataset using different augmentations

Index	Training	Testing	Training data augmentation method	Precision
1	4000	200	Affine transform	0.672
2	4000	200	Affine transform + Brightness transformation + Definition transformation	0.719
3	4000	200	Affine transform + Background fusion	0.758
4	4000	200	Affine transform + Partial erasing	0.775
5	4000	200	Affine transform + Brightness transformation + Definition transformation + Partial erasing + Background fusion	0.802

6.1.4 | Implementation details

The goal of the detection system we proposed is to detect violations in practical power construction sites with high accuracy and efficiency. Given the good real time and accuracy of the YOLO model, we use six variant YOLO models for training to compare detection precision. These YOLO models are YOLOv3, YOLOv3-tiny (YOLOv3 with simplified network size), YOLOv3-spp (YOLOv3 with spatial pyramid pooling model), YOLOv5-s, YOLOv5-m, and YOLOv5-l (three variants of YOLOv5, respectively, with small, medium, and large network size). All models are trained on the basis of YOLO pretraining weights, and use the dataset we constructed above for training and validation. The detailed training parameter settings of series YOLO models are shown in Table 3.

6.2 | Effect of data augmentation

To evaluate the contribution of the data augmentation method used here to the overall experimental results, we conduct a series of data augmentation experiments on the dataset and compare the detection effects of these datasets after training in one detection model. Besides the training dataset, the experiments follow the same parameter settings as before. We build five datasets with the same number, where half of each dataset is the data augmentation images. The testing dataset of the five experiments, which is collected from the realistic power construction sites, is the same.

From Table 4, it can be seen that all of the data augmentation methods improve the precision. It is more effective to use background fusion and partial erasing than to use brightness transformation and definition transformation, while the combination of five types of data augmentation provides the best precision performance. This is because the diversity of the background affects the detection precision. In theory, the more diverse actual images in the training dataset, the better detection performance will be.

6.3 | Performance of different detection models

In our experiment, the performance of six YOLO models in a validation dataset is compared in order to choose a detec-

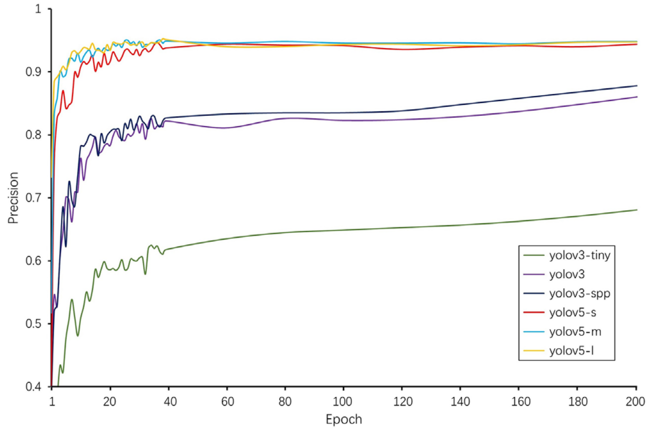

FIGURE 5 Comparison of training precision

tion model that can be applied to realistic power construction sites. The precision trend of six models in the first 200 epochs is depicted in Figure 5, and the final detection results are detailed in Table 5. In Table 5, YOLOv3-tiny has the best speed

FIGURE 4 The detection performance of different models. The green bounding box denotes operators with helmets and safety belts. The operators without helmets and safety belts, as well as height violations, are bounded by the red bounding box. The insulators and ladders are, respectively, bounded by the bounding box in orange and blue

TABLE 5 Performance of different detection models

Model	*P*	*R*	*mAP*	F_1	Speed (ms)
YOLOv3-tiny	0.719	0.891	0.865	0.796	2.8
YOLOv3	0.905	0.961	0.957	0.932	10.8
YOLOv3-spp	0.914	0.955	0.958	0.934	11.0
YOLOv5-s	0.946	0.953	0.964	0.949	8.6
YOLOv5-m	0.951	0.953	0.966	0.952	12.0
YOLOv5-l	0.951	0.960	0.967	0.955	14.0

performance, but the worst precision performance due to its simple model structure. Owing to a more complicated model structure, YOLOv5-l behaviors the best precision, while its speed is the worst. Compared with series of YOLOv3 models, series of YOLOv5 models have better performance in *P*, *mAP*, and F_1. In addition, the detection results of the above six models in the testing dataset are shown in Figure 4. As shown in Figure 4, the YOLOv5-s model has a detection precision for small objects that is not much worse than other more complex YOLOv5 models, and more importantly, its detection precision for safety belt and ladder is significantly better than all other models. Given almost the same precision performance as YOLOv5-l and the best speed performance besides YOLOv3-tiny, as well as its superior detection precision in realistic power operation sites, the YOLOv5-s model is finally chosen as the detection model in our system.

6.4 | Results in realistic power construction sites

To demonstrate the performance of our trained YOLOv5-s, we evaluate it in the testing dataset. The testing dataset is collected in realistic power construction sites and contains plentiful detection objects of all classes. The testing performance is shown in Table 6. Furthermore, according to different power construction sites, we divide the testing dataset into four operation scenes: (1) power inspection; (2) power operation; (3) aerial operation; and (4) night operation. We also detect different objects in the testing dataset according to different

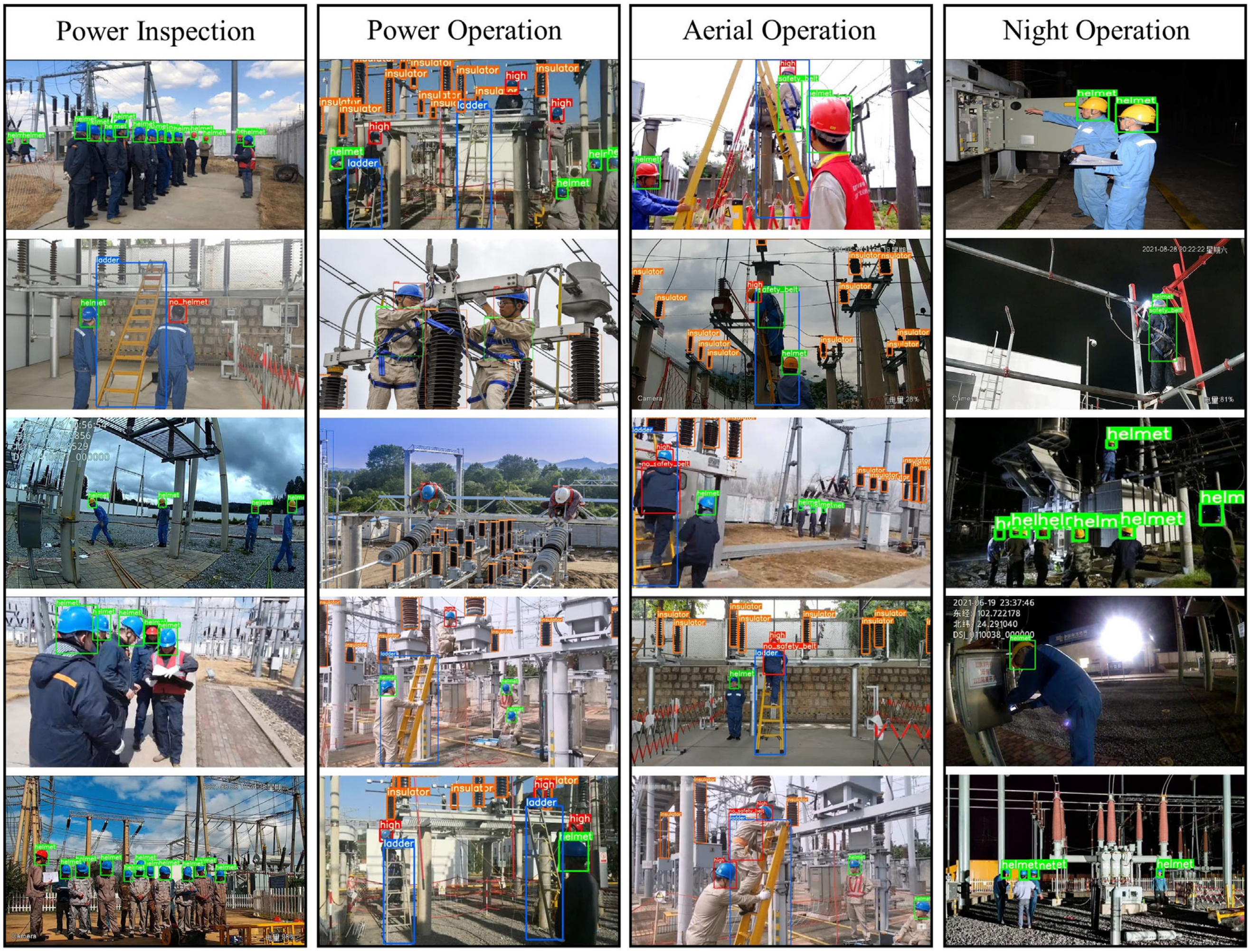

FIGURE 6 The detection performance of different models. The green bounding box denotes operators with helmets and safety belts. The operators without helmets and safety belts, as well as height violations, are bounded by the red bounding box. The insulators and ladders are, respectively, bounded by the bounding box in orange and blue

TABLE 6 Performance of YOLOv5-s on the realistic dataset

Class	*P*	*R*	*mAP*	F_1
no_safety_belt	0.971	0.911	0.971	0.940
safety_belt	0.612	0.938	0.792	0.741
no_helmet	0.834	0.991	0.966	0.906
helmet	0.969	0.977	0.985	0.973
insulator	0.858	0.824	0.885	0.841
ladder	0.880	0.889	0.923	0.884
all	0.852	0.922	0.921	0.886

construction scenes. The detection results are shown in Figure 6. As can be seen from Table 6 and Figure 6, our method can successfully detect all six classes in different practical power operation sites and has good detection performance in both each class and overall.

In addition, we use the testing dataset to test our entire system. The results show that our proposed system can detect at a speed of 8.9 ms per image, demonstrating that the distance logic judgment we added has a negligible impact on detection speed. Also, with good concurrency, our system can save and detect up to 40 different video streams at a frame rate of 30 frames per second in real time.

However, there are still some incorrect detection results because of poor detection precision and wrong logic judgment of the bounding box. In the future, we will improve the detection network structure and modify the algorithm logic of violation detection to detect the power operation violations more accurately.

7 | CONCLUSION

This paper proposes a power safety construction monitoring system that combines the traditional power construction surveillance process with AI to automatically detect violations in the power construction surveillance videos and save the violation clips. In order to compensate for the problem of insufficient training samples, this paper proposes a data augmentation method to better detect objects in realistic power construction scenes. The Substation Remote Licensing Intelligent Safety Construction Management and Control System proposed here has been put into use and is operating stably. The experimental results in the realistic power construction site dataset show that the power safety construction detection system we proposed can detect violations in practical power constructions with good accuracy and robustness.

ACKNOWLEDGEMENTS

The authors thanks to the experimental platform provided by the Laboratory of Pattern Recognition and Artificial Intelligence, also thanks to the dataset for power construction safety detection provided by Yunnan Power Grid Co Ltd Yuxi Power Supply Bureau. This work was supported by National Natural Science Foundation of China [41971392]; Yunnan Province Ten-thousand Talents Program.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The raw/processed data required to reproduce these findings cannot beshared at this time as the data also forms part of an ongoing study.

ORCID

Haifeng Wang https://orcid.org/0000-0002-5937-8812
Yang Yang https://orcid.org/0000-0002-4607-0501

REFERENCES

1. Guo, Y., Yuan, H., Zhuang, Y., Xu, J. & Zhang, Y. : Research on informatization construction of digital substation in smart grid. In: IOP Conference Series: Earth and Environmental Science, Vol. 632, p. 042089. IOP Publishing, Bristol, UK (2021)
2. Zhang, M., Fang, J., Mo, W., Wang, H., Ma, J., Luo, L. & Hao, F. : Development of the monitoring system in the smart distribution substation in guangzhou, china. In: 2020 International Conference on Wireless Communications and Smart Grid (ICWCSG), pp. 159–162. IEEE, New Jersey (2020)
3. Long, D.: Application of mobile intelligent inspection system in substation equipment management. Energy Power Eng. 9(4), 408–413 (2017)
4. Liang, Y., Wang, Z., Li, A., Jiang, Y., Duan, F., Li, Z., Wang, X., Yin, J.: Design of intelligent safety supervision system in electric power substation. In: 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2), pp. 2941–2944. IEEE, New Jersey (2019)
5. Jiangtao, H. : Discussion on the construction of substation security video surveillance system. In: IOP Conference Series: Materials Science and Engineering, Vol. 563, p. 032004. IOP Publishing, Bristol, UK (2019)
6. Lu, P., Zhang, F., Fan, S.: Research on substation intelligent monitoring scheme under big data environment. In: 2018 China International Conference on Electricity Distribution (CICED), pp. 1530–1536. IEEE, New Jersey (2018)
7. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587. IEEE, New Jersey (2014)
8. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. : You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788. IEEE, New Jersey (2016)
9. Redmon, J., Farhadi, A.: Yolo9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7263–7271. IEEE, New Jersey (2017)
10. Redmon, J., Farhadi, A.: Yolov3: an incremental improvement. arXiv:1804.02767 (2018)
11. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y. & Berg, A.C. : Ssd: single shot multibox detector. In: European Conference on Computer Vision, pp. 21–37. Springer, Berlin (2016)
12. Ren, B., Zheng, Y., Wang, Y., Sheng, S., Li, J., Zhang, H., Zheng, C.: Research status and prospect of deep learning in secondary state monitoring of smart substation. In: 2020 Asia Energy and Electrical Engineering Symposium (AEEES), pp. 669–677. IEEE, New Jersey (2020)
13. Li, J., Wang, H., Zhao, Y., Huang, R., Yang, S.: Application research of artificial intelligent technology in substation inspection tour. In: 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), pp. 1097–1100. IEEE, New Jersey (2019)
14. Wang, H., Zhang, X., Sun, Y., Li, J., Li, Y.: Research and application of artificial technology for substation environment surveillance system. In:

2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), pp. 901–905. IEEE, New Jersey (2019)
15. Xiaoyan, W., Pei, C., Rui, L., Bo, Y. & Wenzhao, N. : Research on intelligent analysis technology of power monitoring video data based on convolutional neural network. In: 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), pp. 2461–2464. IEEE, New Jersey (2020)
16. Tang, W., Chen, H.: Research on intelligent substation monitoring by image recognition method. Int. J. Emerging Electr. Power Syst. 22(1), 1–7 (2021)
17. Bo, Y., Huan, Q., Huan, X., Rong, Z., Hongbin, L., Kebin, M., Weizhong, Z., Lei, Z.: Helmet detection under the power construction scene based on image analysis. In: 2019 IEEE 7th international conference on computer science and network technology (ICCSNT), pp. 67–71. IEEE, New Jersey (2019)
18. Wu, F., Jin, G., Gao, M., Zhiwei, H. & Yang, Y.: Helmet detection based on improved YOLO v3 deep model. In: 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), pp. 363–368. IEEE, New Jersey (2019)
19. Gu, Y., Wang, Y., Shi, L., Li, N., Zhuang, L., Xu, S.: Automatic detection of safety helmet wearing based on head region location. IET Image Proc. 15(11), 2441–2453 (2021)
20. Zhou, F., Zhao, H. & Nie, Z. : Safety helmet detection based on YOLOV5. In: 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA), pp. 6–11. IEEE, New Jersey (2021)
21. Long, X., Cui, W. & Zheng, Z. : Safety helmet wearing detection based on deep learning. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), pp. 2495–2499. IEEE, New Jersey (2019)
22. Wu, J., Cai, N., Chen, W., Wang, H., Wang, G.: Automatic detection of hardhats worn by construction personnel: a deep learning approach and benchmark dataset. Autom. Constr. 106, 102894 (2019)
23. Li, Y., Wei, H., Han, Z., Huang, J., Wang, W.: Deep learning-based safety helmet detection in engineering management based on convolutional neural networks. Adv. Civ. Eng. 2020, 9703560 (2020)
24. Hung, H.M. & Hong, H.S. et al.: A deep learning-based method for real-time personal protective detection. Le Quy Don Technical University-Section on Information and Communication Technology (LQDTU-JICT) (13), 23–34 (2019)
25. Jeon, S.Y., Park, J.H., Youn, S.B., Kim, Y.S., Lee, Y.S., Jeon, J.H.: Real-time worker safety management system using deep learning-based video analysis algorithm. Smart Media J. 9(3), 25–30 (2020)
26. Ali, Z., Park, U.: Real-time safety monitoring vision system for linemen in buckets using spatio-temporal inference. Int. J. Control Autom. Syst. 19(1), 505–520 (2021)
27. Li, J., Wang, T., Li, Y., Tian, Y., Wang, S., Zhang, M., Zhai, Y., Sun, S. & Zhao, X. : Working clothes detection of substation workers based on the image processing. In: 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI), pp. 180–184. IEEE, New Jersey (2017)
28. Zhao, B., Lan, H., Niu, Z., Zhu, H., Qian, T., Tang, W.: Detection and location of safety protective wear in power substation operation using wear-enhanced YOLOV3 algorithm. IEEE Access 9, 125540–125549 (2021)
29. Tao, X., Zhang, D., Wang, Z., Liu, X., Zhang, H., Xu, D.: Detection of power line insulator defects using aerial images analyzed with convolutional neural networks. IEEE Trans. Syst. Man Cybern. Part A Syst. 50(4), 1486–1498 (2018)
30. Zhao, Z., Zhen, Z., Zhang, L., Qi, Y., Kong, Y., Zhang, K.: Insulator detection method in inspection image based on improved faster R-CNN. Energies 12(7), 1204 (2019)

How to cite this article: Yan, K., Li, Q., Li, H., Wang, H., Fang, Y., Xing, L., Yang, Y., Bai, H., Zhou, C.: Deep learning-based substation remote construction management and AI automatic violation detection system. IET Gener. Transm. Distrib. 16, 1714–1726 (2022). https://doi.org/10.1049/gtd2.12387