AN INTELLIGENT FUSION OBJECT-DETECTION ALGORITHM FOR SMART SUBSTATION SYSTEM

Ziqiang Zhang

Abstract

Machine learning is playing an increasingly important role in smart substation systems. Object detection algorithms are commonly used in smart substations for procedures, such as helmet detection and personnel clothing inspection. However, object detection algorithms are inadequate for solving complex smart substation scenarios because of their poor generalisation ability. Thus, we introduce an intelligent fusion algorithm named YYSF-4 that has good generalisation ability. YYSF-4 comprises You Only Look Once (YOLO) V1, YOLO V3, a single-shot multi-box detector, and fast-oriented text spotting, and is suitable for use in smart substations. We use real images from substations as a dataset to verify the effectiveness of the YYSF-4 in four scenarios: helmet detection and recognition, personnel clothing detection and identification, personnel detection and identification, and bill detection and recognition. The experimental results show that the mean average precision (mAP) of YYSF-4 in the above four scenarios is higher than the mAPs of other baseline algorithms.

Key Words

Object detection, safety helmet detection, personnel clothing inspection, personnel detection, bill detection

1. Introduction

The safety of substations has always been one of the most important issues in power systems. For example, existing video surveillance systems of substations cannot function as intrusion alarm [1], which means that the timely discovery of unauthorised intrusions into substations is impossible. In addition, under the current transportation inspection mode, the information acquisition method is the traditionally used and the single source of information. Equipment status perception is based mainly on power outage maintenance and offline testing. Thus, advanced methods, such as intelligence and data utilisation, are rarely employed; substation operations and maintenance continue to rely largely on manual operation.

State Grid Corporation of China Gansu Electric Power Company, Lanzhou, Gansu, 730030, China; e-mail: 5084498@qq.com

The use of multi-objective optimisation methods in constructing a new integrated energy system is a current focus of the research [2], [3]. The operation and inspection of substations face enormous challenges given the increasing number of substations. Presently, however, there remains a large difference between traditional and smart substations [4]. Although many intelligent algorithms exist for safety helmet detection [5], personnel dress code detection [6], and moving object detection [7] in substations, they are only adequate for providing early warning of specific types of events, and their generalizability to other scenarios is very poor. The need for smart substations is growing as the role of substations in the electric power system is increasingly crucial. One of the main ways to realise smart substations is to develop smart algorithms to provide real-time and all-round early warning. In this vein, we introduce herein an intelligent fusion algorithm, YYSF-4, for use in power systems. YYSF-4 is an improvement of four typical algorithms and achieves outstanding performance, as is verified in scenarios involving the detection of safety helmets, personnel clothing, personnel, and bills.

2. Related Research

In terms of physical security at smart substations, Huang *et al.* [5] improved the You Only Look Once (YOLO) V3 algorithm and applied it to helmet-wearing detection. Their results showed that the algorithm improved helmet detection speed and accuracy. Zhang *et al.* [6] developed a personnel dress code detection algorithm based on a cascade convolutional neural network. They evaluated this algorithm on private datasets, which revealed that its mean average precision (mAP) was 56.9%. In addition, Wang *et al.* [7] detected moving objects with a Gaussian-model foreground extraction and image dithering approach, which they used to identify dynamic objects in video surveillance footage from smart substations.

In terms of information security of smart substations, Li *et al.* [8] investigated the problem of zero data-frame loss when communication systems of smart substations fail. Focusing on the reliability of smart substation communication networks, the authors proposed a system

model by combining a parallel redundancy protocol with high-availability seamless redundancy network protocols. The feasibility of the method was verified, and the results showed that it improved the reliability of the communication network, providing a reference for developing smart substations. However, currently available automatic information-monitoring acceptance systems of smart substations fail to realise the closed-loop real-time transmission of an entire signal. Chen *et al.* [9] developed a set of closed-loop acceptance systems for smart substation information monitoring based on IEC61850 and multi-machine multi-network parallel verification theory. They also developed a high-efficiency substation data-monitoring method based on all-phase spectrum correction technology. The theory and method were then verified in an engineering project of a provincial power company. Zhang [10] studied the problems of partial or complete relay protection malfunction, which are caused by poor secondary maintenance of the equipment of a smart substation. To realise safety measures for error prevention, online monitoring, and false alarms, the characteristics of the safety measures for the secondary maintenance of the smart substation were first analysed. Subsequently, the key security technologies for preventing errors were expounded in terms of visualisations, anti-error mechanisms, simulation rehearsals, remote control operation, and online monitoring. It was then demonstrated that the application of these technologies enhances safety and security.

Although the aforementioned methods effectively improve security in smart substations, the generalizability of the best method is poor. Moreover, each algorithm is only applicable to a limited number of scenarios. This means it is vital to develop a security protection algorithm that is robust and has high generalizability.

3. YYSF-4

YYSF-4 is an improvement on and fusion of four algorithms, and its structure is shown in Fig. 1. The design principle of YYSF-4, the design of the loss function, and the selection of the corresponding parameters are detailed below.

3.1 Design Principle

The design principle of YYSF-4 comprises five main parts. First, the original image is directly divided into small squares that do not overlap, and feature maps of specified sizes are produced through convolution. Second, multiple scales are used to detect targets of different sizes: the shallow feature map contains the location information, and the deep feature map contains extensive semantic information. Third, convolutional neural networks are used to extract image features. The location and category of the candidate region are also predicted to improve the generalisation ability of the model. Fourth, the positioning and detection in the forward calculation of the network are encapsulated to improve the calculation speed. Fifth, image text detection and recognition are used for joint
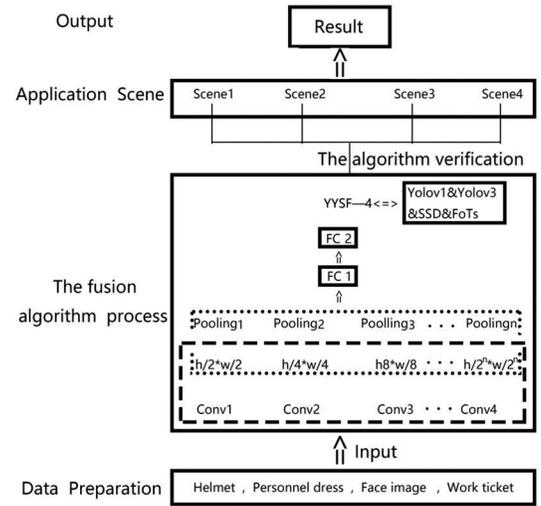

Figure 1. Algorithm fusion of YYSF-4.

training, and both tasks share the convolutional feature layer to save computation time.

Figure 1 shows the structure of YYSF-4, which comprises four effective components: data preparation, algorithm fusion, typical scenario application, and result output. The data preparation component comprises helmets, personnel attire, face data, and work tickets. Two types of helmets are used: red helmets, representing normal staff, and white helmets, representing visitors. The algorithm fusion process is an integration of YOLOV1, YOLOV3, a single-shot multi-box detector (SSD), and fast-oriented text spotting (FOTS). When identifying the job site, the targets are fed to the algorithms in different sequences through a set of instructions, which allows for individual identification. Then, the optimal result is extracted using a set of instructions, which yields the optimal recognition result. YYSF-4 first performs feature processing (such as convolution and pooling) of the input data, and then, after repeated training, it is applied to four different scenarios for verification.

3.2 Loss Function Selection

The goal of YYSF-4 is to optimise the loss function. Four loss functions are selected for this model. The first loss function is shown in (1). Here, L_{coord} is the coordinate error, L_{obj} is the IOU error, and L_{classes} is the classification error.

$$Loss_{N1} = L_{\text{coord}} + L_{obj} + L_{\text{classes}} \qquad (1)$$

L_{coord} can be expressed as :

$$L_{\text{coord}} = \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^{B} 1_{ij}^{obj} \left[(x_i - \widehat{x}_i)^2 + (y_i - \widehat{y}_i)^2 \right]$$

$$+ \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^{B} 1_{ij}^{obj} \left[\left(\sqrt{w_i} - \sqrt{\widehat{w_i}} \right)^2 \right.$$

$$+ \left. \left(\sqrt{h_i} - \sqrt{\widehat{h}_i} \right)^2 \right], \qquad (2)$$

where λ_{coord} is the hyperparameter, (x, y, w, h) are the position coordinates, and 1_{ij}^{obj} denotes that the j-th bounding box predictor in cell i is "responsible" for that prediction.

L_{obj} can be expressed as :

$$L_{obj} = \sum_{i=0}^{s^2} \sum_{j=0}^{B} 1_{ij}^{obj} (c_i - \widehat{c}_i)^2 +$$

$$+\lambda_{noobj} \sum_{i=0}^{s^2} \sum_{j=0}^{B} 1_{j}^{noobj} (c_i - \widehat{c}_i)^2, \qquad (3)$$

where λ_{noobj} is the hyperparameter, and c is the confidence of the prediction box.

L_{classes} can be expressed as :

$$L_{\text{classes}} = \sum_{i=0}^{s^2} 1_i^{obj} \sum_{c \in \text{ classes}} (p_i(c) - \widehat{p}_i(c))^2, \qquad (4)$$

where 1_i^{obj} denotes whether an object appears in cell i, and $p_i(.)$ is the probability of a particular class.

The second loss function is shown in (5), where λ is the weight constant, and its function is to control the ratio between the detection frame loss, obj confidence loss, and $noobj$ confidence loss.

$$loss_{N1} = \lambda_{box} \sum_{i=0}^{N_1 \times N_1} \sum_{j=0}^{3} 1_{ij}^{obj} \left[(t_w - t'_w)^2 + (t_y - t'_y)^2 \right]$$

$$+\lambda_{box} \sum_{i=0}^{N_1 \times N_1} \sum_{j=0}^{3} 1_{ij}^{obj} \left[(t_w - t'_w)^2 + (t_h - t'_h)^2 \right]$$

$$-\lambda_{obj} \sum_{i=0}^{N \times N} \sum_{j=0}^{3} 1_{ij}^{obj} \log(c_{ij})$$

$$-\lambda_{noobj} \sum_{i=0}^{N_1 \times N_1} \sum_{j=0}^{3} 1_{ij}^{noobj} \log(1 - c_{ij})$$

$$-\lambda_{\text{class}} \sum_{i=0}^{N_1 \times N_1} \sum_{j=0}^{3} 1_{ij}^{obj} \sum_{c \in \text{ classes}} \left[p'_{ij}(c) \log(p_{ij}(c)) \right.$$

$$\left. + (1 - p'_{ij}(c)) \log(1 - p_{ij}(c)) \right] \qquad (5)$$

The third loss function is shown in (6), where and represent the score loss function and position loss function, respectively, and N represents the number of grid points. The solutions of and are shown in (7) and (8).

$$L(x, c, l, g) = \frac{1}{N} (L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g)) \qquad (6)$$

$$L_{\text{loc}}(x, l, g) = \sum_{i \in \text{ Pos}}^{N} \sum_{m \in \{x, c, w, h\}} x_{ij}^k \text{smooth}_{L1} (l_i^m - \widehat{g}_j^m) \,(7)$$

$$L_{\text{conf}}(x, c) = -\sum_{i \in \text{ Pos}}^{N} x_{ij}^p \log(\widehat{c}_i^p) - \sum_{i \in \text{ Neg}} \log(\widehat{c}_i^0)$$

$$-\sum_{i \in \text{ Neg}} \log(\widehat{c}_i^0), \qquad (8)$$

where

$$\widehat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)}. \qquad (9)$$

The fourth loss function is shown in (10).

$$L = L_{\text{detect}} + \lambda_{\text{recog}} L_{\text{recog}} \qquad (10)$$

The detection branch loss function is:

$$L_{\text{detect}} = L_{\text{cls}} + \lambda_{\text{reg}} L_{\text{reg}}, \qquad (11)$$

where

$$L_{\text{cls}} = \frac{1}{|\Omega|} \sum_{x \in \Omega} H(p_x, p_x^*),$$

$$= \frac{1}{|\Omega|} \sum_{x \in \Omega} (-p_x^* \log p_x - (1 - p_x^*) \log(1 - p_x)); \quad (12)$$

$$L_{\text{reg}} = \frac{1}{|\Omega|} \sum_{x \in \Omega} \text{IoU}(R_x, R_x^*) + \lambda_\theta (1 - \cos(\theta_x, \theta_x^*)). \,(13)$$

The recognition branch loss function is as follows:

$$L_{\text{recog}} = -\frac{1}{N} \sum_{n=1}^{N} \log p(y_n^* \mid x), \qquad (14)$$

$$p(y^* \mid x) = \sum_{\pi \in B^{-1}(y^*)} p(\pi \mid x). \qquad (15)$$

3.3 Parameter Adjustment

The first parameter-selection step of the improved YOLO V1 is the adjustment of the input image to 448 × 448 and the running of a single convolutional network on the image. The entire network has 24 convolutional layers and two fully connected layers. The output layer uses a linear function as the activation function, and the activation functions of the other layer are leaky rectified linear units. In the improved YOLO V3, the first feature map is downsampled 32 times, the second feature map is downsampled 16 times, and the third feature map is downsampled 8 times. The output dimension of the feature map is $N \times N \times [3 \times (4 + 1 + 8)]$, where $N \times N$ is the number of grid points of the output feature map, giving a total of three anchor boxes. In the improved SSD algorithm, six layers are extracted from the feature map generated by the convolutional neural network. In the improved FOTS algorithm, the shared convolution layer adopts a convolution sharing method similar to U-Net, which merges the bottom and high layers of the feature map.

4. Experiments

The experimental scenarios are ticket detection and recognition, helmet detection and recognition, personnel dress detection and recognition, and personnel detection and recognition. FOTS, YOLO V1, SSD, and YOLO V3 are used for anti-violation recognition.
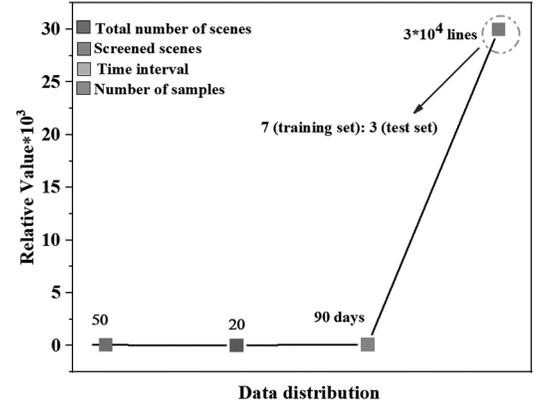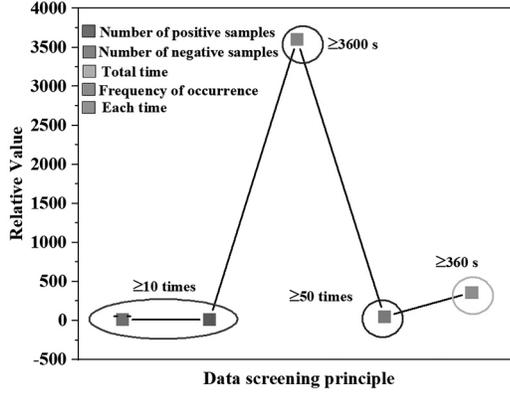
Figure 2. Basis of data filtering.

Figure 3. Data distribution before and after data preprocessing.

4.1 Data Processing

4.1.1 Data Acquisition

The experimental data comprise three parts: positive and negative samples of substation staff wearing safety helmets, positive and negative samples of substation staff wearing overalls, and samples of substation staff during normal operations and maintenance. The original data requires preprocessing as it contains numerous bad data [11]. The preprocessing conditions are as follows. For the safety helmet samples, the colour of the safety helmets must be obvious and free from obstruction and alteration. The total occurrence of both positive and negative helmet samples in the scene must be more than 10, and the total duration of an occurrence must exceed 3,600 s. For personnel dress samples, the samples must be clearly identifiable, and the dress code must be intact, unobstructed, and unaltered. In addition, people in the scene must enter and exit the safe area more than 50 times, and the total duration of an occurrence should exceed 360 s. Figure 2 shows the basic conditions for data screening.

4.1.2 Data Preprocessing

The first data preprocessing step is data filtration. Through careful comparison, we screen the video data of 50 scene cameras in a certain substation, which were obtained from May 2020 to July 2020. In the later stage, through demand screening, 30 scenes are excluded, and the remaining scenes meet the corresponding video production standards for detection and identification requirements. The second data preprocessing step is video frame cutting: different data intervals are selected based on the specific video data for each camera, and Open CV is then used to cut the video frames. If the video contains a flow of many people, the cutting-frame time interval is short, to ensure that enough image samples are obtained; if the video contains a flow of few people, the cutting-frame time interval is long, to ensure that there is a large difference between the images. The third step is data cleaning. As the images cut from the video frame are strongly correlated in time, the images that have similar timestamps and are minimally different in terms of the information they contain are removed, to

Table 1
Hyperparameter Values of Initial Network Parameters

Hyperparameter	Value
Convolution kernel size	3×3
Convolution operation step size	1
Pool size	2
Pooling step	2
Batch size	8
Image size	224×224
Initial learning rate	0.0001
Do b_n processing?	Yes
Weight-decay regular term	0.0005
Number of image channels	3

avoid model overfitting during training. The fourth step is image labelling, which is performed with Label Image, such that the category and location information of the target are labelled. The final step is sample selection, which is conducted following the sample classification index; 70% of the sample is used for model training and 30% is used for model verification. Figure 3 shows the data distribution.

4.2 Experimental Design

4.2.1 Safety Helmet Detection and Recognition

In this experiment, YYSF-4 uses the non-maximum suppression method and determines the category of each box. Ninety-eight boxes are initially processed by returning values less than the confidence threshold to 0. Subsequently, we determine the category of each box. The detection result is output when the confidence value is not 0. GoogLeNet is used as the backbone network for model training. The initial model parameters are shown in Table 1.
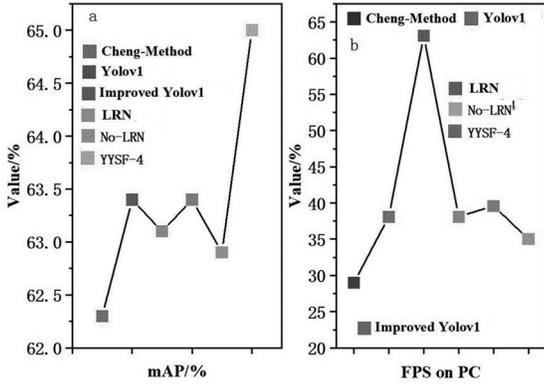
Figure 4. Comparison of YYSF-4 with other algorithms in helmet detection scenarios: (a) mAP and (b) fps.

Figure 5. Comparison of YYSF-4 with other algorithms in personnel clothing inspection scenarios: (a) mAP and (b) processing rate.

Table 2
Initial Network Parameters

Hyperparameter	Value
Number of iterations	300
Batch size	10
Image size	512×512
Learning rate	0.0005
Multi-scale training	True
Training method	Atrous

Table 3
Initial Network Parameters

Hyperparameter	Value
Number of iterations	200
Batch size	8
Image size	608×608
Learning rate	0.001
Multi-scale training	True
Training method	Adam

We use accuracy, frames per second (fps), and mAP as model evaluation indicators. Figure 4(a) shows the mAP profile of YYSF-4, which reaches 65%. This mAP of YYSF-4 is 2.7%, 1.6%, 1.9%, 1.6%, and 2.1% higher than the mAPs of the Cheng Method, YOLO V1 [12], improved YOLO V1 [12], local response normalisation (LRN) [12], and No-LRN [12], respectively. The mAPs of YOLO V1 and LRN are both 63.4%, which is the highest after that of YYSF-4. These mAPs are the same because both algorithms share a similar foundation. Figure 4(b) shows that LRN achieves the fastest processing rate, whereas the processing rate of YYSF-4 is 33 fps, ranking it fifth out of the six algorithms tested. However, in actual application scenarios, mAP is more important than the processing rate. Additionally, we test the F1 value of YYSF-4 and obtain a value of 65.42%, which is higher than that of the other algorithms. Therefore, the above indicators of YYSF-4 show that it is the best choice. These results demonstrate that applying YYSF-4 in substations can assist personnel in decision making.

4.2.2 Personnel Clothing Detection and Identification

The initial model parameters are shown in Table 2.
ResNet-50 is used as the model for training the backbone network for the model application. Figure 5(a) shows that the mAPs of the seven algorithms fluctuate around 70% in personnel clothing detection, with the
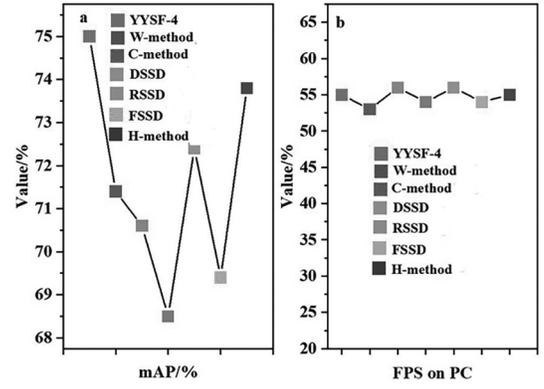
mAP of YYSF-4 being the highest (75%). The mAPs of the W-method [13], the C-method [14], the rainbow SSD (RSSD) [15], and the H-method [16] are all greater than 70%. The mAPs of the deconvolutional single-shot detector (DSSD) [17] and the feature fusion SSD (FSSD) [18] are less than 70%, with the mAP of the latter being the lowest of the seven algorithms. Figure 5(b) shows the processing rates for each of the seven algorithms; those of the C-method and the RSSD are equal-fastest (56 fps) and that of YYSF-4 ranks second (55 fps). The DSSD and W-method have the lowest processing rates (54 and 53 fps, respectively). Importantly, the processing rates of the seven algorithms are similar, differing by only approximately ±2 fps.

4.2.3 Personnel Detection and Identification

Table 3 shows the initial network parameters of personnel detection and identification.
On this basis, we verify and tune the model. At the same time, multiple indicators, such as accuracy, processing speed, and mAP, are used as evaluation indicators for model effect evaluation.
Figure 6(a) shows the five algorithms that are used to simulate the personnel detection scene. The mAP of YYSF-4 is 20.48%, which is 9.31%, 24.76%, and 18.05% higher than the mAPs of the other four algorithms, respectively. Faster R-CNN [19] has the second-best mAP,
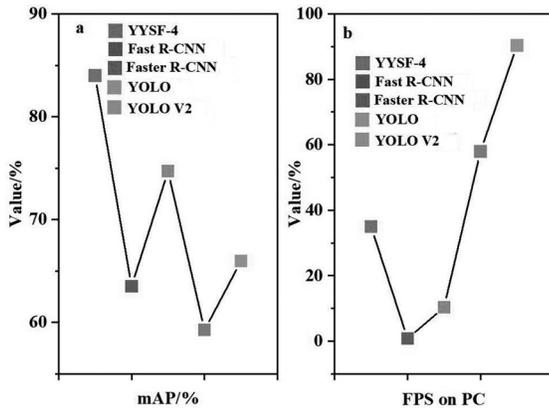
Figure 6. Comparison of YYSF-4 with other algorithms in personnel detection and identification scenarios: (a) mAP value and (b) fps value.

Figure 7. Comparison of YYSF-4 with other algorithms in bill detection and recognition scenarios: (a) P value and (b) F1 value.

Table 4
Text-recognition Branch Settings

Type	KernelT [size, sride]	Out channels
conv_bn_relu	[3,1]	64
conv_bn_relu	[3,1]	64
height-max-pool	[(2,1),(2,1)]	64
conv_bn_relu	[3,1]	128
conv_bn_relu	[3,1]	128
height-max-pool	[(2,1),(2,1)]	128
conv_bn_relu	[3,1]	256
conv_bn_relu	[3,1]	256
height-max-pool	[(2,1),(2,1)]	256
bi-directionalIstm	/	256
fully-connected	/	|S|

which is approximately 10% less than that of YYSF-4. YOLO has the lowest mAP, which is 24.76% less than that of YYSF-4. The order of the other algorithms, in terms of descending mAPs, is YOLO V2, Fast R-CNN, and YOLO. Figure 6(b) shows that the fastest processing rate is that of YOLO V2 (90.36 fps), while the second fastest is that of YOLO (∼60 fps), and the third fastest is that of YYSF-4 (35 fps). The above data show that the processing rates and mAPs of the algorithms show different trends: YYSF-4 is the best choice in terms of mAP, whereas its processing speed was in the middle range.

4.2.4 Bill Detection and Recognition

Table 4 shows the text recognition branch settings of YYSF-4 in the experiment. Figure 7 shows the P and F1 values of the test results for YYSF-4 in comparison with those of SegLink [20] and the rotation region proposal
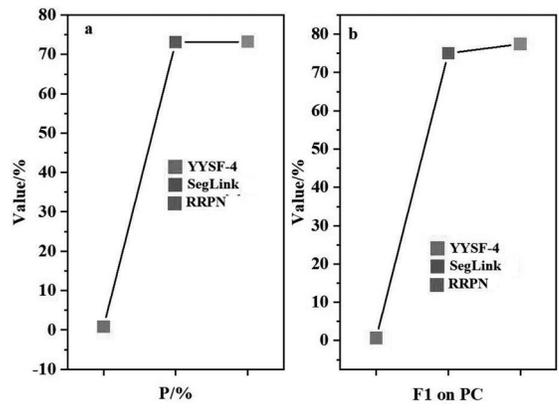
network (RRPN) [20]. Figure 7(a) demonstrates that the P value of YYSF-4 is the best ($P = 80.95\%$), as it is 7.85% and 7.72% higher than those of SegLink and the RRPN, respectively. SegLink and the RRPN have similar P values, differing by only 0.03%. Figure 7(b) shows that the F1 value of the RRPN is the highest, followed by that of SegLink and then YYSF-4. However, the best choice is YYSF-4 according to the crucial indicators.

5. Conclusion

In this study, we develop an intelligent fusion algorithm, YYSF-4, for use in smart substation systems, by improving four typical algorithms. YYSF-4 is verified in four scenarios and achieves outstanding performance. In a helmet detection experiment, the mAP of YYSF-4 is 2.7%, 1.6%, 1.9%, 1.6%, and 2.1% higher than that of the Cheng-Method, YOLO V1, improved YOLO V1, LRN, and No-LRN, respectively. In the personnel dress detection experiment, the mAP of YYSF-4 is the highest (75%). In the personnel detection experiment, the mAP of YYSF-4 is 20.48%, 9.31%, 24.76%, and 18.05% higher than that of YOLO V2, Fast R-CNN, and YOLO, respectively. In the bill detection experiment, the P value of YYSF-4 is the best (80.95%), as it is 7.85% and 7.72% higher than that of the other algorithms, respectively.

References

[1] S. Hussain, J.H. Fernandez, A.K. Al-Ali, and A. Shikfa, Vulnerabilities and countermeasures in electrical substations, *International Journal of Critical Infrastructure Protection*, *33*, 2021, 100406. https://doi.org/10.1016/j.ijcip.2020.100406

[2] A. Mahmoudan, P. Samadof, S. Hosseinzadeh, and D.A.A. Garcia, A multigeneration cascade system using ground-source energy with cold recovery: 3E analyses and multi-objective optimization, *Energy*, *233*, 2021,121185. https://doi.org/10.1016/j.energy.2021.121185

[3] A. Sohani, M.Z. Pedram, K. Berenjkar, H. Sayyaadi, S. Hoseinzadeh, H. Kariman, and M.E.H. Assad, Techno-energy-enviro-economic multi-objective optimization to determine the best operating conditions for preparing toluene in an industrial setup, *Journal of Cleaner Production*, *313*, 2021, 127887. https://doi.org/10.1016/j.jclepro.2021.127887

[4] F. Pang, Q. Li, Q. Bu, Z. Zhou, and C. Zhang, New generation smart substations, *IEC 61850-based smart substations.* (London: Academic Press, 2019), 359–421. https://doi.org/10.1016/B978-0-12-815158-7.00009-3

[5] L. Huang, Q. Fu, M. He, D. Jiang, and Z. Hao, Detection algorithm of safety helmet wearing based on deep learning, *Concurrency and Computation: Practice and Experience*, *33*(13), 2021, e6234. https://doi.org/10.1002/cpe.6234

[6] Z. Na, Y. Zechuan, H. You, B. Xiaoan, and S. Yifan, Personnel dress code detection algorithm based on convolutional neural network cascade, *Proc. 2nd Int. Conf. on Machine Learning, Big Data and Business Intelligence (MLBDBI),* Taiyuan, 2020, 215–221. DOI: 10.1109/MLBDBI51377.2020.00047.

[7] Y. Wang, J. Zhang, L. Zhu, Z. Sun, and J.A. Lu, A moving object detection scheme based on video surveillance for smart substation, *Proc. 14th IEEE Int. Conf. on Signal Processing (ICSP)*, Beijing, 2018, 500–503. DOI: 10.1109/ICSP.2018.8652316.

[8] H. Li, X.J. Zhang, H. Pan, W.Q. Mao, Y.B. Yan, and L.F. Chen, Research on a reliability evaluation method of a communication network for an intelligent substation, *Power System Protection and Control*, *49*(9), 2021, 165–171. DOI: 10.19783/j.cnki.pspc.200807.

[9] Y.Q. Chen, J.H. Chen, J.B. Qiu, B.J. Liu, and P. Du, Research on automatic acceptance system of monitoring information in a smart substation, *Power System Protection and Control*, *48*(11), 2020, 143–150. DOI: 10.19783/j.cnki.pspc.190366.

[10] Y. Zhang, Study on the technology of errors prevention for secondary maintenance of intelligent substation, *Electric Power System Equipment*, *18*(1), 2020, 139–140.

[11] J.X. Zhao, S.S. Guo, and D.J. Mu, DouBiGRU-A: Software defect detection algorithm based on attention mechanism and double BiGRU, *Computers & Security*, *111*, 2021, 102459. https://doi.org/10.1016/j.cose.2021.102459

[12] Y.Q. Chen, Z.Z. He, Z. Ma, R.X. Bi, and H.Y. Mao, Intelligent target detection algorithm for embedded FPGA, *Microelectronics & Computer*, *38*(6), 2021, 87–92. DOI: 10.19304/j.cnki.issn1000-7180.2021.06.015.

[13] J.W. Wen, Y.W. Zhan, C.H. Li, and J.B. Lu, Design of Atrous filter to strengthen small object detection capability of SSD, *Application Research of Computers*, *36*(3), 2019, 861–865+872. DOI: 10.19734/j.issn.1001-3695.2017.09.0967.

[14] H.J. Chen, Q.Q. Wang, G.W. Yang, J.L. Han, C.J. Yin, J. Chen, and Y.Z. Wang, SSD object detection algorithm with multi-scale convolution feature fusion, *Journal of Frontiers of Computer Science and Technology*, *13*(6), 2019, 1049–1061.

[15] J. Jeong, H. Park, and N. Kwark, Enhancement of SSD by concatenating feature maps for object detection, *arXiv preprint arXiv:1705.09587*, 2017.

[16] Q.S. Hou and J.S. Xing, SSD object detection algorithm based on KL loss and grad-CAM, *Acta Electronica Sinica*, *48*(12), 2020, 2409–2416.

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

[18] Z.X. Li and F.Q. Zhou, FSSD: Feature fusion single shot multibox detector, *arXiv preprint arXiv:1712.00960*, 2017.

[19] X.X. Yue, J.X. Jia, X.D. Chen, and G.A. Li, Road small target detection algorithm based on improved YOLO V3, *Computer Engineering and Applications*, *56*(21), 2020, 218–223.

[20] B.J. Zou, W.J. Yang, M. Liu, and L. Z. Jiang, A three-stage text recognition framework for natural scene images, *Journal of Zhejiang University (Science Edition)*, *48*(1), 2021, 1–8.

Biographies

Ziqiang Zhang graduated from Hohai University. He is currently the Deputy Chief Engineer of the State Grid Gansu Electric Power Company Division. The majors are data monitoring and analysis and big data application, data asset management and quality control, digital new technology management and application, hydropower dispatching, and comprehensive utilisation of water resources. During his work, he has won a number of awards and honorary titles, including 2 second prizes for innovation achievements. In addition, he has won 2 first prizes, 2 second prizes, and 2 third prizes of provincial scientific and technological progress awards.