ISSN: 2664-9640

Summary of Power Field Violation Detection Based on Deep Learning

Zhu Shi [1], Sijing Deng [2], Yuhang Fu [2]

[1] School of Sichuan University of Science & Engineering, Zigong 643000, China;

[2] School of Sichuan University of Science & Engineering, Zigong 643000, China;

Abstract

With the rapid development of deep learning and computer vision, related technologies have been widely used in many fields. Violation detection occupies an extremely important position in target detection. Real-time monitoring and detection of personnel in the construction site through video surveillance equipment is a very important application scenario. At present, many construction sites still use manual methods to observe whether the constructors violate the rules, but this method has great limitations. Human energy is limited. As the working time increases, the efficiency will gradually decrease. In response to this situation, this article mainly introduces the identification of three types of violations of behavior, wear, equipment and facilities. Compared with manual methods, it is more efficient and greatly saves labor costs.

Keywords

Deep learning, Target Detection, Construction Site, Violation detection.

1. Introduction

Traditional violation detection is mainly a "person detection person" operation mode, but it is limited by the working hours and patrol range of supervisors, and cannot meet the needs of the current complex construction sites. With the continuous development of image processing technology, the traditional detection of violations on the construction site has gradually changed to detecting the surveillance video screen, extracting the characteristics of the target and identifying it. Wear testing mainly detects the wearing of personal protective equipment of construction workers, such as: Liu Yunbo et al. [1] judged the color and wearing condition of construction workers' helmets by detecting the distribution of pixels in the video, and warned workers who did not wear helmets to ensure the safety of construction workers as much as possible; Action behavior detection is mainly to detect human behavior. For example, Zhang Jinlin et al. [2] proposed a monocular visual recognition algorithm that uses a single camera to recognize and track multiple semaphore actions, which greatly improves the acquisition of information; Equipment and facility testing is mainly to detect the status of construction site equipment. For example, literature [3] introduces infrared thermal imaging technology and machine vision graphics technology to realize image forming, status identification, fault diagnosis and maintenance of power equipment, which greatly improves The accuracy of condition monitoring and the response efficiency of the system reduce the burden on staff. Compared with manual monitoring, these methods can achieve reasonable allocation of resources and effective management of personnel. The improvement of the construction environment and the safety of construction personnel have been greatly improved. However, there are many uncertain factors in the actual construction environment, such as rain, snow and shelter problems, so these methods have great limitations. In recent years, with the continuous development of computer technology, deep learning has been greatly developed. Many researchers combine deep learning with target detection, which has many advantages

ISSN: 2664-9640

compared with traditional target detection algorithms. Neural network-based target detection can be roughly divided into two categories: one is a region-based regional convolutional neural network, Such as Fast R-CNN, Faster R-CNN; The other is the regression-based SSD (Single Shot Multi Box Detector) and YOLO (You Only Look Once) series.

2. Organization of the Text

2.1. Wear violation detection

In the existing construction environment, it is very important for the personnel on the construction site to wear personal protective equipment such as safety helmets, safety belts and protective glasses. These are the protective equipment of the workers. In the event of an accident, they can effectively protect the workers. Safety, to ensure the safety of their lives and property. Traditional wear detection is basically performed for the task of image recognition, using machine learning and image processing technology to achieve, but it requires manual design of features, which has high requirements on the environment, and the manually extracted features are not robust. , And the generalization ability is poor. The target detection algorithm based on CNN is divided into two-step method and one-step method. The design idea of two-step method: first use CNN network or algorithm to acquire a priori box (anchorbox), and then perform classification and position regression. Representative algorithms mainly include R- CNN[4], Fast R-CNN[5], Faster R-CNN[6]; and the one-step method uses end-to-end design ideas, and directly performs anchor acquisition and category in the same CNN And location prediction, representative algorithms include YOLO, SSD, YOLOv2 and YOLOv3, etc. [7-11]. Comparing the two types of algorithms, the one-step method, the monitoring speed is fast, the model training is simple, but the detection accuracy is slightly weaker; but the detection accuracy of the YOLOv3 algorithm is faster than the detection rate.

Traditional algorithms mainly use skin color, head and face information with image processing and machine learning to achieve. For example, literature [12] uses the YcbCr color model to locate the human face with skin color, and then combines image processing and neural network to achieve wearing detection. Literature [13] uses a combination of hu matrix and support vector machine to complete the detection. Literature [14] uses the Vibe algorithm to detect the human body, then uses the embossing algorithm to detect the head, and finally uses the HOG algorithm to extract features and combine with SVM to achieve helmet wearing detection. Literature [15] uses a deformable component model as a feature carrier, and combines the geometric shape, texture, and color features of the image to achieve this. Literature [16] divides the entire detection part into two parts, first combining the frequency domain information of the image and the direction gradient histogram for worker detection, and then combining the color and ring Hough transform features for helmet detection.

Xu Shoukun et al. [17] aimed at the difficulty and low accuracy of existing helmet wearing detection algorithms for partial occlusions, different sizes and small targets, and proposed an improved Faster R-CNN and multi-component combination. Safety helmet wearing detection method. Use multi-scale training and increase the number of anchor points on the original Faster R-CNN to enhance the robustness of the network in detecting targets of different sizes, and introduce an online difficult sample mining strategy to prevent the imbalance of positive and negative samples, and then wear a safety helmet on the detected ones. Workers and safety helmets use a multi-component combination method to eliminate false detection targets. Compared with the original Faster R-CNN, it improves the relative detection accuracy and is more adaptable to the environment. In [18], in view of the complex application environment of helmet detection and identity recognition, the relatively small frame size of the target, and the diversity of scenes, a method of helmet wearing detection and identity recognition based on improved Faster R-CNN was proposed. Based on the original Faster R-CNN, the feature layers of multiple stages are fused and multi-scale detection is performed, and the size of the candidate target frame is modified to achieve the optimal. Xiao Tigang et al. [19] aimed at the problem of low detection accuracy and slow detection rate of helmet wearing in intelligent monitoring, and proposed an improved wearing detection algorithm YOLOv3-WH on the basis of YOLOv3. By increasing the input scale, adding depth can be separated. Convolutional structure; using four scales to detect, increase the up-sampling structure layer by 4 times; and optimize the K-Means clustering algorithm to make the loss function convergence value smaller and the average IOU value larger. In the comparison of experimental data, YOLOv3-WH Compared with YOLOv3, the number of detection frames per second (FPS) has increased by 64%, and the average detection accuracy (mAP) has increased by 6.5%, which improves the detection speed and greatly improves the accuracy of the detection. Cao Yan et al. [20] aimed at when the detected object is small, because it occupies a small pixel in the picture. The problem of feature information loss occurs in the process of feature extraction. An improved model based on the SSD model-the hierarchical SSD model is proposed. Through the hierarchical processing of the feature pyramid, Conv-LSTM is introduced to separate the high-feature information from the low-feature information layer. Make full use of the feature information of small targets to adapt to the complex environmental background of construction sites and difficult problems of small targets.

2.2. Behavior violation detection

Existing construction sites have the characteristics of many sources of danger, uneven quality of construction personnel, and complex cross-construction surfaces. Therefore, supervision of dangerous behaviors of construction personnel has always been indispensable, including smoking, illegal leaning, and illegal crossing of fences. As well as illegally climbing high-rise buildings, etc., effectively identifying these violations can greatly improve the supervision efficiency of the construction site, realize intelligent safety management, and promptly alert the offenders to ensure the safety of construction personnel. Action recognition based on deep learning is an end-to-end method. The classification is completed by self-learning behavior representations in videos on the network, and has a wide range of applications in the safety supervision of construction workers. According to the characteristics of the backbone network, the action recognition network based on deep learning mainly includes: (1) 2D CNN; (2) 3D CNN; (3) Spatio-temporal decomposition network. Among them, 2D CNN is mainly used for single-channel images and multi-channel images; 3D CNN is realized by convolving a 3D kernel to a video clip; the spatiotemporal decomposition network mainly includes spatiotemporal decomposition and convolution of decoupled spatiotemporal filters And the separation of spatiotemporal features and channel separation convolution [21].

Reference [22] aimed at detecting the behavior of the fence crossing on the job site, and proposed a method combining two-dimensional convolution and three-dimensional convolution, in which three-dimensional convolution was used to extract the temporal features of the input clip; two-dimensional convolution was used to extract the spatial features of the current frame , To solve the positioning problem, introduce the SE module on the basis of the three-dimensional convolution architecture 3D-ResNext-101 to improve the accuracy of the algorithm; use Darknet-19 as the two-dimensional convolution architecture to extract features, perform channel fusion and classification regression, and achieve fence crossing Behavior detection and recognition. Shi Jiacheng et al. [23] proposed the LRDCN model on the basis of LRCN[24], using the dual-stream model and CNN to extract video features, using GRU to correlate the front and back features, and using encoding-decoding and maximum likelihood estimation methods. Known action sequences introduce new character actions, thereby updating network parameters faster and increasing network depth under limited computing speed conditions.

ISSN: 2664-9640

Fan Yinhang et al. [25] proposed a human action recognition algorithm that combines a 3D convolution residual network and a lightweight multi-scale convolution module. The intermediate feature map is divided into several sub-images, and then the 3D convolution Feature map fusion, extract the multi-scale features of the target and assign different weights to each channel, and finally classify. Literature [26] proposed a three-dimensional dense convolutional network (3D DenseNet) based on attention mechanism for human action recognition in video, using dual-stream network as the basic framework, spatial network using three-dimensional dense network for feature extraction, and finally The final result is obtained after the fusion of spatiotemporal features and the fusion of the classification layer. Literature [27] proposed a recognition method based on a three-dimensional convolutional neural network to solve the problem that 2D CNN extracts features from a single frame but does not include the motion information encoded between consecutive frames in the actual video.

In [28], in order to make full use of the spatio-temporal features in the video to improve the accuracy of action recognition, a spatiotemporal feature fusion action recognition framework using a sparse sampling scheme is proposed. The sparse sampling is used to obtain the RGB image and the optical flow image of the video, which are sent to The spatio-temporal features extracted by the VGG-16 network are merged with the middle-level spatio-temporal fusion features and sent to the 3D CNN to identify the action category. Literature [29] proposes a graph convolution model based on the spatio-temporal fusion of attitude motion (PM-STFGCN) in order to fuse the motion characteristics of the spatio-temporal domain, using the spatio-temporal fusion module (PM-STF) based on attitude motion to fuse the temporal and spatial domains. Spatial features and self-adaptive enhancement features.

2.3. Equipment and facilities violation detection

In recent years, with the continuous development of deep learning computer vision technology, image processing technology has become more and more widely used in online equipment monitoring. It can not only effectively replace manual inspections, improve work efficiency and automation, but also work under high pressure and danger. In harsh environments, real-time monitoring of equipment status can effectively ensure the safety and reliability of equipment operation [30]. The traditional method is mainly divided into time domain feature extraction and frequency domain feature extraction, and statistical analysis of numerical data, such as maximum value, minimum value, mean value, variance, etc., is mainly used for feature extraction. Compared with traditional recognition algorithms, target recognition algorithms based on deep convolutional neural networks such as Faster-RCNN, SSD and YOLO can automatically learn target features from a large amount of image data.

For human action recognition, machine learning methods mainly include traditional SVM, decision tree, and naive Bayes. Literature [31] used acceleration sensors and gyroscopes, placed these sensors in different positions of the human body for data collection, and used Naive Bayes, decision trees and K-nearest neighbors for identification and classification. Literature [32] used acceleration sensors, magnetic field sensors, gyroscopes and linear acceleration sensors, and the recognition rate was only slightly improved with multiple sensors. Literature [33] extracted 43 features, including maximum, minimum, mean, variance, peak interval, numerical interval distribution, etc. Literature [34] uses neural networks, but PCA is used for feature extraction, and there are some statistical features in the time domain.

Compared with traditional detection methods, the target recognition algorithm based on deep learning has a higher accuracy rate. Zhang Xiaohua et al. [35] designed a recognition algorithm based on image processing and deep learning in order to realize the detection and recognition of the equipment position in the cabinet. The specific operation is to establish the equipment image data set to train the YOLO model, obtain the appropriate weighting coefficient and make the loss function The value is the smallest, and the actual environment is detected and identified.

Literature [36] uses convolutional neural networks to implement deep learning pattern recognition through convolution operations, pooling operations, and network structure and settings to identify the pressure, speed, component swing, and abnormal state of component swing speed of manufacturing equipment. Lu Wanjie et al. [37] proposed the use of SVM algorithm to establish the type recognition of coal mine equipment based on deep learning, and the type of coal mine equipment and equipment serial number corresponded one by one to realize the precise identification and classification of coal mine equipment by inspection robots. . Gao Lu et al. [38] proposed a Faster-RCNN-based fault detection and recognition method in order to realize the automatic judgment of equipment failures. By collecting monitoring images around the equipment, training the images and obtaining the corresponding model, the equipment can be switched on and off. Recognition of status, indicator light abnormality and display abnormality. Literature [39] uses an improved YOLO target detection framework, uses convolutional neural network to extract features, uses batch normalization method to standardize the model, and finally predicts the target bounding box through the RPN network to realize the positioning and abnormality of the cable equipment State recognition. Literature [40] uses the Mobilenet-SSD deep network model to train the map features in the database, and deploys the trained model parameters to the edge computing device to realize the identification and location of the equipment oil leakage.

3. Conclusion

This article mainly summarizes the three types of violation detection on the construction site, and summarizes the practical application of each type of violation detection. Compared with the traditional method, the detection method based on deep learning improves the accuracy while ensuring real-time performance. It also has a good detection accuracy for the situation where there are many targets and overlapping parts on the construction site. Deep learning has high speed and accuracy, and has gradually replaced traditional target detection algorithms to become the mainstream.

Acknowledgements

This work is supported by Science and Technology Project Funding of State Grid Corporation of China(Contract No.521997180016)；

The Project of Sichuan provincial science and Technology Department (Grant No.,2019YJ0477, 2018GZDZX0043,2018JY0386,2020YFG0178);

The artificial intelligence key laboratory of Sichuan province Foundation (2019RYY01).

Enterprise informatization and Internet of things measurement and control technology key laboratory project of Sichuan provincial university (2018WZY01,2019WZY02).

The Project of Sichuan Provincial Academician (Expert) workstation of Sichuan University of Science and Engineering(2018YSGZZ04).

The Project of Zigong Science and Technology Bureau(2019YYJC13).

Graduate Innovation Fund of Sichuan University of Science and Engineering(y2020019)

References

[1] Liu Yunbo, Huang Hua. Research on Monitoring Technology of Wearing Safety Helmet on Construction Site[J]. Electronic Science and Technology, 2015, 28(04): 69-72.

[2] Zhang Jinlin, Wang Lulu, Rui Ting, Rong Xiaoli. Semaphore recognition system based on neural network [J]. Journal of PLA University of Science and Technology (Natural Science Edition), 2003(04): 65-68.

[3] Qu Yinlin. Research on automatic identification of power equipment based on digital image processing[J]. Management and Technology of Small and Medium-sized Enterprises (First issue), 2019(02):168-169.
[4]Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.
[5]Girshick R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2015:1440-1448.
[6]Ren Shaoqing,He Kaiming,Girshick Ross,Sun Jian. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.[J]. IEEE transactions on pattern analysis and machine intelligence,2017,39(6) .
[7]Redmon, Joseph, Divvala, Santosh, Girshick, Ross, et al. You Only Look Once: Unified, Real-Time Object Detection[J].,2015.
[8] Wang Song, Fei Shumin. Research and improvement of SSD (Single Shot MultiBox Detector) target detection algorithm [J]. Industrial Control Computer, 2019, 32(04): 103-105.
[9]Hsu WeiYen,Lin WenYen. Ratio-and-Scale-Aware YOLO for Pedestrian Detection.[J]. IEEE transactions on image processing: a publication of the IEEE Signal Processing Society,2020,PP.
[10]Redmon J, Divvala S, Girshick R, et al. You only look once:Unified,real-time objectdetection[C]// Proceedings of the IEEE conference on computer vision and pattern recognition.2016:779-788.
[11]Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European conference oncomputer vision.Springer, Cham, 2016: 21-37.
[12]Chai D, Bouzerdoum A. A BAYESIAN APPROACH TO SKIN COLOR CLASSIFICATION IN YCBCR COLOR SPACE[C]// Tencon. IEEE, 2000.
[13] Liu Xiaohui, Ye Xining. Application of skin color detection and Hu moment in helmet recogniti on[J]. Journal of East China University of Science and Technology (Natural Science Edition), 2014, 40(03): 365-370.
[14]Li Qirui. Research and implementation of a helmet video detection system based on human body recognition[D]. University of Electronic Science and Technology of China, 2017.
[15]Jia Junsu, Bao Qingjie, Tang Huiming. Safety helmet wearing detection based on deformable component model[J]. Computer Application Research, 2016, 33 (3):953-956.
[16]Rubaiyat AHM, Toma TT, Kalantari-Khandani M, et al.Automatic detection of helmet uses for construction safety[C]//Pros of International Conference on Web Intelligence Workshops. Piscataway, NJ:IEEE Press,2017:135- 142.
[17] Xu Shoukun, Wang Yaru, Gu Yuwan, Li Ning, Zhuang Lihua, Shi Lin. Research on safety helmet wearing detection based on improved Faster RCNN[J]. Computer Application Research, 2020, 37(03): 901-905.
[18]Wu Dongmei, Wang Hui, Li Jia. Safety helmet detection and identification based on improved Faster RCNN[J]. Information Technology and Information Technology, 2020(01): 17-20.
[19] Xiao Tigang, Cai Lecai, Gao Xiang, Huang Hongbin, Zhang Chaoyang. Improved YOLOv3 detection method for helmet wearing [J/OL]. Computer Engineering and Applications: 1-13 [2021-03-04].
[20] Cao Yan, Song Zhengwei, Zhou Nan. A detection method for construction workers' helmet wearing based on grade SSD[J]. Industrial Control Computer, 2021, 34(01): 55-56.
[21]Qian Huifang, Yi Jianping, Fu Yunhu. Overview of Human Action Recognition Based on Deep Learning [J/OL]. Computer Science and Exploration: 1-20 [2021-03-07].
[22] Fang Kai. Detection method of fence crossing behavior based on deep learning [J]. Computer System Applications, 2021, 30(02): 147-153.
[23]Shi Jiacheng, Chen Zhi, Hu Chen, Wang Renjie, Ye Kehuai. Real-time character action recognition based on long-acting recursive deep convolutional network[J]. Software Guide, 2021, 20(02): 49-53 .
[24]DONAHUE J, HENDRICKSL A, GUADARRAMA S, et al. Long-term recurrent convolutional networks for visual recognition and description[C]. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2015: 2625-2634.
[25] Fan Bank, Zhao Haifeng, Zhang Shaojie. Human action recognition algorithm based on 3D convolution residual network[J]. Application Research of Computers, 2020, 37(S2): 300-301+304.

ISSN: 2664-9640

[26] Zhang Congcong, He Ning, Sun Qixiang, Yin Xiaojie. 3D DenseNet human action recognition method based on attention mechanism [J/OL]. Computer Engineering: 1-9 [2021-03-10].
[27] Zhu Yunpeng, Huang Xi, Huang Jiaxing. Research on human action recognition based on 3D CNN[J]. Modern Electronic Technology, 2020, 43(18): 150-152+156.
[28] Wang Qian, Sun Xiankun, Fan Dongyan. Human action recognition based on deep learning spatiotemporal feature fusion[J]. Sensors and Microsystems, 2020, 39(10): 35-38.
[29]Zhong Qiubo, Zheng Caiming, Pu Songhao. Research on skeleton action recognition and interaction based on temporal and spatial fusion[J]. Journal of Intelligent Systems, 2020, 15(03): 601-608.
[30]Tang Fangli. Application of image recognition technology in power equipment monitoring[J]. Communication Power Technology, 2019, 36(01): 94-95.
[31]Wu W, D asgupta S, Ramirez E E, et al. Classification accuracies of physical activities using smartphone motion sensors[J]. Journal of medical Internet research, 2012, 14(5): e130-e130.
[32] Martin H, Bernardos A M, Iglesias J, et al. Activity logging using lightweight classification techniques in mobile device[J]. Personal and Ubiquitous Computing, 2013, 17(4):675-695.
[33]KwapiszJ R,Weirs CU M,Moore S A.Activity recogmtion using cell phone accelerometers[J].ACM SigKDD Explorations Newsletter, 2011,12(2):74-82.
[34]Zhang L, Wu X, Luo D. Recognizing Human Activities from Raw Accelerometer Data Using Deep Neural Networks[C]// Proc of 2015 IEEE 14th International Conference on Machine Learning and Applications.Miami:IEEE, 2015:865-870 .
[35] Zhang Xiaohua, Ma Yu, Yang Chenhui, Sui Yushan. Intelligent recognition method of equipment installation location based on convolutional neural network [J]. Industrial Instrumentation and Automation, 2021(01): 13-17.
[36] Li Yingda, Jia Ning. Deep learning-based optical ultra-precision manufacturing equipment state pattern recognition[J]. Laser Journal, 2020, 41(12): 161-165.
[37] Lu Wanjie, Fu Hua, Zhao Hongrui. Mine inspection robot equipment recognition based on deep learning algorithm[J]. Journal of Engineering Design, 2019, 26(05): 527-533.
[38] Gao Lu, Ma Yuanjing. Equipment fault detection and recognition based on Faster R-CNN[J]. Computer System Applications, 2019, 28(04): 170-175.
[39]Zhou Ziqiang, Chen Qiang, Ma Bihuan, Qi Donglian. Application of an improved YOLO target detection method in the recognition of abnormal state of cable equipment[J]. Electrical Measurement and Instrumentation, 2020, 57(02): 14-20.
[40] Wu Jianhua, Liang Lihui, Zhang Zhe, Liu Yunpeng, Pei Shaotong. Detection and identification of oil leakage in substation equipment based on deep learning[J]. Guangdong Electric Power, 2020, 33(11): 9-15.