Research on an Intelligent Monitoring Model Training and Monitoring Method Based on Deep Learning

Ke Zeng, Runbo Lu *, Jianyu Li *, Runming Yu *, Hongbiao Lun and Qiwei Liu

No. 1 Substation Management Office of Dongguan Power Supply Bureau of Guangdong Power Grid Co., Ltd, Dongguan, Guangdong, China

* Corresponding author: Runbo Lu, Jianyu Li, Runming Yu

Abstract**:** In the field of security surveillance technology, the false alarm problem has always been the key factor in the performance improvement of the video surveillance system. To solve this problem, this paper proposes a new monitoring model, and studies the model training and monitoring methods, devices, electronic devices and storage media. By collects multiple sets of image information in the target area, generates 3 D images, marks the trust objects and alert areas, and then trains a monitoring model that can accurately identify the category of moving objects. This method can avoid unnecessary alarm information, effectively improve the accuracy and efficiency of the monitoring system, and significantly reduce false positives.

Keywords**:** video monitoring; 3D image; target recognition; trust object; alert area

1. Introduction

Power system is the infrastructure of modern society, and its safe and stable operation is crucial to economic and social activities. The safety of the key facilities in the substation especially needs to be strictly guaranteed. However, the traditional manual monitoring method is not only inefficient, but also easy to the interference of human factors, leading to frequent monitoring blind spots or false positives. In addition, non-target objects such as small animals and fluttering objects in the natural environment often trigger the alarm of the traditional monitoring system, resulting in a large number of invalid alarms, increasing the operation and maintenance costs and work burden.

In order to improve the monitoring efficiency and accuracy of key areas of the power system, reduce false positives, and ensure timely response when really needed, we propose an intelligent monitoring model training and monitoring method based on deep learning [1]. The method is specifically designed for the characteristics and requirements of the power system, and can effectively identify and distinguish the mobile objects in the target area, including potential threats such as unauthorized personnel, vehicles, as well as trusted objects such as maintenance personnel and small animals. By marking trust objects and alert areas in 3 D images and training the monitoring model using these data, this method can significantly reduce the false alarm rate while ensuring the accuracy of monitoring.

The monitoring model training and monitoring methods in this study not only improve the intelligent level of power system monitoring, but also optimize the allocation of human resources and reduce the operation and maintenance cost by reducing false positives. In addition, the implementation of this method can also help to find the potential safety risks in advance, so as to take preventive measures to enhance the safety protection capability of the power system

2. Principle of Monitoring Model Training and Monitoring Methods

This study provides a monitoring model training method for identifying the moving object in the target area and identifying the category of the moving object, and the multiple specified directions of the target area is set with several cameras [2,3], the monitoring model construction method comprising:

Acquiring multiple sets of image information acquired by the camera, each group of image information includes image information for each of the specified directions;

For each set of the image information, the image information is stitched together to generate a three-dimensional image [4];

A trust object is marked in the above three-dimensional image, including at least one of the animals meeting the preset dimensional parameters and the moving object moving within a fixed range;

According to the preset alert center position and alert distance, then the alert area is marked in the 3 D image, and then the training image is obtained;

The monitoring model is trained using the training image and the face information parameters of the designated person, who is allowed to enter the alert area.

In addition, for the monitoring method, this study is mainly used to monitor the moving objects in the target area, and multiple cameras are set in the target area.

The image information is collected through the camera, and the image information is input into the trained monitoring model to identify the moving object in the target area. When a moving object exists, and the category of the moving object is determined, and when the moving object is a trusted object or a designated person, it is judged as a normal environment, and there is no need to trigger the alarm device, otherwise the judgment is an abnormal environment will trigger the alarm device.

3. The Composition of the Monitoring Device and the Monitoring Model Training Device

A monitoring device consists of a camera acquisition module, a moving object determination module, a category determination module, and an environment determination module:

(1) Camera acquisition module for collecting image information through the camera.

(2) Moving object determination module for inputting image information into the trained monitoring model to identify whether there is a moving object in the target area.

(3) Category determination module is used to determine the category of a moving object when it exists.

(4) Environment judgment module is used to determine a normal environment when the mobile object is a trusted object or a designated person, otherwise an abnormal environment.

The monitoring model training device consists of image information acquisition module, 3 D image generation module, trusted object marking module, training image acquisition module and monitoring model training module:

(1) Image information acquisition module for obtaining multiple sets of image information collected by the camera, each group of image information includes image information of each specified orientation.

(2) 3 D image generation module, which is used to assemble the image information for each group of image information.

(3) A trust object marking module for marking a trust object in a 3 D image, the trust object includes at least one of the animals that meet the requirements of the preset size parameters and the floating object moving within a fixed range.

(4) Training image acquisition module for marking the alert area in the 3 D image according to the preset alert center position and the alert distance to obtain the training image.

(5) Monitoring model training module, used to train the monitoring model with the training image and the face information parameters of the designated person. The designated person is the person allowed to enter the alert area

Furthermore, it should include electronic devices that can perform monitoring model training methods and monitoring methods, and computer-readable storage media for storing computer instructions for monitoring model training and monitoring methods.

4. Specific Implementation Mode of Monitoring Model Training and Monitoring Methods

According to the principle flow chart of the monitoring model training method, the specific implementation of the monitoring model training steps is analyzed in detail:

(1) First, obtain multiple sets of image information collected by the camera. Image information is the image information corresponding to the target region. The target area can be the place that needs to be monitored or the area where the items need to be monitored, such as power distribution rooms and transmission towers. For distribution room and transmission tower, are high voltage, when non-professionals near or into the site, there may be personal danger, on the other hand, distribution room and transmission tower is indispensable in the power supply link, and operational cost is higher, if the distribution room and transmission tower is damaged, will greatly affect the power transmission, also improve the power maintenance costs.

There are multiple designated location cameras in the target area, which can be a binocular or multiocular camera. The multiple designated directions of the target area are multi-directional information collection of the target area to obtain the image information of different directions. For example, the camera can collect the left view, right view, top view, front view and rear view of the target area [5]. All the cameras synchronously collect the image information and upload it to the monitoring center. Each set of image information includes the image information of each specified orientation, i. e., a set of image information is the image information collected by all the cameras at the same time. It should be noted that when setting up the camera, you can use tools such as Bluetooth to make the remote wireless connection to the monitoring center, and then the remote operation connection center will adjust the binocular camera to the appropriate position.

The image information may include information about the size, appearance, shape, and action trajectory of the moving object.

(2) For each group of image information, the image information is stitched together to generate 3D images. A point cloud coordinate system can be established according to one of the cameras. For each group of image information, point cloud data can be obtained according to the image information, and the point cloud data of other cameras can be converted into point cloud data under the point cloud coordinate system, and three-dimensional images can be obtained according to the obtained point cloud data.

You can also obtain the initial image according to the image information, set the image splicing line and crop and splicing the initial image, and use the visual sensor to generate the stitched video image. Target area can set a certain number of fixed reference, fixed reference for the image stitching line setting benchmark, can obtain the plane position relationship between the moving object and fixed reference, and can obtain the spatial position relationship between the moving object and fixed reference from three-dimensional three-dimensional space, when the plane position relationship and spatial position relationship mismatch, then judge the moving object in three-dimensional three-dimensional mapping error, need to remapping.

(3) Mark the trust object in the 3 D image. A trust object is a trusted moving object, including at least one of the

animals meeting the preset size parameters and a moving object moving within a fixed range.

For animals that meet the requirements of preset size parameters, the requirements can be less than 20cm, width less than 10cm and height less than 10cm. For example, small birds, small reptiles, etc., namely small animals within the range of preset size parameters, the size of these objects can meet the preset size parameters requirements. Small animals are often considered to have no personal danger or disruptive motive when entering the target area, so animals that can meet the requirements of preset dimensional parameters are considered trusted in order to avoid misdetection. For larger animals, it can be considered that damage is possible. For example, when the cow approaches the transmission tower, it may hit the transmission tower and cause damage.

For the easy to float objects within the fixed range, even if these easy to float objects float, they only move within the fixed range, and the easy to float objects can be ropes, plants and trees, etc. For example, ropes are used in the power distribution room. The door at the door of the power distribution room is long except weeds. When the wind blows, the ropes and weeds will be easy to move, and they will be detected as moving objects. In order to avoid misdetection, it can be considered that the floating objects moving within the fixed range are the objects of trust.

(4) Mark the alert center in the 3 D image according to the preset alert center position. For example, obtain the 3 D coordinates corresponding to the position of the alert center, and then annotate them in the 3 D image as the alert center.

(5) The warning area is obtained with the warning center as the basis point and the warning distance as the regional range parameter. When the number of preset alert distance is multiple, with the alert center and the area range parameters of the alert area, including: with the alert center as the center, the alert distance as the radius of multiple circular area, as the alert area, the alert degree of the alert area is negatively correlated with the alert distance.

In one example, as shown in Figure 1, the number of preset alert distances is two, including the first alert distance and the second alert distance, with the alert center as the center (reduction of the alert center, not shown in Figure 1), the first alert area 1 is generated with the alert center as the center of the radius, and the second alert area 2 is generated with the alert center as the radius. The first alert area 1 is higher than the second alert area 2.

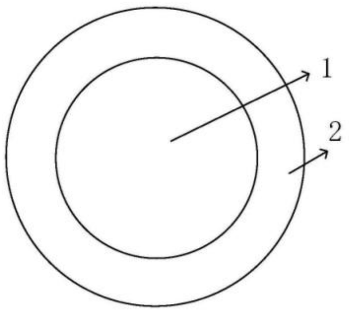

Figure 1. Schematic diagram of the alert area

(6) According to the preset alert center position and alert distance, mark the alert area in the 3 D image, and obtain the training image.

The alert area is the key area within the target area, and the target area includes the alert area. The position of the alert center is generally the location of the important items to be monitored, or the central position of the alert area. The alert distance is the distance of the alert range based on the alert center, and the alert distance is used to limit the size of the alert area.

According to the preset alert center position and alert distance in the 3 D image, the alert area can be marked in the alert image. Specifically, the alert center position can be marked in the 3 D image, and then the alert range can be set with the alert center mark as the base point and the reference alert distance.

(7) The monitoring model is trained by training image and designated face information parameters. Trust object is the object allowed to enter the alert area, training image marked with trust objects, using the training image training monitoring model, in the trained monitoring model to identify the image information of the target area, can quickly identify the trust objects including animals that meet the preset size parameters, movement within a fixed range.

The designated personnel are allowed to enter the alert area, such as the management personnel, staff of the target area, then the designated personnel to enter the alert area is compliant. Training the monitoring model according to the face information parameters of the designated person, then the identity of the designated person can be quickly identified.

Specifically, the face parameter information of the designated person can be the full face information recorded by the designated person, including the left face parameter information, the right face parameter information, the positive face parameter information, etc.

When the trained monitoring model is used to monitor the target area, the alarm device will not be triggered when the object, the animal with preset volume parameters or the designated personnel enter the alert area.

The alert center is marked in the 3 D image through the preset alert center position, the alert area is obtained with the alert center and the alert distance as the area range parameter, and the alert area is marked in the 3 D image, and finally the method of training the image is obtained. The alert area can be accurately marked in the training image to improve the detection accuracy of moving objects in the alert area.

In addition, the detailed implementation of the monitoring method:

(1) Image information was collected via the video camera. Image information is the image information corresponding to the target region. The target area can be the place that needs to be monitored or the area where the items need to be monitored, such as power distribution rooms and transmission towers.

There are multiple designated location cameras in the target area, which can be a binocular or multiocular camera. A plurality of designated directions in the target area is multi-directional information acquisition in the target area to obtain the image information of different directions.

When a moving object enters the target area, binocular (multi-ocular) cameras at different locations will find the moving object, and binocular (multi-ocular) cameras at different locations will obtain the size, appearance, shape and action trajectory of the moving object respectively, take it as image information, and send it to the data processing background respectively.

(2) Input the image information into the trained monitoring model to identify whether there are moving objects in the target area. Monitoring model root the monitoring model training method is trained as described above.

(3) When the moving object exists, determine the category of the moving object.

(4) When the moving object is a trusted object or a designated person, it is judged to be a normal environment, otherwise it is an abnormal environment. Conversely, when the moving object is not the trusted object or the designated person, the abnormal environment. In the specific implementation, when the environment is judged to be abnormal, the alarm device is triggered. Specifically, the sound and light alarm can be triggered, or the alarm information can be generated and sent to the terminal of the staff member to inform the staff member to process it in time. If the moving object leaves the alert area within the preset time, close the speaker of the alarm device, and the preset time can be 20 seconds.

5. Summary and Outlook

This study has successfully developed an innovative monitoring model training and monitoring method, specifically for the key regional security monitoring needs of the power system. By adopting deep learning techniques, our method is able to accurately identify and classify moving objects within the target region, including unauthorized intruder and trusted objects, from multiple sets of image information. The experimental results prove the remarkable effect of this method in improving the monitoring accuracy and reducing the false alarm rate, and provide an effective technical solution for the safety monitoring of the power system. Our study also demonstrates the potential of the proposed monitoring devices and electronic devices in real power systems that can be effectively integrated into existing monitoring systems, providing real-time monitoring and rapid response capabilities.

Although this research has made encouraging achievements in the field of power system monitoring, but there are still many directions worthy of future research to further explore: (1) multiple scene adaptability: this research mainly for power system monitoring, the future work can extend the model to other monitoring scenarios, such as Banks, transportation, public safety, etc., to verify its universality and adaptability. (2) Cost-effectiveness: it is necessary to conduct a detailed cost-effectiveness analysis before the promotion and application. Future studies could evaluate the relationship between system cost and monitoring effectiveness under different configurations to guide actual deployment and optimize resource allocation.

Through these future research directions, we expect to further improve the level of intelligence of the monitoring system, and provide more powerful and reliable technical support for the security monitoring of power systems and other critical infrastructure.

Acknowledgement

The research was funded by: Project No.: 031900KZ24040083; Project name: 3D Visual Intelligent Safety Management Device for Substation.

References

[1] X. Zhou, X. Xu, W. Liang, et al., "Deep-Learning-Enhanced Multitarget Detection for End–Edge–Cloud Surveillance in Smart IoT," in IEEE Internet of Things Journal, vol. 8, no. 16, pp. 12588-12596, 15 Aug.15, 2021, doi: 10.1109/JIOT.2021.3077449.

[2] Rudd S, Kirkwood J, Davidson E, et al., “Intelligent Monitoring of Distribution Automation,” Pac World, 2012.

[3] J. Chen, K. Li, Q. Deng, et al., "Distributed Deep Learning Model for Intelligent Video Surveillance Systems with Edge Computing," in IEEE Transactions on Industrial Informatics, doi: 10.1109/TII.2019.2909473.

[4] Yassine Himeur, Somaya Al-Maadeed, Hamza Kheddar, et al., “Video Surveillance Using Deep Transfer Learning and Deep Domain Adaptation: Towards Better Generalization,” Engineering Applications of Artificial Intelligence, Volume 119,2023,105698, ISSN 0952-1976.

[5] Gupta, I., Seeja, K.R., “Crowd Density Estimation for Video Surveillance Using Deep Learning: A Review,” In: Senjyu, T., So–In, C., Joshi, A. (eds) Smart Trends in Computing and Communications. SmartCom 2024 2024. Lecture Notes in Networks and Systems, vol 948. Springer, Singapore, 2024. https://doi.org/10.1007/978-981-97-1329-5_23.