Research on Detection of Large Coal Blockage at the Transfer Point of Belt Conveyor Based on Improved Mask R-CNN

Xiaoqiang Shao, Hua Zhu*, Defeng Guo, Runyang Zheng, Jinyang Wei
Xi’an University of science and Technology, Xi’an 710000, China

Abstract. Aiming at the problem of coal accumulation and blockage at the transfer point due to the presence of large coal in the coal belt conveyor, a method based on the improved Mask R-CNN for the detection of the jam at the transfer point of the belt conveyor is proposed. The method firstly adds the SENET and SKNet models to the Mask R-CNN feature extraction part ResNet50, enhances the function of extracting features, and cancels the segmentation part of Mask R-CNN. In order to make the network better converge, the loss function in the RPN network is optimized, and the original IOU is replaced with GIOU, which solves the problem that the IOU cannot be optimized when it is 0 and the IOU cannot distinguish the anchor from the ground truth.

1. Introduction
The research on the jam detection of the brittle belt conveyor has always been concerned by the coal industry. Due to the backwardness of technology and the limitation of application scenarios, the traditional jam detection method has not failed to make a major breakthrough. At present, the detection of the jam at the transfer point of the belt conveyor requires two methods: manual detection and sensor detection. Among them, the manual inspection mainly relies on the human eye to observe the labor force of the workers, and cannot meet the coal mine main transportation system required by the smart mine. Unmanned and intelligent requirements. The coal-filled sensors can be roughly divided into three categories: the first type, the coal-fired sensor based on the travel switch. Due to the complex environment of the coal mine, it is often not affected by the external environment such as coal dust and moisture, and it is often unable to timely and accurately alarm. Its durability, sensitivity and reliability. The sex is very unsatisfactory; the second type, based on the mercury switch or the kerosene switch, can not achieve all-round high-precision measurement, and its own anti-interference ability is poor; the third type, based on the electrode-based principle of the coal-fired sensor. It needs to clean up too much coal in the electrode holder regularly, especially after spraying water, the coal dust and water should be wiped clean and the maintenance frequency is high. It is known that the main problems of the above three types of sensors include: First, the sensors are all contact sensors, which are prone to mechanical friction, which can damage the equipment and reduce the detection accuracy. At the same time, the friction is easy to cause electric spark. There is a big hidden danger to the safety of the coal handling system [1]. Second, the anti-interference ability is poor and vulnerable to environmental impact. The coal mine environment is complex, and the working state of the traditional sensor is easily disturbed by the environment, electromagnetic, vibration and the like. The coal-fired sensor of the travel switch is easily affected by the external environment such as coal dust and moisture, and often cannot be timely and accurately alarmed. Its durability, sensitivity and reliability
are not ideal. The electrode type coal-fired sensor accumulates excessive coal or its electrode seat. After spraying water, it may not work properly. Accuracy and false positive rate are also closely related to the dry humidity of coal; the coal-fired sensor of mercury switch or kerosene switch can not achieve all-round high-precision measurement, and its own anti-interference ability is poor; third, the detection effect is poor. Only the judgment result of whether or not to pile coal is output, and there is no intermediate process data, and the cause type or development trend of the coal seam abnormality cannot be analyzed, and the detection output is seriously delayed, and the damage of the coal seam failure to the belt conveyor cannot be effectively reduced. For these two methods, there is a problem that the detection accuracy is not high, the device is damaged, and the durability is poor. This paper proposes a method based on improved Mask RCNN for the detection of the jam at the transfer point of the belt conveyor. In this method, the SENet and SKNet models are added to the feature extraction part ResNet50, and the loss function in the RPN network is optimized. The original IOU is replaced with GIOU to effectively learn the rich features and make it easier to deal with large foreign objects. Identification. Finally, it is indirectly judged whether the transfer point is stuck by determining whether the position of the coal block changes, and the change of the position of the module can be judged between several frames to determine the distance between the frames of the target detection. If the distance is greater than a certain threshold, the coal mass is not clogged, and when the distance between the coal blocks is less than a certain threshold, clogging occurs. Finally, the purpose of improving the detection accuracy is achieved.

2. Mask R-CNN Large Coal Detection Algorithm

At present, the application of deep learning algorithm has become a major trend in the field of target detection. The main idea of this detection method is to use convolution operation to extract the deep features of large coal in the picture, and to identify and determine the extracted features.

Out target location. Convolutional neural network (CNN) has achieved good results in image classification and target detection. In the process of CNN's continuous development and growth, traditional target detection methods are gradually replaced.

The target detection algorithms based on deep learning can be divided into three categories according to their different principles: The first type of target detection algorithm is based on the search target detection algorithm. This type of method mainly uses deep learning to judge the tendency of visual attention, according to the judgment result of visual attention. Keeping close to the position where the real target appears, and then identifying the target of the corresponding position, but the search times are too many when searching for the target, which reduces the recognition efficiency of the algorithm, and the detection effect of the algorithm is not good when there are too many similar targets in the image. The second type is the target recognition algorithm based on regression. The representative algorithms of this method are YOLO, SSD, etc. The candidate frame extraction process is not shown in this algorithm, and belongs to the one-stage target detection algorithm. In the target recognition process, the target candidate area is canceled, and a whole neural network is used to directly input an entire image into the neural network, and the target is directly regressed, and then the coordinates of the bounding box and the object's category and confidence are predicted to complete the target detection. This saves the time required for identification, and the recognition speed is greatly improved compared to the R-CNN series, which can reach up to 155 frames per second. However, YOLO divides the picture into 7*7 grids, which seems a bit rough. In the regression stage, the error will be deviated due to this error, resulting in inaccurate positioning of the object, low detection accuracy, and poor detection of small objects. Each grid of YOLO can only recognize one object. When dealing with scenes with relatively dense objects such as bird migration and dense crowds, multiple objects appear in one grid, and the detection effect will be poor [2]. The third category is the target recognition algorithm based on regional recommendations. The main representative algorithms of this method are R-CNN, Fast R-CNN, Faster R-CNN and Mask R-CNN. The CNN series is a very good target recognition algorithm. It has a high accuracy when identifying targets, and is undoubtedly a good choice for scenes with low real-time requirements.
Large coal is the main cause of the jam at the transfer point. It is necessary to accurately identify the large coal in the transfer point. There is a certain requirement for the accuracy of the algorithm. The Mask R-CNN network has been developed more maturely in target recognition. The accuracy rate has good performance, and the real-time requirements are not high. Therefore, this paper uses Mask R-CNN network to identify large coal.

2.1. Mask R-CNN network structure and optimization

The Mask R-CNN network is the best paper for ICCV 2017, highlighting the latest achievements in machine learning in computer vision in 2017. In the latest development of machine learning in 2017, the single-task network structure has gradually become less noticeable, replaced by an integrated, complex, multi-tasking multitasking network model. Mask R-CNN can achieve both target detection and instance segmentation in a network. The network structure can be extended based on the Faster R-CNN network.

Mask R-CNN follows the idea of Faster R-CNN. The feature extraction part adopts the architecture of ResNet-FPN, and adds a Mask prediction branch. The Mask R-CNN network can be divided into four parts: ResNet-FPN, RPN, ROI Align, bounding box identification (classification and regression) + mask prediction. The Mask R-CNN network structure is shown in Figure 1.

Figure 1. Mask R-CNN network structure diagram.

Mask R-CNN has a network for extracting features. The network for extracting features is used in the Mask R-CNN network with resnet50. There are five large feature structures in the resnet50, c1 c2 c3 c4 c5, which are the five structures. The structure is multi-layer convolution, which represents the 5 layers, and then other SENE, SKNET modules are added behind these layers to enhance the extraction feature.

For the original network, there will be problems of inaccurate resolution and missed detection. The SENET and SKNET modules are added to the network, and the role is to obtain stronger features [3]. Given an input x, the number of characteristic channels is c_1, and a feature of the feature channel number c_2 is obtained by a general transformation such as convolution. Each two-dimensional feature channel is transformed into a real number, which has a global receptive field to some extent, and the output dimension matches the number of feature channels input. It characterizes the global distribution of responses on the feature channels and allows global layers of receptive fields to be obtained close to the input layer. Then, the weights are generated for each feature channel by the parameter w. These weights are regarded as the importance of each feature channel after feature selection, and then weighted to the previous feature by multiplication by channel, and the pair in the channel dimension is completed. Recalibration of the original features. SKNET is a lightweight embedded module, which mainly enables the network to adaptively adjust the domain size according to multiple scales of input information. This network is mainly divided into three operations: Split, Fuse and Select. Here, Split refers to a complete convolution operation of different convolution kernel sizes on the input vector. This part of Fuse is roughly the same as the processing of the SE module. It should be noted that the two matrices a and b of the output, where matrix b is a redundant matrix. The Select operation corresponds to the Scale in the SE module. The difference is that Select uses two weight matrices, a and b, for weighting, and then sums to get the final output vector. It can be seen that
SENET is directly added to the fully connected layer after the complete convolution operation (1×1 convolution + 3×3 convolution + 1×1 convolution), learning the inter-channel dependencies, and then weighting the learned channel weights. Return to the original vector. SKNET replaces the 3×3 convolution part of ResNeT, and replaces it with two or more convolution operations of different convolution kernel sizes plus learning channel weights. The output vector continues with the 1×1 convolution operation. ResNet50 improved network structure shown in Figure 2.

![Figure 2. Improved ResNet50 network structure diagram.](image)

2.2. The core process of improved Mask R-CNN bulk coal detection

Mask-RCNN algorithm steps:

1) According to the characteristics of the detection target, the training data sources are prepared. The sources of these data sources mainly include experiments on the belt transport system built by the laboratory and the relevant experimental collection of the main inclined shaft of the Dongxia Mine and the Shimen mine head.

2) Pre-process the samples that need to be trained, then use the LabelMe annotation tool to mark the data and make the Mask mask.

3) Input it into an improved pre-training model neural network (ResNet50 [4], currently only the ResNet50 version of the Keras version) to obtain the feature map of the training sample image, and set a predetermined ROI for each point in the feature map (region of Interest), thereby obtaining a plurality of candidate ROIs.

4) Next, these candidate ROIs are sent to the region proposal network (RPN) for binary classification (foreground or background) and BB (bounding box) regression, and some candidate ROIs are filtered out, and the remaining The ROI performs the ROIAlign operation.

5) Finally, these ROIs are classified (N classification), BB regression, and MASK generation.

6) Repeat steps 4) to 5) and train all samples to get the final optimized adjustment test model.

3. Experimental results and analysis

In this experiment, in order to better explain the generalization and detection accuracy of the experiment, the test data is taken in the training set before the model training to randomly extract the
image captured by the main inclined shaft of the Dongxia Mine, the Shimen mine head and the laboratory. The number of verification sets is one-fifth of the training set, and the verification set data selects large pieces of coal with different backgrounds and different coverage levels. After testing, the Faster R-CNN detection algorithm has more missed detections under severe conditions of large coal cover, and there are certain deviations in positioning. The overall detection accuracy is not high. The Mask-RCNN detection method in this paper is better than the Faster R-CNN test for the detection of large coals in the case of severe coal cover. It can not only obtain the positioning frame for detecting the building target, but also obtain the contour of the large coal. The value of Mask provides the possibility to further improve the edge of the large coal seam in the future, and the overall detection accuracy is higher.

4. Conclusion

In this paper, the combination of large coal detection and deep learning is used. Based on the Mask-RCNN target detection method, the images from the main inclined shaft of the Dongxia Mine, the Shimen mine head and the laboratory are selected as data sources, under the Keras deep learning framework. The ResNET50 network part is improved and then combined with the Mask-RCNN algorithm for multi-thread training model. Finally, a large coal detection model with training optimized weight parameters is obtained. This method effectively avoids the defects of traditional target detection, does not need to manually set the target feature value, and the coal mine underground image is easy to obtain low cost, and can acquire the characteristics of the same large coal target at different azimuth angles, and overcomes the problem from different angles. The problem of inaccurate detection improves the ability to learn models. The experiment proves the reliability of the algorithm and framework used in this paper. The experimental results achieve the expected model effect and improve the detection accuracy of large coal. Compared with the traditional large coal detection method, it is more rapid, intelligent and automated. However, the detection effect needs to be improved in the case of severe coverage of large coal.

References

[1] Cui Yayi. Intelligent Control of Coal Mine Underground Belt Conveyor [J]. Energy and Energy Conservation, 2016(02): 165-166.
[2] Li Dajun, He Weilong, Guo Bingxuan, Li Maosen, Chen Minqiang. A Building Target Detection Algorithm Based on Mask-RCNN [J]. Science and Mapping Science, 2019, 44(10): 172-180.
[3] Liu Xueping, Li Yiqian, Liu Li, Wang Zhe, Liu Yu. Improved YOLOV3 Target Recognition Algorithm Embedded in SENet Structure [J/OL]. Computer Engineering: 1-6[2019-11-20]. https://doi.org/ 10.19678/j.issn.1000-3428.0052861.
[4] Wang Heng, Li Xia, Liu Xiaofang, Xu Wenlong. Study on Pathological Image Classification of Breast Cancer Based on ResNet50 Network [J]. Journal of China Jiliang University, 2019, 30(01): 72-77.