

Article

Real-Time Conveyor Belt Deviation Detection Algorithm Based on Multi-Scale Feature Fusion Network

Chan Zeng ^{1,2}, Junfeng Zheng ^{1,2} and Jiangyun Li ^{1,2,*}

¹ School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China; chanzengupup@163.com (C.Z.); g20188721@xs.ustb.edu.cn (J.Z.)

² Key Laboratory of Knowledge Automation for Industrial Processes, Ministry of Education, Beijing 100083, China

* Correspondence: leeji@ustb.edu.cn

Received: 5 August 2019; Accepted: 24 September 2019; Published: 26 September 2019



Abstract: The conveyor belt is an indispensable piece of conveying equipment for a mine whose deviation caused by roller sticky material and uneven load distribution is the most common failure during operation. In this paper, a real-time conveyor belt detection algorithm based on a multi-scale feature fusion network is proposed, which mainly includes two parts: the feature extraction module and the deviation detection module. The feature extraction module uses a multi-scale feature fusion network structure to fuse low-level features with rich position and detail information and high-level features with stronger semantic information to improve network detection performance. Depthwise separable convolutions are used to achieve real-time detection. The deviation detection module identifies and monitors the deviation fault by calculating the offset of conveyor belt. In particular, a new weighted loss function is designed to optimize the network and to improve the detection effect of the conveyor belt edge. In order to evaluate the effectiveness of the proposed method, the Canny algorithm, FCNs, UNet and Deeplab v3 networks are selected for comparison. The experimental results show that the proposed algorithm achieves 78.92% in terms of pixel accuracy (PA), and reaches 13.4 FPS (Frames per Second) with the error of less than 3.2 mm, which outperforms the other four algorithms.

Keywords: conveyor belt; deviation detection; multi-scale feature fusion

1. Introduction

Mining operations require conveyor belts to move mined material, such as coal, from the working face over a long distance to the processing plant [1]. As the main transportation equipment in the coal industry and other industries [2], conveyor belts will deviate during operation because of uneven stress on the surface, which results in material spilling, property damage and even personnel injury. It has long been recognized that a belt condition monitoring system for early detection of unusual belt deviation is desirable. Thus, it is essential to study the mechanism of conveyor belt deviation [3], respond to the intelligent mine by focusing on monitoring and controlling [4] and combine it with the new generation of information technology such as cloud computing, big data and artificial intelligence to propose a more intelligent method for conveyor belt deviation detection.

An increasing number of researchers have been paying attention to the detection of conveyor belt deviation, and it is a significant problem in coal mining [5–8]. Initial detection of conveyor belt deviation mainly relies on manual inspection, which is labor intensive and prone to errors and omissions. Then, mechanical and photoelectric detection devices were introduced [9]. Mechanical detection is used to drive the linkage mechanism through the contact between the roller and belt,

while photoelectric detection mainly relies on wireless sensor modules to achieve this. Huang et al. used the joint sensor and the hall sensor to monitor the deviation of the conveyor belt [10]. When the joint of a conveyor belt with a magnetic disc passes through the joint sensor for length measuring, the hall sensor would sense a series of pulses. Then, when the conveyor belt deviates beyond a certain range, the hall sensor cannot detect the pulse signal in order to determine whether the belt deviates. However, the mechanical device is only suitable for conveyor belts with low rotational speed and is seldom used in practice. While photoelectric detection equipment mainly relies on a wireless sensor module, its detection accuracy depends too much on the installation density of the sensor. In a harsh and complex environment, the sensor is easily damaged, and it is difficult to ensure the reliability of the detection. With the development of computer technology, machine vision has gradually become an active research direction of conveyor belt deviation monitoring. Li et al. proposed a detection approach based on visual computing that was comprised of three steps, namely image filtering and edge enhancement, edge detection using the Canny algorithm [11] and the Hough transform [12] according to the linear characteristics of the conveyor belt edge [13]. Yang et al. introduced a gray-scale average method for the fast segmentation of conveyor and background [14]. By mapping the binary image to a one-dimensional fault edge function, a deviation eigenvector consisting of a deviation angle and offset was designed for deviation detection. Gao et al. presented a multi-view image adaptive enhancement method for conveyor belt fault detection and developed a mine conveyor belt online detection system to realize real-time detection of belt deviation, surface damage and longitudinal tear [15]. Traditional computer vision methods are sensitive to noise, so they can detect other edges such as a coal mine and material transported on the conveyor belt. In the place where the conveyor belt occluded, there will be discontinuity in the detection results. In addition, the parameters of the algorithm need to adjust many times for different scenarios.

In recent years, deep learning has made remarkable achievements in computer vision. Deep learning is broadly used in image recognition, object detection and classification [16–18]. It has achieved state-of-the-art results in those fields. In this work, a novel algorithm based on a multi-scale feature fusion network is presented to achieve real-time conveyor belt deviation detection. With the input of the video stream image captured by the monitoring system in order to the embedded equipment for processing, it can realize the remote monitoring of conveyor belt operation with improved accuracy and avoidance of interference from harsh and complex environments. The algorithm, which is based on a multi-scale feature fusion network proposed in this paper, has the following characteristics:

- (1) Multi-scale fusion extracts global and local edge features effectively in a network structure;
- (2) The traditional convolutions are replaced with depthwise separable convolutions for network compression, which improves detection speed and performance to meet the requirements of various production scenarios;
- (3) A new weighted loss function is presented to optimize the network, which has a greater improvement in accuracy;
- (4) The experimental results show that the proposed method can achieve 78.92% in terms of pixel accuracy and the processing speed is 13.4 FPS with the error of less than 3.2 mm.

2. Materials and Methods

In this section, an intelligent algorithm based on a multi-scale feature fusion network of conveyor belt deviation surveillance is presented. Figure 1 shows the overall architecture of the proposed algorithm, including the feature extraction module and deviation detection module.

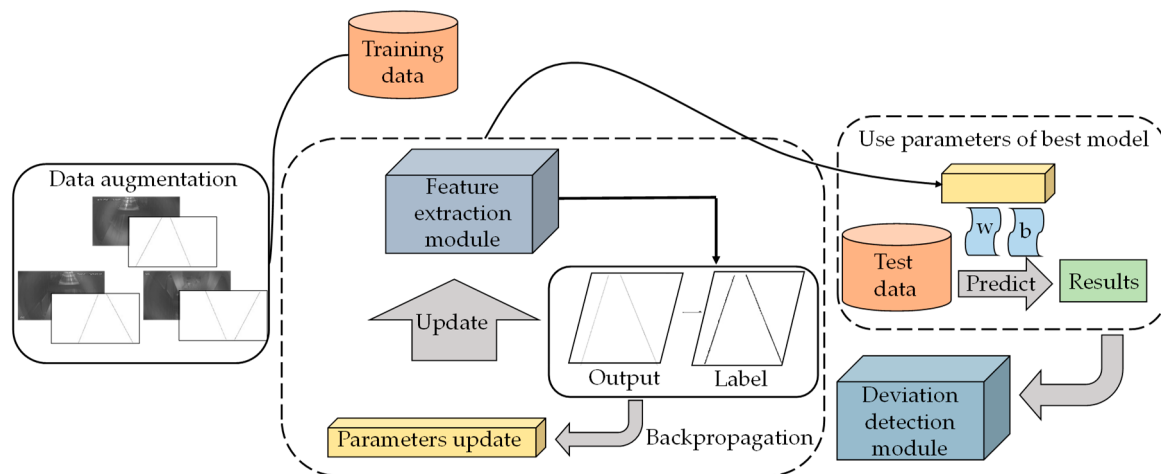


Figure 1. The overall architecture of the proposed algorithm. According to the specific task required of belt deviation detection, this paper develops an algorithm based on a multi-scale feature fusion network comprising of a feature extraction module and a deviation detection module. Then, training data are fed into the feature extraction module and parameters are updated with backpropagation. In addition, parameters of the best model are used to predict results on test data. Finally, the deviation detection module identifies and monitors the deviation fault.

To enrich the training data, the following data augmentation techniques are adopted: (i) flip horizontal; (ii) flip vertical. In the training process, the training data are processed by the feature extraction module for image edge feature extraction, and parameters are updated with backpropagation. In the test process, parameters of the best model obtained by the cross validation method are used to predict the conveyor belt edge image on test data. Finally, the deviation detection module judges the deviation by calculating the conveyor offset.

2.1. Multi-Scale Feature Fusion

Conveyor belt images are a special image data and differ hugely from natural images. In this paper, conveyor edge detection is transformed into pixel-level classification in image segmentation [19,20]. FCNs (Fully Convolutional Networks) [21], a classical full convolution network, replaced the last three full connection layers of VGG (Visual Geometry Group) [22] with convolution layers, which reduced parameters and kept the spatial information of the image. There are some advantages in FCNs that are very suitable for processing conveyor images, so this paper takes the FCNs as the basic architecture to study the deviation detection task of conveyor images.

The resolution of the feature image is reduced because of several convolutions and poolings. FCNs addresses this by adding links that combine the final prediction layer with lower layers with finer strides. Then, the stride 32, 16 or 8 predictions are upsampled back to the image, which corresponds to FCN-32s, FCN-16s and FCN-8s. Considering that FCN-8s, which gets finer precision, only fuses the output of the last three stages of pooling by per-pixel adding, it still loses much detailed edge information for conveyor belt images. In this paper, a multi-scale feature fusion network, which builds on the FCNs, is proposed for global and local feature extractions. As shown in Figure 2, the output of each stage after pooling is upsampled by 2, 4, 8 and 16 times the original image. In addition, by changing the fusion strategy of adding pixels-by-pixels in FCNs, the channel fusion method is used to fuse the low-level features with the rich location and detail information and the high-level features with stronger semantic information. Therefore, the network can learn abundant multi-scale features and solve the problem of detail recovery in edge detection.

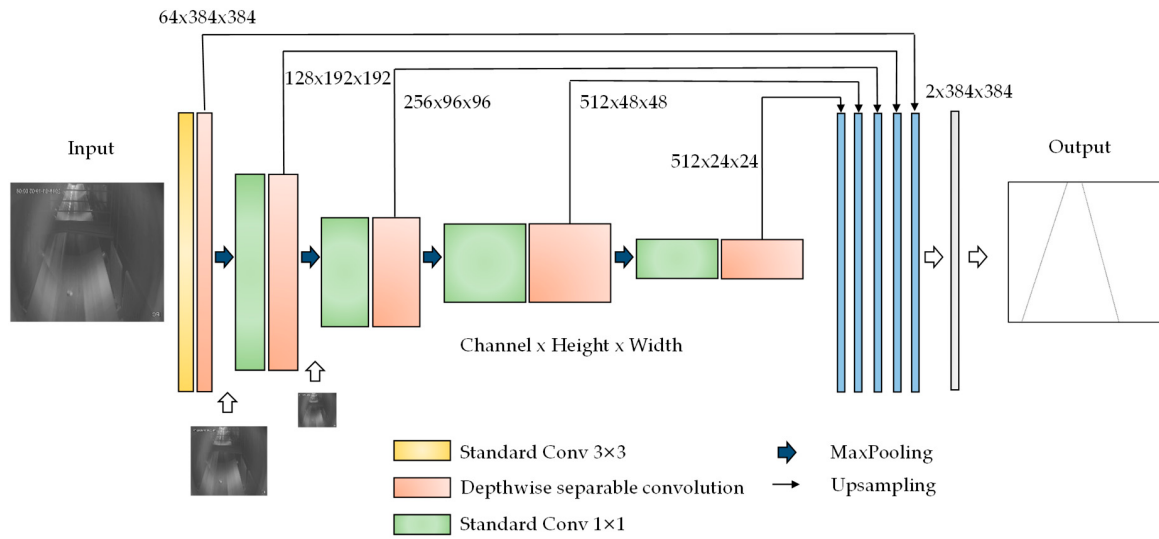


Figure 2. Illustration of our multi-scale feature fusion network for conveyor deviation detection. The resized original images are imposed at the first two stages for strengthening information flow. The side outputs of the multi-scale feature fusion network are upsampled to the original image and then fused to get the final prediction.

2.2. Network Compression

Depthwise separable convolution is an effective solution for model speeding up, which has been explored for efficient network design, e.g., Xception [23] and MobileNet [24,25]. The depthwise separable strategy changes the standard convolution with two layers, which could reduce the number of parameters and network complexity without hurting the learning capacity. The structure of depthwise separable convolution is shown in Figure 3. Depthwise separable convolution comprises of two layers: depthwise convolutions and pointwise convolutions. The $k \times k$ depthwise convolution is used to apply a single filter per each input channel (input depth is M). Pointwise convolution, a simple 1×1 convolution, is then used to create a linear combination of the output of the depthwise layer.

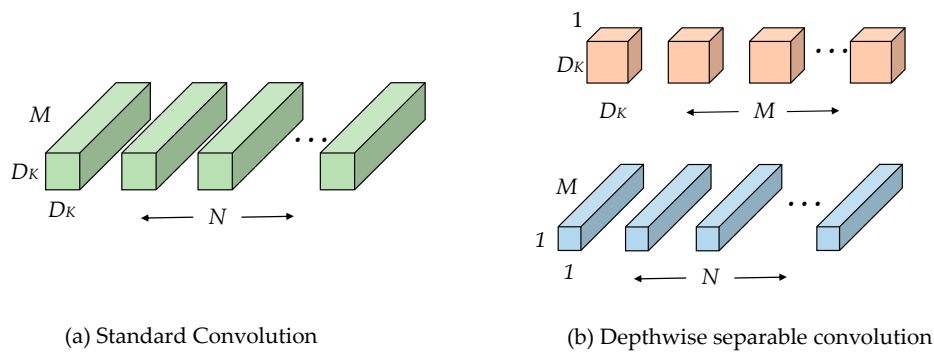


Figure 3. (a) The standard convolution. (b) Depthwise separable convolution comprises of two layers: depthwise convolutions and pointwise convolutions.

In this paper, 3×3 depthwise separable convolution is used, and the computational cost is $8 \times$ less than the standard convolutions. To ensure the detection accuracy of the original network while using depthwise separable convolution, the number of channels are expanded by 1×1 convolution in each stage, so that the depthwise separable convolution can capture useful features in a higher spatial dimension. In addition, inspired by ESPNet [26,27], the original image at the first two stages is used for strengthening information flow. The detailed structure and parameter are displayed in in Figure 4. Conv means the standard convolution, Dwconv represents the depthwise separable convolution

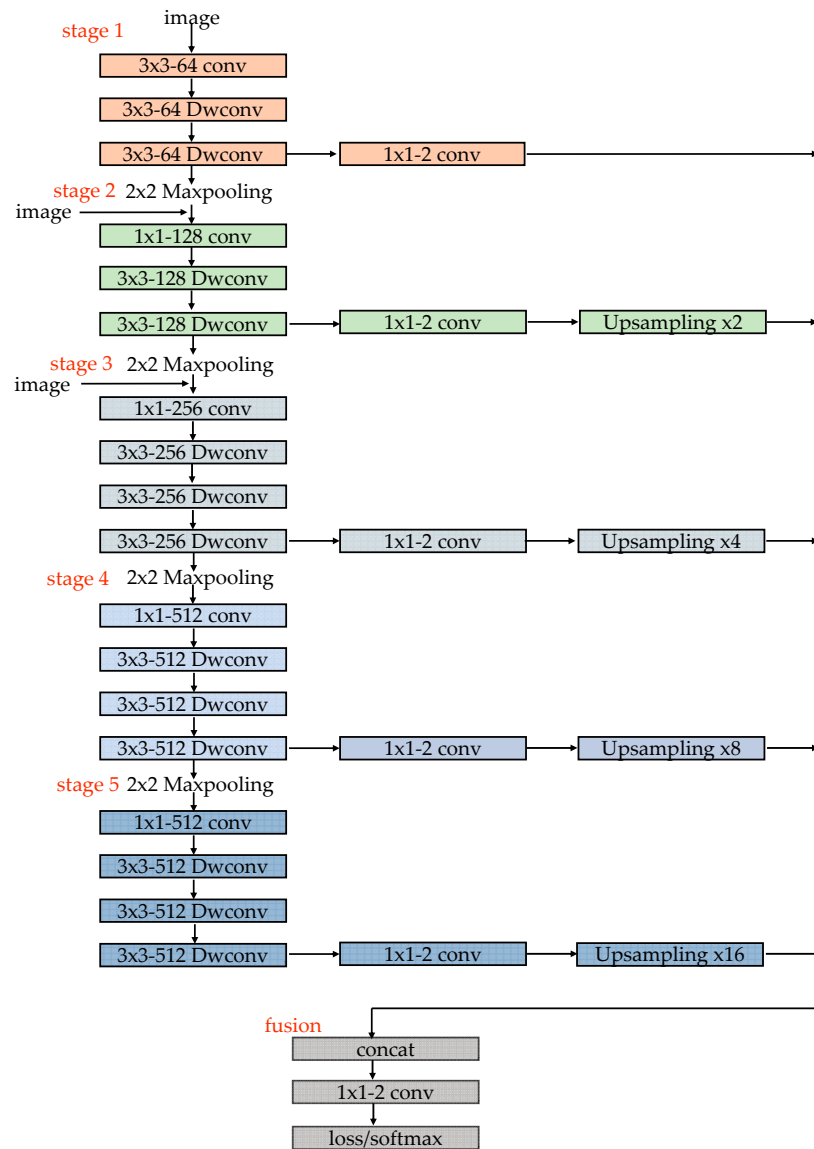


Figure 4. Specific structure and parameters of our multi-scale feature fusion network. Its convolution layers are divided into five stages, in which a maxpooling layer is connected after each stage. Feature maps of each stage are connected to a conv layer with kernel size 1×1 and channel depth 2 and then upsampled to original image. Finally, a weighted cross-entropy loss/softmax layer is followed to get the output.

2.3. Weighted Loss Function

In the conveyor image, the number of edge and non-edge pixels are seriously unbalanced and about 90% are non-edge pixels. However, the traditional loss function does not balance the edge pixels and non-edge pixels reasonably. To solve this problem, this paper improves the traditional cross-entropy loss function and uses weighted cross-entropy loss to increase the weight of the edge pixels. The weighted cross-entropy loss can be expressed as:

$$\text{loss} = -w \times g \times \log(\sigma(y)) - (1 - g) \times \log(1 - \sigma(y)). \quad (1)$$

In Equation (1), w represents the weight coefficient of the edge pixels in the image and g presents the label for the category of the pixels (edge pixels or non-edge pixels). y is the prediction result and σ denotes the *sigmoid* function. By increasing the weight of the edge pixels, the network needs to pay

more attention to the edge part in order to achieve more accurate edge detection results. In this paper, the different w are used to conduct experiments and the results are in Section 4.

2.4. Deviation Detection Module

When the conveyor belt deviates, the deviation of the center line position of the two edges of the conveyor in the result and the label is used as the fault feature information. Assuming (x_{ia}, y_{ia}) is the coordinate of the starting point of the left edge in the result, while (x_{ib}, y_{ib}) is the coordinate of the starting point of the right edge, and the abscissa of midpoint between two edge x_A and x_B can be expressed as:

$$x_A = \frac{\sum_{i=1}^j x_{ia}}{j}, \quad (2)$$

$$x_B = \frac{\sum_{i=j+1}^n x_{ib}}{n-j}, \quad (3)$$

where j represents the pixel width of left edge and $n - j$ represents the pixel width of right edge, then the abscissa of the center line of the left and right starting point are as follows:

$$d_1 = \frac{|x_B + x_A|}{2}. \quad (4)$$

The criteria for judging the deviation of conveyor is defined as:

$$|d_1 - d_2| > \alpha w. \quad (5)$$

In Equation (5), d_2 is the abscissa of the center line of the two edges of conveyor belt in normal operation, w represents the width of belt and α is constant. When the belt deviates, the deviation in the width direction will exceed 5% of the bandwidth, so the value of α is set to 0.1. By comparing $|d_1 - d_2|$ and $0.1w$, it can be judged whether the conveyor belt is running a fault.

3. Experiments

In this section, the experimental contents are discussed, and the proposed method with the traditional edge detection method and other deep learning methods are compared.

3.1. Datasets

Aiming for the specific task of conveyor belt detection, it is necessary to extract the edge features of the belt in the image, while ignoring other edge features, such as beams, coal on the conveyor belt and so on. However, the commonly used dataset for edge detection is global edge detection. Therefore, a specific dataset of belt edge is made for the task in this paper. The data that come from the conveyor video are about one hour long and are collected by the camera under the mine; then the OpenCV is used to extract 500 valid images of coal, no-coal and coal being transported, with an average size of 1000×560 pixels.

The larger the training dataset, the stronger the generalization of the training model; therefore, the training data are increased by flipping 1500 images vertically and horizontally. In each image, the edge of the belt is traced by a white line with one-pixel width, and the rest is regarded as background filled with black so it can make the label. Then the label is binarized to $\{0, 1\}$ to facilitate the training and testing of subsequent networks. We randomly divide the dataset in to 1200 training images and 300 test images according to the ratio of 4:1. In this paper, the five-fold cross-validation is adopted on training images and tested on the rest of the 300 images. Figure 5 shows the original image and the corresponding label.

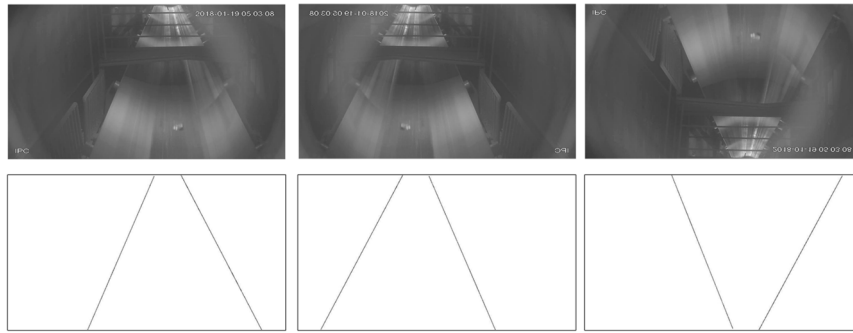


Figure 5. Three examples from the dataset. Three conveyor images (the first line) and their ground truths (the second line).

3.2. Training Details

The Pytorch framework is used to implement this multi-scale feature fusion network, and the model is trained on Nvidia GeForce 1080Ti GPU with a batch of 4 for 1000 epochs. The Adam optimizer is used with an initial learning rate of 0.001. The BatchNormLayer decay factor is set to 0.9 for the exponential moving average. The loss function is weighted with the cross-entropy loss described above, and the weight coefficient of the positive sample is set to 3 through several experiments. After about 3 h of training with a five-fold Cross-validation, the best model is saved and evaluated on 300 test images.

3.3. Evaluations

Formally, the evaluation metrics for the proposed strategy are pixel accuracy (PA) and testing accuracy. In this paper, PA represents the ratio of the number of edge pixels predicted correctly to the total number of edge pixels in the label, while the testing accuracy is the product of camera resolution and average pixel width. Lower testing accuracy represents a finer predicted edge. The camera resolution is measured as follows:

When the camera is at an acute angle to the ground, as the model shown in Figure 6, θ is the field of view of the camera, h is the camera installation height, β is the angle between the camera and the ground and L is the maximum distance that the camera can detect. The mathematical relationship between them is as follows:

$$L = \frac{2h \tan \theta}{\sin \beta}. \quad (6)$$

Assuming that the parameter of the CCD camera is vxw , the resolution is expressed as:

$$pre = \frac{L}{\sqrt{v^2 + w^2}} = \frac{2h \tan \theta}{\sin \beta \sqrt{v^2 + w^2}}. \quad (7)$$

Through calculation, the resolution of the camera used in this paper is 0.8 mm.

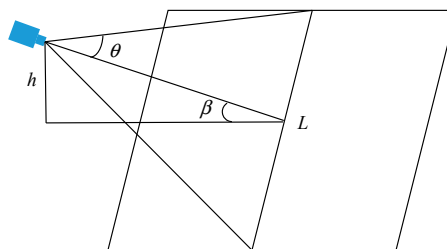


Figure 6. The structure and parameters when the camera is at an acute angle to the plane.

3.4. Comparisons With Canny Algorithm

Traditional edge detection is based on gray image, which detects the discontinuous gray scale of the image. Where the gray value of a pixel transforms is the edge of the gray image. The Canny algorithm and its variants are classical edge detection algorithms in the traditional machine learning field. We carry out the experiments with the Canny algorithm and compare them with the results of the proposed method. As shown in Figure 7a, all the edges in the image are detected with the Canny algorithm. The covered part of the conveyor belt cannot be automatically repaired. However, it can be clearly seen from Figure 7b, and the detection result of the proposed method has almost no breakpoint and avoids interference from other objects on the conveyor belt.

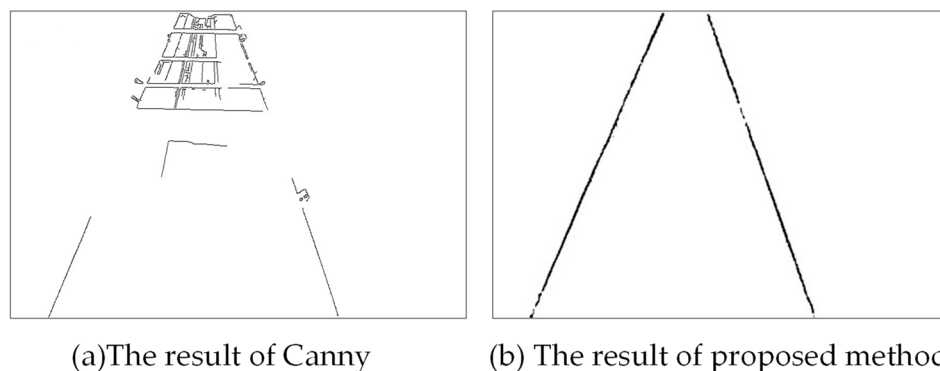


Figure 7. Comparison of the results between the Canny and the proposed method. (a) The result of Canny. (b) The result of the proposed method.

3.5. Comparisons with State-of-the-Art Methods

In this section, experiments are conducted on the conveyor dataset mentioned above, and our method is compared with the state-of-the-art approaches. Among them, UNet [28] is an improved FCN network, and Deeplab v3 is a very successful version of the image segmentation network Deeplab series [29].

UNet is the cornerstone of medical image segmentation. In recent years, many effective methods in medical image field have been improved on UNet. It consists of two parts: a feature extraction part that increases the receptive field to obtain context information, and an upsampling fusion structure that can achieve precision positioning. We use the dataset to train the UNet network for 1000 epochs, and we save the best model to conduct experiments on the test images. Then we post-process the result of UNet by image erosion. The result images are shown in Figure 7, column 5.

Besides, compared with UNet, the Deeplab series network is more popular in image segmentation tasks. Deeplab v3 proposes a more general framework that replicates the last block of ResNet and cascades it. An ASPP (Atrous Spatial Pyramid Pooling) structure with a BN (Batch Normalization) layer is used to increase the receptive field and extract more abundant features without increasing the parameters. We train the Deeplab v3 with the same training parameter settings and then test it. For the result of Deeplab v3, we use the least square method for linear fitting. The results are presented in Figure 8, column 6.

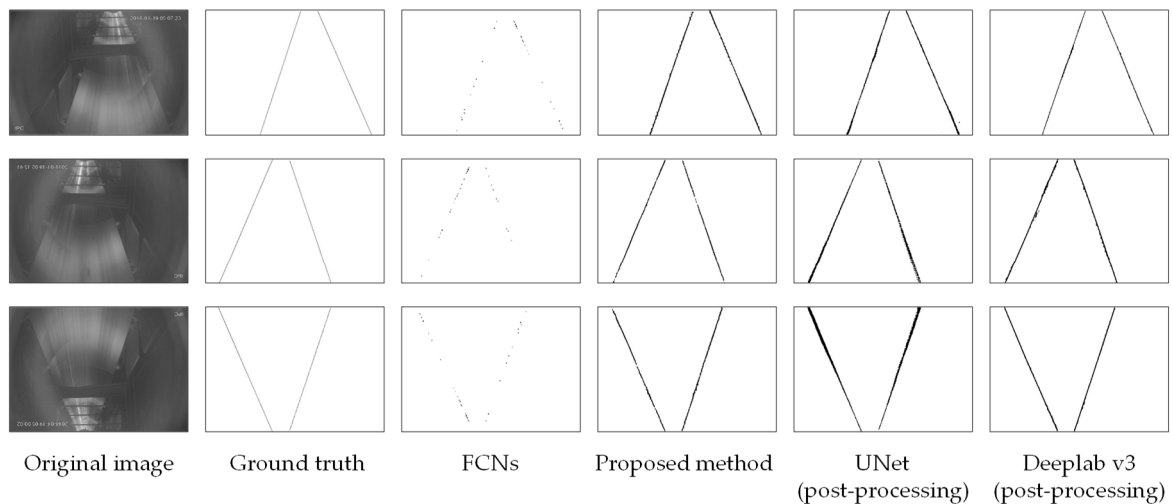


Figure 8. Comparison of experimental results in three methods. From the third to the sixth column: FCNs, proposed deep network, UNet and Deeplab v3 with post-processing. The proposed method shows superior performance, particularly with its ability to correctly segment the conveyor edge with almost no breakpoint as compared to FCNs and Deeplab v3. Besides, the proposed work gets finer results when compared to UNet.

3.6. Impact of the Weighted Loss Function

A series of experiments are carried out to verify the scientificity of the weighted loss function. The experiment results are shown in Figure 9. The left represents the influence of the weight coefficient (w) on pixel accuracy, while the right shows a relationship between w and the testing accuracy. When setting w to 3, it can achieve the best pixel accuracy and testing accuracy.

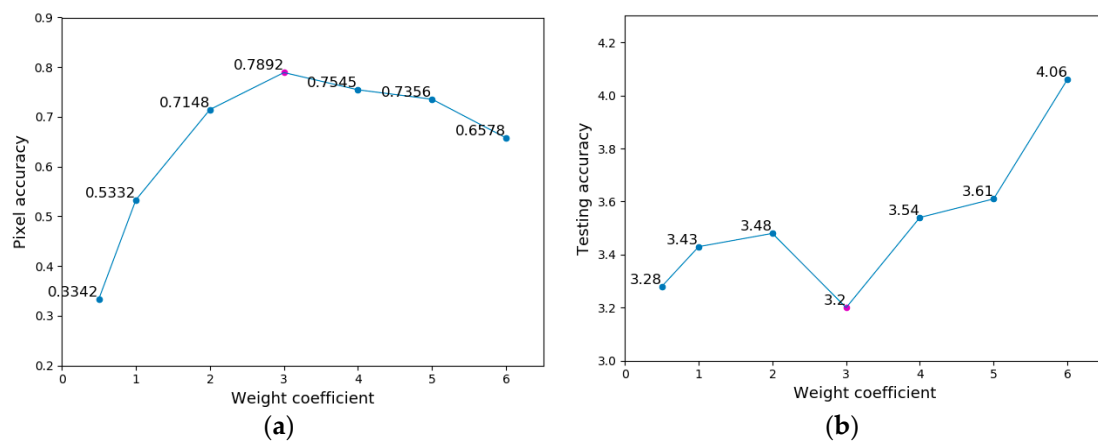


Figure 9. The performance on the test dataset with different weight coefficients. (a,b) represent the influence of weight coefficients on pixel accuracy and testing accuracy, respectively.

4. Discussion

To evaluate the effectiveness of this method, the detection results from subjective and objective aspects are analyzed and discussed.

4.1. Subjective Evaluation Results

As seen from Figure 8, it is obvious that the conveyor edge detected by FCN-8s is a series of points, which can roughly describe the conveyor edge. However, there are many breakpoints and discontinuity, and the detection edge is blurred, which cannot meet the requirements of the task

of conveyor belt edge detection. Besides, the original UNet result is rough with one end thick and the other thin, and the problem still exists after image erosion, which is that we cannot achieve the detection effect. In addition, the structure of Deeplab v3 is complex with a long training time, while there are many breakpoints and rough edges in the result before post-processing. After linear fitting, the breakpoint problem is solved, but the rough edges still exist. We can see intuitively from the figure that in our method result, the conveyor edge is finer, the breakpoints are fewer and the boundary is clearer, which can reflect the conveyor belt edge well.

Although UNet is helpful for precise localization in image segmentation after many times of up-sampling by bilinear interpolation and successive symmetric fusion, it cannot fully learn edge features because the positive and negative samples are imbalanced in the edge detection task in this paper, and the high-dimensional features are eliminated in the process of upward fusion layer by layer, which leads to rough detection results. Deeplab v3 is generally used in complex image segmentation tasks with a pre-training model based on the ResNet network and various scales of atrous convolution getting different sizes of field. The fusion of the multi-scale atrous convolution and the output after global average pooling is helpful for solving the problem of missing image edge details. However, for the simple edge detection dataset in this task, Deeplab v3 is too complex, and atrous convolution cannot effectively extract features for the conveyor belt occlusion part. As a result, no good experimental results can be obtained.

4.2. Objective Evaluation Results

In this section, the results of FCN-8s, the proposed method, UNet and Deeplab v3 in pixel accuracy, testing accuracy, testing speed and training time are compared. As can be seen from Table 1, our method achieves the best results in terms of pixel accuracy and testing accuracy. FCN-8s cannot measure the testing accuracy for the serious breakpoint, while the result of UNet is not fine. Deeplab v3 has a longer training time due to the more complex network. Considering the inference time, the proposed method significantly outperforms the others by reducing the parameters of the depthwise separable convolution significantly, while the testing speed of FCN-8s, UNet and Deeplab v3 are too slow to meet the requirements of real-time detection in an industrial field. In a few words, our approach has the following advantages compared with other methods:

- (1) The proposed method has a high pixel accuracy without any post-processing operations.
- (2) The simple training and fast testing of the proposed method are advantageous to practical industrial application.
- (3) When there are objects on the belt, our method can automatically repair the occluded with avoidance of interference.

Table 1. Comparison of performance indicators.

Method	Pixel Accuracy /%	Testing Accuracy /mm	Testing Speed /FPS	Training Time /h
FCN-8s	34.58	-	2.2	2.43
Proposed method	78.92	3.2	13.4	3.07
UNet	69.51	7	2.1	3.52
Deeplab v3	71.44	4.6	1.3	8.74

5. Conclusions

In this work, a real-time conveyor detection method based on the multi-scale feature fusion network is developed, which can realize remote on-line monitoring of conveyor belt operation and avoid the interference of a harsh and complex environment. The standard convolutions are replaced with the depthwise separable convolutions, which have 8x less computational cost. Meanwhile, a new weighted loss function is designed to optimize the network, which improves the pixel accuracy from 53.32% to 78.92%. Compared with the other three representative networks, the effectiveness of the

proposed method is verified, while the processing speed of 13.4 FPS can be achieved, and the real-time performance can meet the requirements of various production scenarios. The final test error is less than 3.2 mm, which meets the application requirements of the industrial field. In the future, we will continue to improve our algorithms for superior performance.

Author Contributions: J.L. provided the data and conceived the idea; C.Z. designed the algorithm, carried out the experiments and wrote the paper; J.Z. provided good ideas and revised it critically for important intellectual contents. All the authors have approved the final version to be published.

Funding: This work was supported by the Natural Science Foundation of Beijing Municipality (No. 4182038), the National Natural Science Foundation of China (No. 61671054), the Fundamental Research Funds for the China Central Universities of USTB (FRF-DF-19-002), and the Open Project Program of the National Laboratory of Pattern Recognition (No. 201800027).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Huang, J.; Tian, C.; Xing, L.; Bian, Z.; Miao, X. Green and Sustainable Mining: Underground Coal Mine Fully Mechanized Solid Dense Stowing-Mining Method. *Sustainability* **2017**, *9*, 1418. [[CrossRef](#)]
- Ma, H.-W.; Fan, H.-W.; Mao, Q.-H.; Zhang, X.-H.; Xing, W. Noise Reduction of Steel Cord Conveyor Belt Defect Electromagnetic Signal by Combined Use of Improved Wavelet and EMD. *Algorithms* **2016**, *9*, 62. [[CrossRef](#)]
- Wang, X.J.; Ding, Y.F.; Liu, H.B.; Chi, Y.L.; Xu, P. Study on Running Deviation and Adjusting Deviation of Belt Conveyor's Structure Based on Synchronous Technology. *Adv. Mater. Res.* **2013**, *634*, 3794–3797. [[CrossRef](#)]
- Li, W.; Dai, K. Design of coal mine intelligent monitoring system based on ZigBee wireless sensor network. In Proceedings of the International Conference on Mechanics, Materials and Structural Engineering, Jeju Island, Korea, 18–20 March 2016.
- Yang, Y.; Miao, C.; Li, X.; Mei, X. On-line Conveyor Belts Inspection Based on Machine Vision. *Opt. Int. J. Light Electron Opt.* **2014**, *125*, 5803–5807. [[CrossRef](#)]
- Liu, T.; Tan, C.; Wang, Z.; Xu, J.; Man, Y.; Wang, T. Horizontal Bending Angle Optimization Method for Scraper Conveyor Based on Improved Bat Algorithm. *Algorithms* **2019**, *12*, 84. [[CrossRef](#)]
- Minaee, S.; Wang, Y. Screen Content Image Segmentation Using Robust Regression and Sparse Decomposition. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2016**, *6*, 573–584. [[CrossRef](#)]
- Minaee, S.; Wang, Y. An ADMM Approach to Masked Signal Decomposition Using Subspace Representation. *IEEE Trans. Image Process.* **2017**, *28*, 3192–3204. [[CrossRef](#)] [[PubMed](#)]
- Pang, Y. A novel embedded conductive detection system for intelligent conveyor belt monitoring. In Proceedings of the IEEE International Conference on Service Operations and Logistics, and Informatics, Shanghai, China, 21–23 June 2006; pp. 803–808.
- Huang, M.; Pan, H.S.; Hu, C. A System for Real-time Monitoring and Protecting of Steel Cord Belt Conveyors. *J. China Univ. Min. Technol.* **2006**, *35*, 673–678. (In Chinese)
- Canny, J. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *8*, 679–698. [[CrossRef](#)] [[PubMed](#)]
- Illingworth, J.; Kittler, J. A Survey of the Hough Transform. In *Computer Vision, Graphics, and Image Processing*; Academic Press Professional, Inc.: San Diego, CA, USA, 1988; Volume 44, pp. 87–116.
- Li, Z.L.; Zhao, W.B. Deviation Monitoring of Coal Belt Conveyor Based on Visual Computing. *Saf. Coal Mines* **2014**, *45*, 118–121. (In Chinese)
- Yang, Y.L.; Miao, C.Y.; Kang, K. Machine Vision Inspection Technique for Conveyor Belt Deviation. *J. North Univ. China (Nat. Sci. Ed.)* **2012**, *33*, 667–671. (In Chinese)
- Gao, R.; Miao, C.Y.; Miao, D. Multi-View Image Adaptive Enhancement Method for Conveyor Belt Fault Detection. *J. China Coal Soc.* **2017**, *42*, 594–602. (In Chinese)
- Xie, S.; Tu, Z. Holistically-Nested Edge Detection. *Int. J. Comput. Vis.* **2015**, *125*, 3–18. [[CrossRef](#)]
- Włodarczyk-Sielicka, M.; Polap, D.; Sielicka, W. Automatic Classification Using Machine Learning for Non-Conventional Vessels on Inland Waters. *Sensors* **2019**, *19*, 3051. [[CrossRef](#)] [[PubMed](#)]
- Polap, D. Human-Machine Interaction in Intelligent Technologies Using the Augmented Reality. *Inf. Technol. Control.* **2018**, *47*, 691–703. [[CrossRef](#)]

19. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)] [[PubMed](#)]
20. Hamaguchi, R.; Fujita, A.; Nemoto, K.; Imaizumi, T.; Hikosaka, S. Effective use of dilated convolutions for segmenting small object instances in remote sensing imagery. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, Lake Tahoe, NV, USA, 12–15 March 2018.
21. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
22. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:14091556.
23. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807.
24. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:17004861.
25. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 4510–4520.
26. Mehta, S.; Rastegari, M.; Caspi, A.; Shapiro, L.; Hajishirzi, H. ESPNet: Efficient spatial pyramid of dilated convolutions for semantic segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 561–580.
27. Mehta, S.; Rastegari, M.; Shapiro, L.; Hajishirzi, H. Espnetv2: A light-weight, power efficient, and general purpose convolutional neural network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–21 June 2019; pp. 9190–9200.
28. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany; Volume 9351, pp. 234–241.
29. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:170605587.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).