

Retraction

Retracted: GCN-Unet: A Computer Vision Method with Application to Industrial Granularity Segmentation

Mobile Information Systems

Received 5 December 2023; Accepted 5 December 2023; Published 6 December 2023

Copyright © 2023 Mobile Information Systems. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] S. Pang, X. Xiao, Y. Cui et al., “GCN-Unet: A Computer Vision Method with Application to Industrial Granularity Segmentation,” *Mobile Information Systems*, vol. 2022, Article ID 1466261, 9 pages, 2022.

Research Article

GCN-Unet: A Computer Vision Method with Application to Industrial Granularity Segmentation

Shuyang Pang,¹ Xuewen Xiao,² Yuao Cui,³ Shangwei Mao,¹ Xin Cao,¹ Hongsheng Jia,¹ Hao Wang,³ Fenghua Tong,³ and Xiaohui Zhang ¹

¹CISDI Information Technology Co.,Ltd., 1 N Cisi Street, Chongqing 401122, China

²CISDI Engineering Co.,Ltd., 1 N Cisi Street, Chongqing 401122, China

³Chongqing University-University of Cincinnati Joint Co-Operative Institution, Chongqing University, 174 N Shapingba Street, Chongqing 400044, China

Correspondence should be addressed to Xiaohui Zhang; xiaohui.zhang@cisdi.com.cn

Received 9 June 2022; Revised 20 July 2022; Accepted 2 August 2022; Published 25 August 2022

Academic Editor: Yang Gao

Copyright © 2022 Shuyang Pang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the field of metallurgical industry, identifying the granularity of raw materials is an essential process during transportation. We propose an image segmentation method by the GCN (global convolutional network)-Unet to extract the contour edge of raw materials granules. To obtain legible images of raw materials, a stationary industrial high speed camera is used to photograph the operating belt conveyor from above. Then, a well-trained GCN-Unet model is used to compute the images and output the results with the contour edge of granules and tiny parts of the materials segmented. We combined the U-Net with several global convolutional network models and boundary refinement blocks and compared the prediction results of the GCN-Unet and the U-Net, showing that the GCN-Unet has a better prediction ability with fewer parameters (7,876,675, while the U-Net has 31,101,448 parameters) and a higher calculating speed (about twice faster than the U-Net). Based on the CNN (convolutional neural network), our computer version method can almost replace traditional manual sampling inspection method for the corresponding overall analysis and the automatically identification process.

1. Introduction

Raw material transportation is an essential link in the field of metallurgical industry. Generally, tubular belt conveyors or belt corridors are used to transport raw materials. With the closed transformation of material transportation, the grain size of materials in the closed corridor is difficult to be detected. Due to the long distance of the corridor, the traditional manual inspection method is inefficient and there are personnel safety problems, which has a great impact on the actual production. Traditional manual inspection mainly operates through the mode of manual collection, manual analysis, and manual written transmission of information. An operator takes a few kilograms of raw materials from several tons as a sample, extracts granular materials by sieve, and manually records the granularity. The most striking drawback of this method is that the sample, which is less

than 5% of the amount of whole materials, is too insufficient to be representative. One reason to explain the difficulty of replacing manpower by automation equipment is that the existing various sensors are not able to recognize granularity directly. The other is that the high speed of the conveyor, which is around 2 m/s, causes some difficulties for measuring methods.

The material profile can be identified by the image segmentation algorithm, so as to obtain granularity. In order to obtain clear sample images of all raw materials, set up high-definition cameras supplemented by high-brightness relighting lights, take pictures of fast-running materials, detect material particles transported on the current belt in real-time, and use deep learning techniques to split material particles in pictures. A granularity can be well presented through extracting the materials granules and its contour edge.

Using computer vision technology, intelligent industrial system based on convolutional neural network algorithm can solve the problem of granularity recognition. To realize the automation in mature modern applications, the research goals of industrial AI are to develop AI algorithms and AI systems in order to get better recognition and prediction result [1].

In general, the contributions of the paper are as follows:

- (1) A novel segmentation method is proposed to effectively discriminate the granularity. The adopted GCN-Unet structure greatly enhances the inference efficiency and improves the accuracy of classification and localization in semantic segmentation on the basis of GCN method.
- (2) The combined structure of U-Net and GCN is of great innovative and practical importance and is successfully applied in industrial scene, greatly increasing the real-time inference efficiency.

2. Related Work

In recent years, visual perception technology has been successfully applied in industrial production. The authors of [2] proposed the offline-online NARX neural networks to predict the water content of sintering and control the water content in sintering process. An identification method using an automated visual helmet is analysed in [3] based on artificial neural networks (ANN) [3]. Industry AI enables production to adapt to complex and changing industrial environments and perform diverse industrial tasks, thereby improving productivity, product quality, and equipment performance [4].

U-Net is an outstanding image segmentation method evolving from FCN (Fully Convolutional Networks). U-Net is constructed with a symmetrical structure, with its left part as an encoder and its right part as a decoder. Every dimension of it can contain more features by patching channels, which can also retain more information for locating the target object. In [5], H-DenseUNet is suggested, which is modeled after the auto-context algorithm for segmenting liver and tumors for effectively extracting intra-slice features for hierarchically aggregating volumetric contexts. All above applications prove the efficiency and accuracy of U-Net on fine image segmentation.

The inspiration of dataset annotation method and network designing comes from the medical field. Networks such as CNL-UNet [6] and MC-Net [7] built based on the U-Net [8] are widely used and perform well in semantic segmentation for different objects. Considering the similarity of medical cell images in DIC-HeLa dataset [8] and granular raw materials images, our method of data set annotation also refers to medical images as well, as Figure 1 shows. (b) mainly segment cells by recognizing the monolithic regions, while our methods in (d) focus on the edge of each granularity. By this method, model learns to ignore irregular gaps between granularity that are not clinging to each other and only depict gaps as lines, thus the segmentation speed could be much higher and more applicable in the industrial scene.

There are three essential parts of our approach: Firstly, we labeled the edge of granular raw materials and the tiny parts of them with white color. We annotated 90 original images and, after diverse augmentation operation, we obtained 1710 images and their labels of the same amount. 2% of them are in the validation set and others are used to train the network. Secondly, we propose the construction of the GCN-Unet by replacing the convolutional layers with several global convolutional network modules and adding an amount of boundary refinement blocks. Additionally, we design a suitable loss function for the situation that the amount of foreground and background pixels is unbalanced and that can avoid the network prediction to be heavily biased towards the background, yielding the GCN-Unet to be trained successfully.

3. Methods

3.1. Network Architecture. The network architecture and the details of its components are illustrated in Figure 2(a). In the input layer, a single-channel image of size 288×288 enters. After several steps, including GCN models, boundary refinement blocks, convolutions using 2×2 and 1×1 kernels, and max-pooling layers, the segmentation results of the images are output. In the output layer, Sigmoid is used to do activation and generate the segmentation results.

Similar to the U-Net [8], the whole network has a pooling-down path and an up-sampling path on the left and right sides, respectively. The former path conducts contraction while the latter one does extension. That is, after each step of the contraction path, the number of feature channels is doubled, and after each step of the extension path, the number of feature channels is halved. A GCN model followed by a BR block and a 2×2 max-pooling down-sampling operation with stride 2 form a contraction step. The GCN module is proposed to improve the accuracy of classification and localization in semantic segmentation, while the BR is introduced to improve the positioning of object boundary. The expansion steps include the activation function based on the corrected linear unit (up-sampling the feature map, and then 2×2 convolution), cascade the corresponding feature mapping, GCN model, and BR block. In the last layer, the 1×1 convolution reflects the number of feature vector classes in 16 components.

Classification tasks require the model to keep the transformation of input (such as ResNet [9] and VGGNet [10]) unchanged, so it is contradictory to verify the excellent performance of GCN model in classification and localization. The classification task uses dense connection layers (such as full connection layer [11]), and the positioning task needs to transform sensitive models (such as FCN [12]). At the same time, deconvolution [12] and depooling are introduced to obtain high-resolution feature maps. In order to achieve classification and location at the same time, the GCN model adopts the fully convolution model structure, retains the spatial information in the original input image, and allows the prediction of each pixel. Finally, each pixel is classified on the feature map [13] of up-sampling. In order to

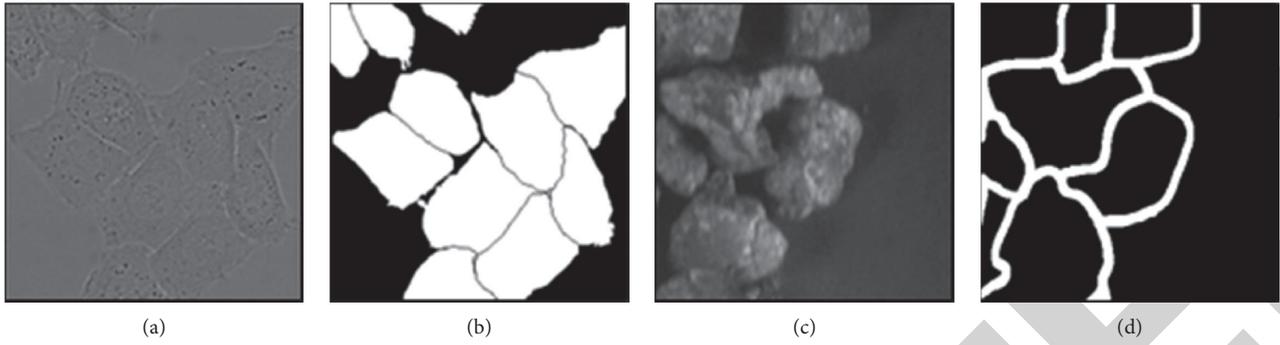


FIGURE 1: Data comparison of HeLa cells image (a) and label (b) and granular raw materials image (c) and label (d).

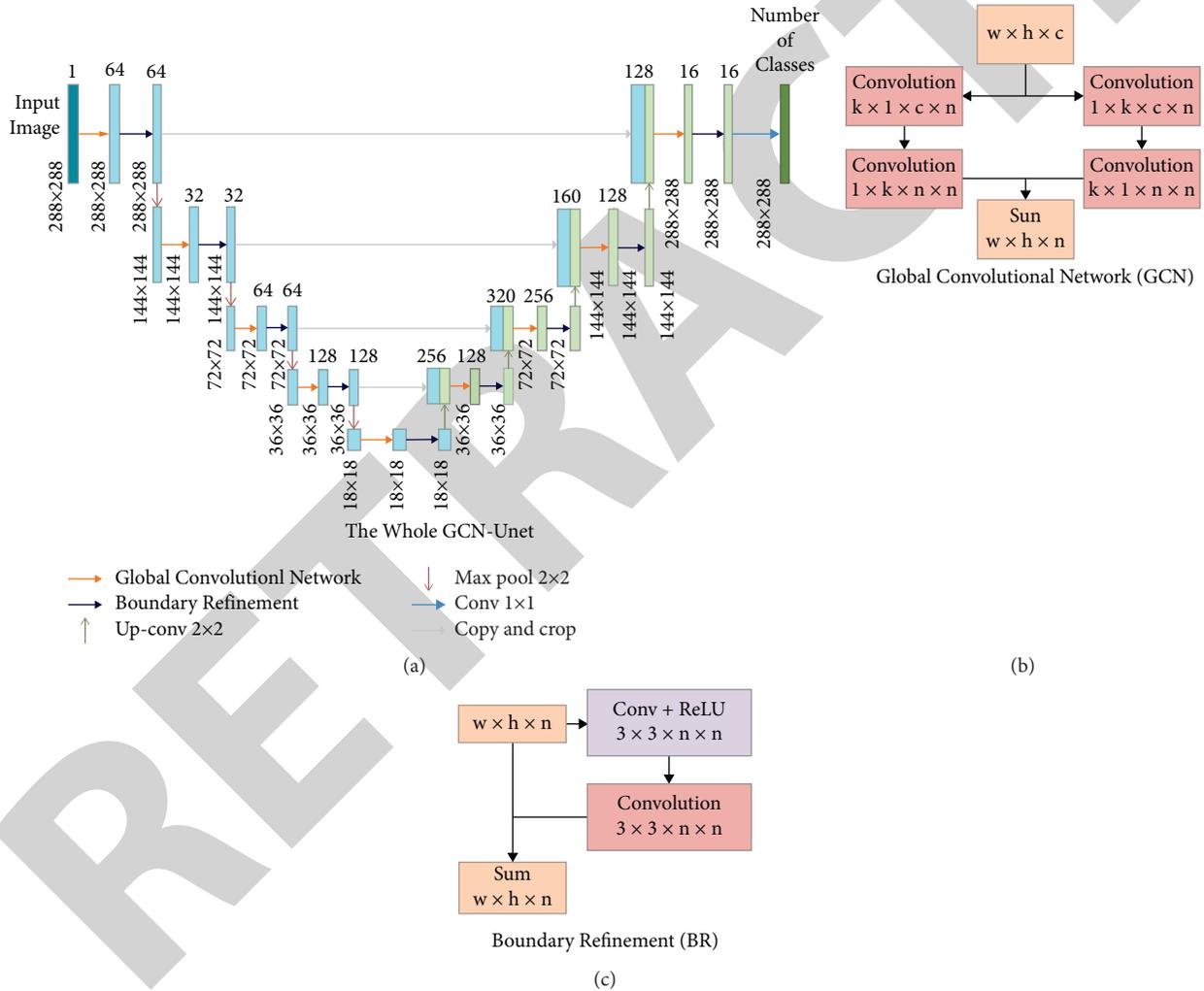


FIGURE 2: An overview of GCN-UNet architecture (example for 288×288 pixels) (A). The arrows signify different operations. The details of global convolutional network (GCN) and boundary refinement (BR) block are shown in (B) and (C), respectively.

discard location information, an incomplete connection or global pooling layer is used. In addition, in order to make the feature mapping have a dense connection with each pixel classifier, the network structure uses a large convolution kernel size. This method also enhances the ability to handle different transformations [13].

A GCN model combines $1 \times k + k \times 1$ convolution and $k \times 1 + 1 \times k$ convolution, so that the dense connection region in the feature map is close to [13]. Each expected class obtains the corresponding multi-scale semantic score mapping by this method. Similar to [12], a deconvolution layer is used to sample up lower resolution score maps,

which is then combined to higher ones. The combined maps are up-sampled for the last time and the final semantic map is generated to output the final prediction results.

BR block can achieve further improvement of the localization ability by modeling the boundary alignment, which refines the predictions near the object boundaries as a residual structure [13]. The details are shown in Figure 2(c).

3.2. Weighted Loss Function. The weighted loss function consists of the dice loss function [14] and a binary cross-entropy (BCE) loss function. After conducting several experiments, the coefficients of the two components are decided to be 1 and 0.2. The weighted loss function is written as follows:

$$\text{Total}_{\text{loss}} = bce_{\text{loss}} + 0.2 \text{ dice}_{\text{loss}}. \quad (1)$$

BCE loss function is widely used in binary classification tasks [15]. These are tasks that answer a question with only two choices. Several independent such questions can be answered at the same time, as in binary image segmentation. The BCE loss here is based on sigmoid to do binary classification, as [14] shows, where N is the number of samples, as this loss is equal to the average of the categorical cross-entropy loss on the two-category task.

$$bce_{\text{loss}} = -\frac{1}{N} \sum_{i=1}^N y_i \ln(\bar{y}_i) + (1 - y_i) \ln(1 - \bar{y}_i). \quad (2)$$

The loss function of the dice coefficient appeared frequently with outstanding performance in granularity segmentation networks, such as Fully Convolutional Networks [16]. Dice loss was first proposed in the article VNet [14]. Dice loss is beneficial for image segmentation involving an extreme imbalance of the positive and negative samples [17]. For granular raw materials data, the contour edge of granules as well as tiny materials of interest occupies only a small area in the scan. This often leads to the loss or partially detection of the foreground and the network prediction is heavily biased towards the background. To solve the problem, the loss function based on the dice coefficient is used to reweight the sample and enhance the importance of the foreground area, making it higher than the background area. The dice coefficient D is written as follows:

$$D = \frac{2 \sum_{i=1}^N p_i g_i}{\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2}. \quad (3)$$

where N is the total number of pixels, p_i is a single component of the predicted binary segmentation area P , and g_i is that of the ground truth binary area G . This formulation of dice can be differentiated, yielding the gradient computed with respect to the j -th pixel of the prediction [14].

$$\frac{\partial D}{\partial p_j} = 2 \left[\frac{g_j (\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2) - 2 p_j (\sum_{i=1}^N p_i g_i)}{(\sum_{i=1}^N p_i^2 + \sum_{i=1}^N g_i^2)^2} \right]. \quad (4)$$

4. Experiment

In the experiment part, several original scenes of industrial raw materials transported by belt conveyor were obtained. Dozens of the images were selected and annotated as the original images of the dataset. In the data preparation stage, the original label after annotation is normalized and adjusted to generate the standard label. After data enlargement and segmentation of the standard label and original large image, the training set is generated. With the same data set, U-Net and GCN-UNet were trained to segment the granular raw materials, and the accuracy and speed of them are compared.

4.1. Dataset. The dataset used in this project are single-channel images of ore, coke, pellet, and sinter scenes. The size of these images is 2048×2448 . Besides belt convey as the background, one type contains only particles. The other contains particles and extremely small materials. The purpose of the algorithm is to recognize the contour of the particles. The particle contour and tiny materials are labeled with white (pixel value 255). The background and the particle area color are labeled with black (pixel value 0) as shown in Figure 3. The size of these single-channel labels is 2048×2448 .

4.2. Image Augmentation. Image augmentation increases scene diversity. As for the image, the most commonly used methods to increase the training sample include affine transformation, such as rotation, distortion, and mirror transformation. Several types of data augmentation are applied: three flip operations towards a different direction, two histogram equalization operations, contrast enhancement, mean blur, change of light intensity, and elastic deformation as Figure 4 shows.

The purpose of image augmentation is to avoid overfitting and increase the generalization ability of the model. Here, another important reason to use image augmentation is to balance the uniqueness of contour shape and increase training data set. Based on the purpose, we have test the best training set number to be around 1500–2000 for U-Net to learn the contour pattern of granularity, so we augment the volume of training dataset 89 to 1691. With different angles, brightness, contrast, and granules distortion, the dataset is more extensive and universal than the original one.

4.3. Data Cutting. One original image has around 5 million pixels, which makes it unfriendly for computation speed. The original image should not be scaled because details would be lost. Therefore, cutting the original image into small ones is a good way to fasten the computation speed without losing the details. As shown in Figures 5 and 6, images and labels of size 2048×2448 are cut into $72 \times 288 \times 288$ partial images. Totally 6480 partial images are obtained for 90 original images.

4.4. Training Details. As Figure 7 shows, the accuracy of the GCN-UNet in validation set starts from 86.43% and reaches 92.77% after training 20 epochs while the U-Net starts from

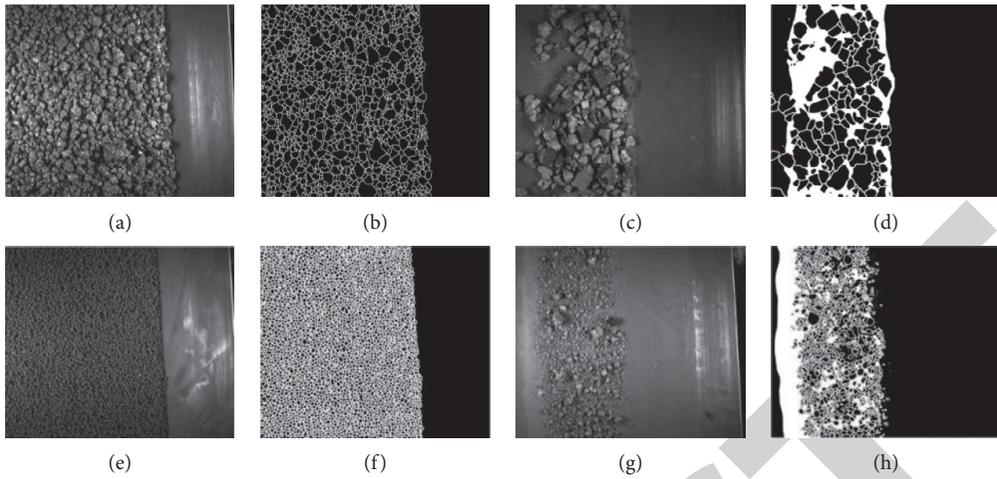


FIGURE 3: Images and corresponding labels of granular raw materials. Use two colors to annotate images with particle and background and three colors for those with particle, background, and tiny materials. (a) Image of sinter. (b) Label of sinter (c) Image of coke. (d) Label of coke. (e) Image of pellet. (f) Label of pellet. (g) Image of ore. (h) Label of ore.

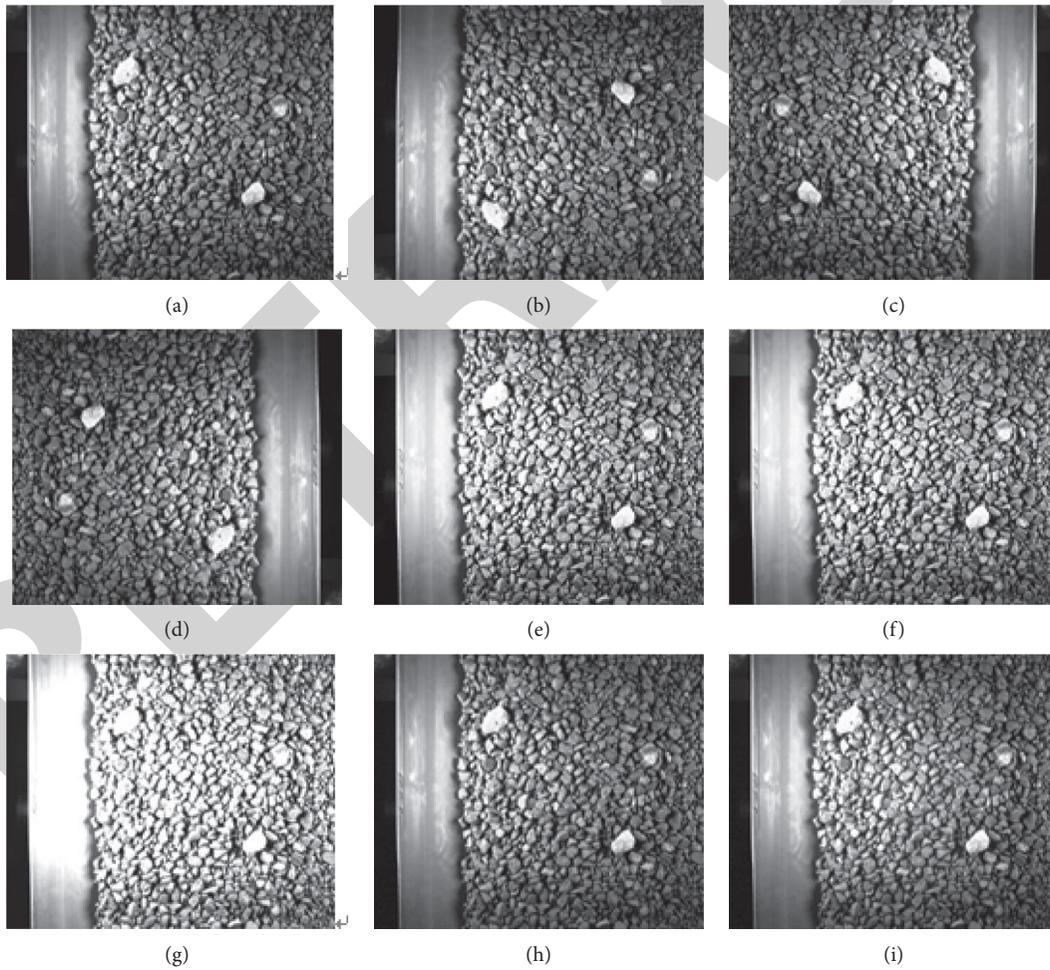


FIGURE 4: Continued.

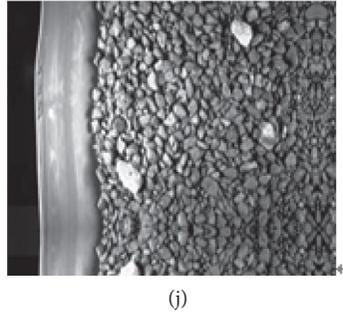


FIGURE 4: Data augmentations. (a) Original image, (b) vertical rotation, (c) horizontal flip, (d) diagonal rotation, (e) histogram equalization, (f) adaptive histogram, equalization, (g) contract enhancement, (h) fuzzy, (i) light enhancement, and (j) elastic deformation.

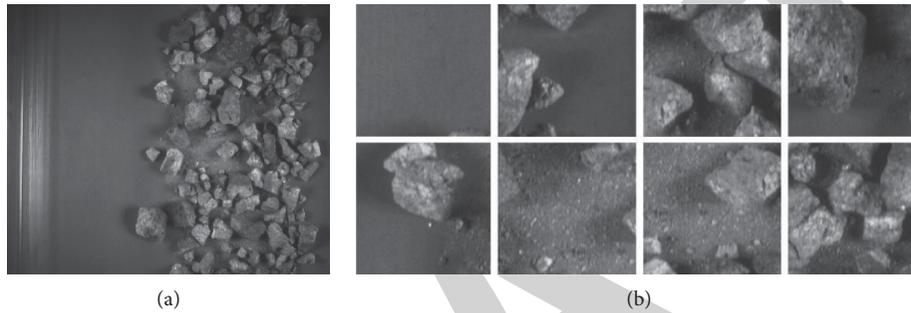


FIGURE 5: Image cutting before training. (a) Original image and (b) cutten images.

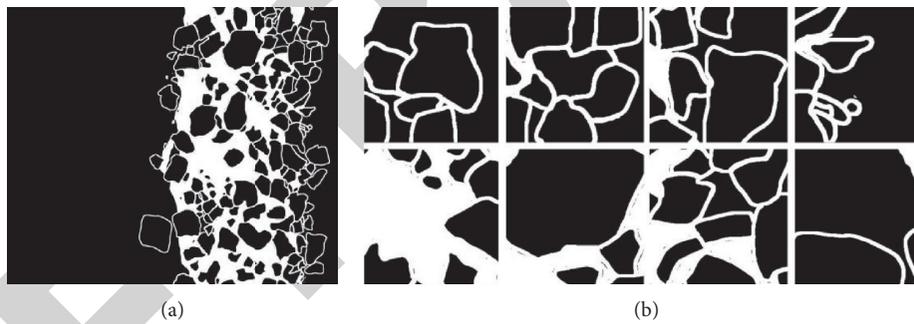


FIGURE 6: Image cutting before training. (a) Original label and (b) cutten labels.

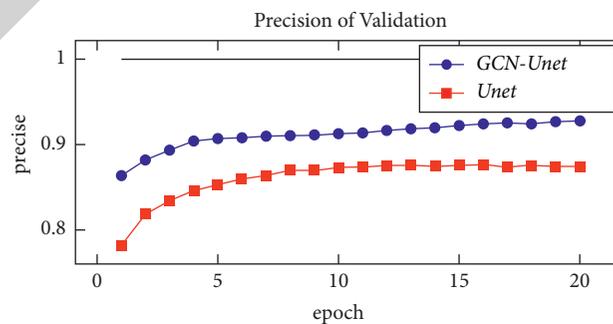


FIGURE 7: Accuracy increases during training.

78.21% and ends at 87.45%. Two training accuracy lines shows an upward trend, while GCN-UNet reaches higher. Figure 8 shows the prediction results compared to the ground truth.

5. Discussion: Compared with U-Net

As Table 1 shows, the traditional U-Net has 31,101,448 parameters, while the GCN-UNet gets 7,876,675 ones,

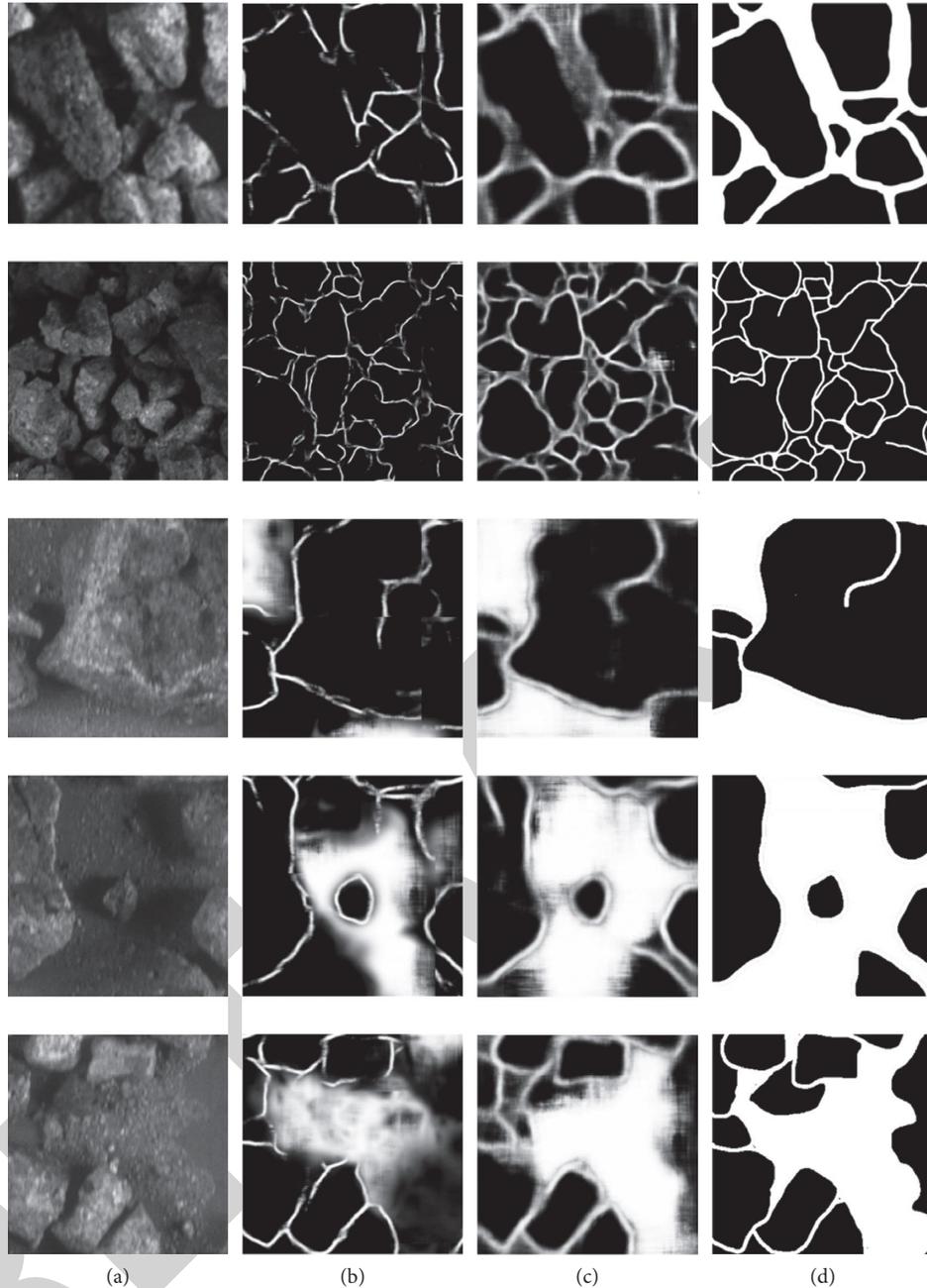


FIGURE 8: Precise contrast: (a) original images; (b) prediction by the U-net; (c) prediction by the GCN-Net; and (d) the ground truth. Degree of confidence is represented by grayscale.

TABLE 1: Hyper-parameters of GCN-Net.

Batch size	16
Input image size	288
Kernel_initializer	he_normal
Loss	MSE
Activation	Relu

around 25% of the former. The substantial reduction in parameters greatly improves the network operation rate. As Table 2 shows, the computation speed of GCN-Net of a 2048×2448 pixels image is twice as fast as the U-Net.

Moreover, after introducing appropriate loss function and proper training, the segmentation results are far better than the original network. The prediction time of the GCN-Net is more or less than half of that of the U-Net. The result is obtained by operating on NVIDIA Tesla V100-PCI-E-16 GB.

Compared with the U-Net, the GCN-Net can identify the tiny parts of the materials. Therefore, particle can be seen more clearly. For the splice between each small image, the GCN-Net can make the outline of the particle connect more smoothly. The compartment is shown in Figure 8. As Figures 8(b) and 8(c) shows, degree of confidence is represented by grayscale. Pure white indicates a confidence of

TABLE 2: Parameter and speed contrast.

	GCN-Unet	U-Net
Average speed	0.515 s	1.08 s
Total params	7876675	31101448

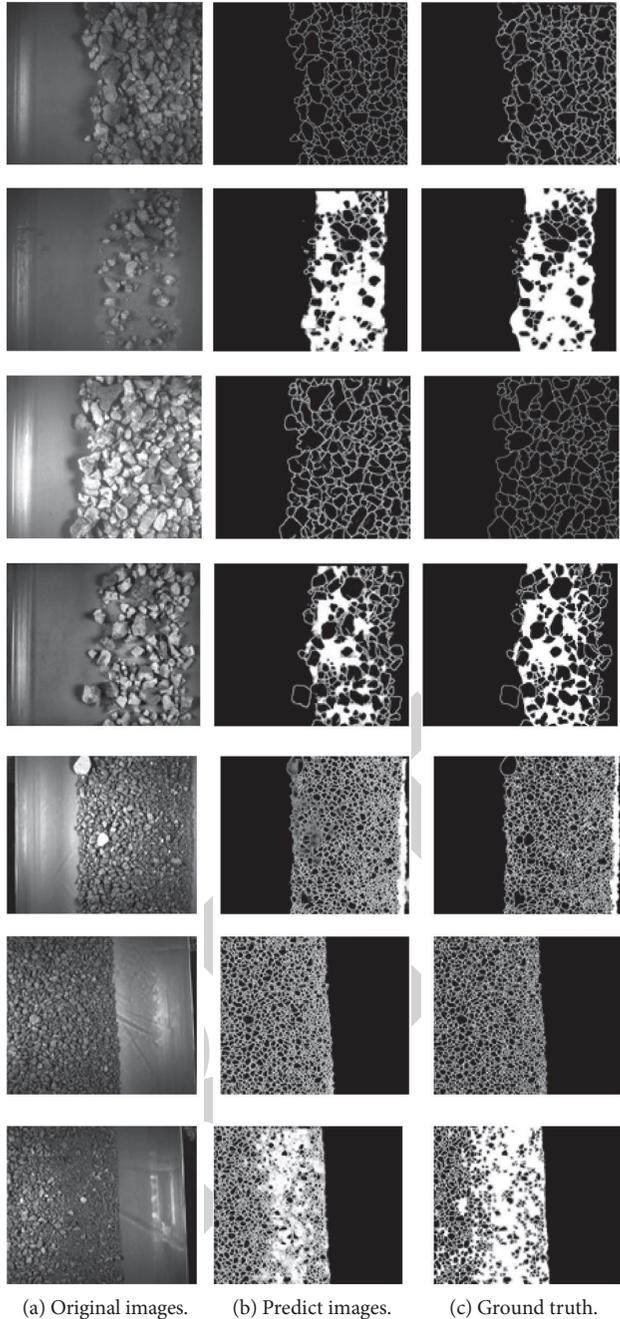


FIGURE 9: Prediction results by GCN-Unet: (a) original images; (b) predict images; and (c) ground truth.

100%. For original large images captured by camera, the result predicted by GCN-Unet is shown in Figure 9. Column Figure 9(a) shows the original images, Figure 9(b) shows predict images, and Figure 9(c) shows the ground truth.

6. Conclusion

Artificial intelligence technology is rapidly and profoundly affecting various industries, through visual perception technology real-time monitoring of closed corridor is the inevitable trend of the development of green manufacturing of intelligent raw materials field. At present, the visual intelligence perception technology in the intelligent raw material field is still in the early stage of development, cannot meet the need for automatic identification of the production anomalies of raw material transportation, and cannot be used as a commercial technology for large-scale promotion.

To identify the granularity of raw materials during transportation, a computer vision method based on the GCN-Unet algorithm is presented to replace traditional manual sampling inspection method for the corresponding overall analysis and the automatic identification process. The GCN-Unet, based on the traditional U-Net [8], combined with 8 several global convolutional network models [13] and boundary refinement blocks [13], has a better prediction ability with fewer parameters (7,876,675, while the U-Net has 31,101,448 parameters) and a higher calculating speed (about twice faster than the U-Net).

Because of data augmentation, our dataset is of good universality. After training the dataset, our model can perform good recognition under different brightness scenes.

Compared to the traditional U-Net model with a precise 88.45% and 1s per image, the GCN-Unet has 5.32% higher precision and uses half of the original computation time. It is certain that the improved U-Net architecture can be applied successfully to many industrial tasks.

The visual intelligence perception technology to be developed in this project can effectively promote the process of green manufacturing of steel raw materials and the development of industrial solutions of intelligent raw materials, promote the deep integration of artificial intelligence technology and industry, and then drive the innovation and transformation of the steel industry.

As has been mentioned above, our proposed GCN-Unet is of great practical and innovative importance for the following reason: compared with the traditional UNet, our method not only increases inference speed with fewer parameters but also has finer segmentation performance, which enables the proposed structure to be applied in universal industrial scenes.

Data Availability

No data were used to support this study.

Conflicts of Interest

There are no potential conflicts of interest in this study.

Acknowledgments

This work was financially supported by National Key RD Project (2020YFB12800) and Chongqing Municipal Technology Innovation and Application Development Project (sctc2020jscx-msxmX0158).

References

- [1] T. Yang, X. Yi, S. Lu, K. H. Johansson, and T. Chai, "Intelligent manufacturing for the process industry driven by industrial artificial intelligence," *Engineering*, vol. 7, no. 9, pp. 1224–1230, 2021.
- [2] Y. Jiang, N. Yang, Q. Yao, Z. Wu, and W. Jin, "Real-time moisture control in sintering process using offline-online NARX neural networks," *Neurocomputing*, vol. 396, pp. 209–215, 2020.
- [3] H. Wu and J. Zhao, "Automated visual helmet identification based on deep convolutional neural networks," *13th International Symposium on Process Systems Engineering (PSE 2018)*, vol. 44, pp. 2299–2304, 2018.
- [4] Industrial Internet Industry, "Industrial intelligence white paper Alliance," 2020, https://www.miit.gov.cn/ztzl/rdzt/gylhw/cgzs/art/2020/art_e1842c433fce43e39a45ce96b50213a.html Available from.
- [5] X. Li, H. Chen, X. Qi, Q. Dou, and P. A. Fu Heng, "H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Transactions on Medical Imaging*, vol. 37, no. 12, pp. 2663–2674, 2018, Dec.
- [6] M. B. ShuvoShuvo, R. Ahommed, S. Reza, and M. Hashem, "CNL-UNet: a novel lightweight deep learning architecture for multimodal biomedical image segmentation with false output suppression," *Biomedical Signal Processing and Control*, vol. 70, p. 102959, 2021.
- [7] H. You, L. Yu, S. Tian, and W. MaXingXinCai, "MC-Net: multiple max-pooling integration module and cross multi-scale deconvolution network," *Knowledge-Based Systems*, vol. 231, p. 107456, 2021.
- [8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," 2015. ArXiv, abs/1505.04597.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016.
- [10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition, arXiv, preprint," arXiv:1409.1556, 2014.
- [11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [12] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015.
- [13] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun, "Large kernel matters-improve semantic segmentation by global convolutional network," in *Proceedings of the 20 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 2020.
- [14] F. Milletari, N. Navab, and S. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, California, CA, USA, 2016.
- [15] M. Kesici, C. B. Saner, Y. Yaslan, and V. M. Genc, "Cost sensitive class-weighting approach for transient instability prediction using convolutional neural networks," in *Proceedings of the 2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*, Turkey, 2019.
- [16] H. Wei and P. Peng, "The segmentation of retinal layer and fluid in SD-OCT images using mutex dice loss based fully convolutional networks," *IEEE Access*, vol. 8, pp. 60929–60939, 2020.
- [17] L. Wang, C. Wang, Z. Sun, and S. Chen, "An improved dice loss for pneumothorax segmentation by mining the information of negative areas," *IEEE Access*, vol. 8, pp. 167939–167949, 2020.