HONGDI JING [1,2*], WENXUAN HE[3], MIAO YU[1,2*], XIN LI[1,2],
XINGFAN ZHANG[1,2], XIAOSONG LIU[1,2], YANG CUI[1,2], ZHIJIAN WANG[1,2]

RESEARCH ON ORE FRAGMENTATION RECOGNITION METHOD BASED ON DEEP LEARNING

The degree of ore fragmentation in mining sites is closely related to crushing efficiency, equipment safety, beneficiation efficiency, and mining costs. Aiming to address the challenges of high labour intensity and low accuracy during manual ore fragmentation measurement at the mine site, this paper proposes a method for ore fragmentation recognition based on deep learning. This method not only uses the residual neural network structure to form the backbone feature extraction network of CSPDarkNet21 under the Darknet framework but also selects the simple two-way fusion feature PANet as the feature extraction network under the condition of only needing to identify large ore. PANet is simplified from three feature layers to one feature layer, which speeds up model training and prediction. The research results show that with a 6% decrease in accuracy, the model training time is reduced by 13 times, and the model running efficiency is improved by 21.2 times, significantly shortening the model development time. At the same time, CIOU calculates the loss value to make model training more stable. After the ore identification is completed, the real size of the ore can be obtained by calculating the pixel area of the prediction frame using the ore fragmentation judgement method.

Keywords: Underground mines; ore fragmentation; visual identity; recognition; deep Learning

1. Introduction

The physical and mechanical properties of rocks are significantly influenced by their internal structure, especially the presence of joints and fractures [1]. They can be natural or caused by

[1]   CHINESE ACADEMY OF SCIENCES,SHENYANG INSTITUTE OF AUTOMATION, SHENYANG 110016, CHINA
[2]   CHINESE ACADEMY OF SCIENCES, INSTITUTES FOR ROBOTICS AND INTELLIGENT MANUFACTURING, SHENYANG 110169, CHINA
[3]   ANSTEEL GROUP MINING CORPORATION LIMITED, ANSHAN 114001, CHINA

*   Corresponding authors: jinghongdi@163.com, yumiao1@sia.cn
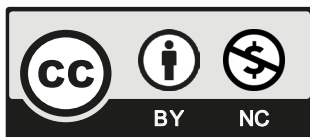
human activities such as mining and blasting. Pre-existing cracks in brittle substances seem to be the main cause of their breakage under various loading conditions [2]. Therefore, reasonable blasting and rock-cutting techniques can effectively control the fragmentation of rock mass into ore blocks, which is also a research focus for scholars in the mining industry [3,4].

The problem of ore fragmentation has always been an important factor restricting mining safety, efficiency and cost. The increase in ore fragmentation will cause difficulty in transporting the ore in the concentrator. Similarly, the secondary blasting of ore will cause problems such as the burden of ventilation and the decrease of mine transportation efficiency. Eventually, the mining cost of ore will rise and the profitability of mining companies will decline, which will hinder the safe and efficient development of the mining industry.

Traditional ore fragmentation measurement mainly relies on manual labour, which leads to a huge workload of surveying personnel. Moreover, the accuracy and efficiency of manual measurement are not in line with the development trend of the mining industry. With the continuous development of visual recognition algorithms, the depth of neural networks is also increasing. In many respects, its recognition accuracy can surpass humans, so it is possible to use visual recognition technology to solve the problem of ore fragment recognition. At present, semantic segmentation is the main method of artificial intelligence recognition of ore fragmentation, such as Unet [5], PSPNet [6], RestNet [7] and so on. The operation and training efficiency of these algorithms is lower than target recognition algorithms, and the resolution rate is slightly higher.

In recent years, many researchers have achieved great breakthroughs in the improvement and innovation of ore identification and segmentation. Among them, Hongdong Wang [12] successfully classified the ore under the microscope based on the K-Means method, and the recognition accuracy reached 90.44%. Chengzhao Liu [9] established a comprehensive ore recognition model based on depth learning, machine learning and other algorithms, which has high recognition accuracy. Ye Zhang et al. [10] established a mineral micrographic model based on the INCEPTION-V3 network structure and ensured that the final accuracy of the recognition model was more than 90%. Sascha T. Ishikawa et al. [11] use an artificial neural network to train spectral data and obtain a mineral intelligence classification model, the model recognition accuracy can reach 83%. In the study of Mariusz Młynarczuk et al., polarisation microscopy and the K-nearest neighbour algorithm were used to study the 27,000 images of 9 kinds of rocks, and the recognition accuracy of the model can reach 99%.

This study is based on convolution under the DarkNet framework, and the residual convolutional neural network is used as the backbone feature extraction network. Under the simple two-way fusion feature extraction network, a small-size feature layer is used to predict output results. The ore fragmentation recognition method in this study is finally proposed, which is called OIR for short. The specific process of the OIR method is shown in Fig. 1.

2. Network Selection

2.1. Network structure

The OIR method is mainly composed of CSPDarkNet21 [13] backbone network, simple two-way feature fusion PANet [14] and ore fragmentation judgement structure. Among them, as the main feature extraction network of the algorithm, CSPDarkNet21 classifies and locates
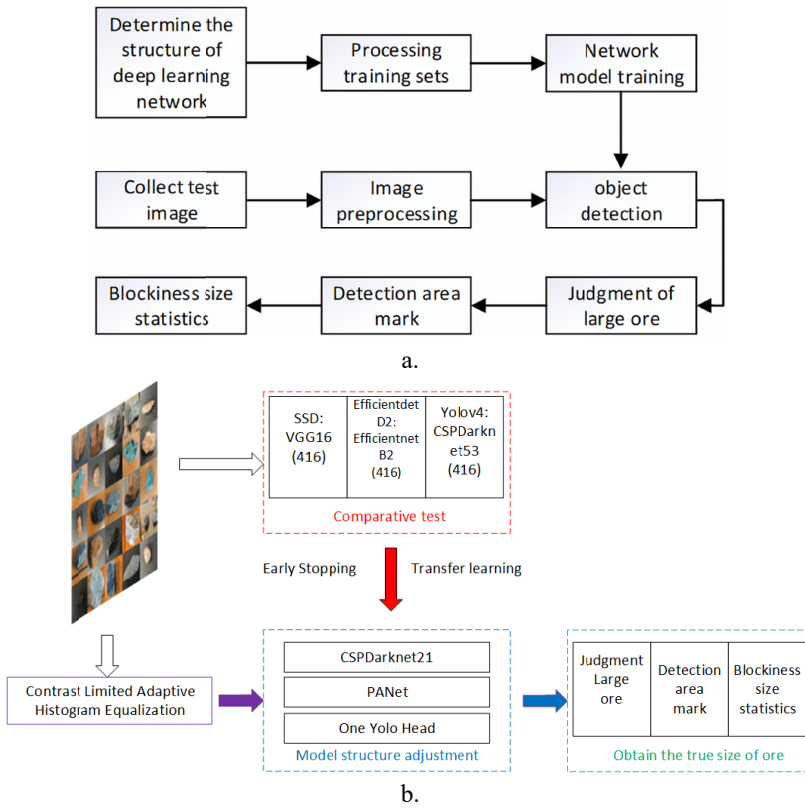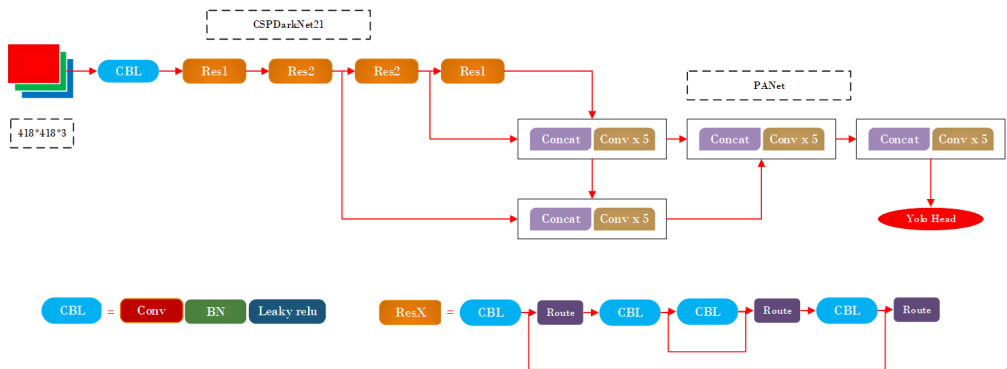
Fig. 1. Training and testing process of OIR method

the image candidate frames. On the other hand, PANet is responsible for the feature fusion and judgement of the characteristic layers in OIR, and the ore fragmentation judging structure is finally used to judge the ore fragmentation. The specific structure is shown in Fig. 2.

Fig. 2. CSPDarrkNet21 Network structure

2.1.1. Backbone feature extraction network

CSPDarkNet21 is composed of Residual block [15] and DarknetConv2D_BN_Leaky. Among them, LeakyRelu [16] was selected as the activation function to speed up the training efficiency. After the images pass through the backbone feature extraction network, three size feature layers could be output, and then they will be input to PANet for processing and judging features. The specific structure is shown in TABLE 1.

TABLE 1

CSPdarknet21 Network structure parameters

Network structure	Type	Convolution kernel number	Image output size
Input			416×416×3
DarknetConv2D_BN_Leaky	Conv	32	416×416×32
Residual block_1	Conv	64	416×416×64
Max pool	Step	2	208×208×64
Residual block_2	Conv	128	208×208×128
Residual block_3	Conv	128	208×208×128
Max pool	Step	2	104×104×128
Residual block_4	Conv	256	104×104×256
Residual block_5	Conv	256	104×104×256
Max pool	Step	2	52×52×256
P3			52×52×256
Residual block_6	Conv	512	52×52×512
Residual block_7	Conv	512	52×52×512
Max pool	Step	2	26×26×512
P4			26×26×512
Residual block_8	Conv5	1024	26×26×1024
Max pool	Step	2	13×13×1024
P5			13×13×1024

1) DarknetConv2D_BN_Leaky convolution
   DarknetConv2D_BN_Leaky is composed of a 2D convolutional layer and a regularisation layer under the framework of Darknet [16]. The regularisation parameter is 5e-4, and LeakyRelu should be used as the activation function of the network. Darknet is a neural network framework written in C language and CUDA, which enables DarknetConv2D_BN_Leaky to be installed quickly and easily, and GPU computing is also feasible.

2) Residual block
   Compared with the initial Residual block [17], the residual block is split by the CSPnet structure, the main part of the residual block remains unchanged, and the other part is directly connected to the output item after a little processing, finally a large residual edge is constituted. Among them, the Residual block structure is mainly composed of DarknetConv2D_BN_Leaky and some other small structures.

2.1.2. PANet Feature Extraction Network

When extracting features, the high-level feature maps mainly reflect the overall informa-tion of the objects, while the low-level feature maps mainly reflect the texture information of the objects. Compared with FPN [18] (one-way fusion), PANet's bottom-up path is enhanced. At the same time, low-level precise positioning is used to shorten the propagation path of in-formation. The features of all pyramid layers are used by each proposal, which avoids random allocation of proposals, and the fusion of fully connected layers is adopted, which increases the source of prediction information.

The backbone feature extraction network generates three-size feature layers which are input to the PANet feature extraction network for sampling and fusion. The final predicted position is only in the high-level feature layer, that is, the feature layer with the smallest size. The specific structure is shown in Fig. 3.
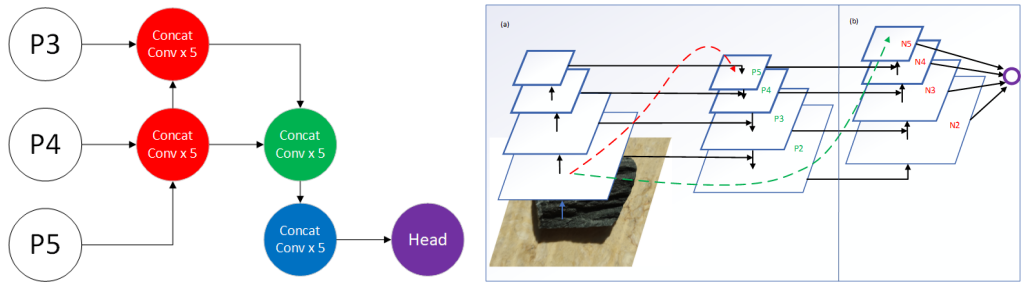
Fig. 3. PANet simple two-way fusion feature network

2.2. Construction of loss function

In the method of this paper, CIOU [19] is used as the loss function of regression optimiza-tion. Compared with IOU, factors such as the scale, overlap rate, and penalty items of the objects and the anchor boxes are considered in CIOU, which makes the regression of the anchor box more stable, and the training process is not prone to divergence. The specific CIOU is shown in Formula 1.

$$CIOU = IOU - \frac{\rho^2\left(b, b^{gt}\right)}{c^2} - \alpha v \tag{1}$$

Among them, $\rho^2(b, b^{gt})$ respectively represent the Euclidean distance between the centre points of the anchor box and the ground truth box. c represents the diagonal distance of the smallest closed area that contains both the anchor box and the ground truth box. IOU is the overlap value of the anchor box and the ground truth box. The calculations of α and v are shown in formula 2 and formula 3.

$$\alpha = \frac{v}{1 - IOU + v} \tag{2}$$

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right)^2 \tag{3}$$

Finally, calculate the loss value:

$$LOSS = 1 - IOU + \frac{\rho^2\left(b,b^{gt}\right)}{c^2} - \alpha v \tag{4}$$
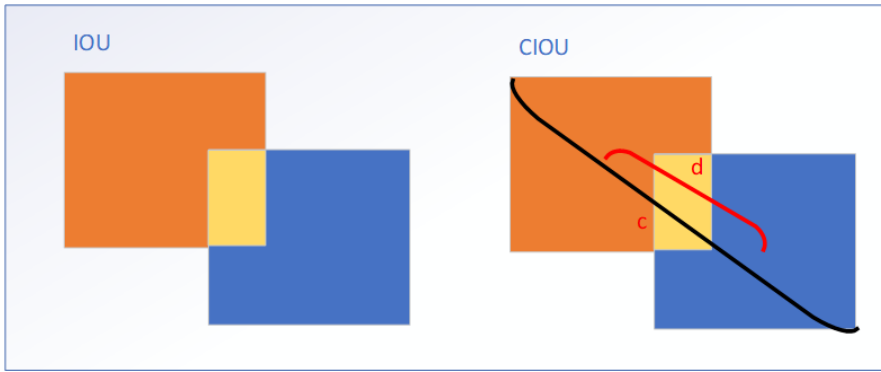
IOU and CIOU are shown in Fig. 4.

Fig. 4. Schematic diagram of IOU and CIOU

3. Experiment

3.1. Data collection

IOU and SCORE are used by traditional neural networks models, such as Faster-RCNN [20], YOLOV4 and Efficientnet [21] to judge objects. Based on the traditional two judgement modules, the ore fragmentation judgement module is added to the algorithm of this paper. Finally, the coordinate points and the measuring scale of the obtained anchor boxes are used to determine the ore fragmentation.

As we all know, the image sizes of the same object captured by the camera at different heights are different. The farther the camera is from the object, the smaller the object will be. However, different camera parameters make it impossible to have a uniform ratio to obtain the relationship between the area of the image pixel points and the actual object area. Therefore, this paper proposes a comparison ruler structure for measuring the ratio before testing. The calculation method is shown in Formula 5, and the specific schematic diagram is shown in Fig. 5.

$$T = \left(\frac{\sqrt{\left(X_{max} - X_{min}\right)^2 + \left(Y_{max} - Y_{min}\right)^2}}{S_r} * \frac{416}{S_{max}}\right)^2 \tag{5}$$

Among them, X_{max}, X_{min}, Y_{max}, and Y_{min} are the pixel coordinate points at both ends of the ruler. The Pythagorean Theorem is used to calculate the distance between the ruler and the pixels in the picture. S_r is the length of the actual ruler, and the size of the image input during the neural network training process is fixed and scaled in equal proportions. The size of the training picture in this algorithm is 416×416, and S_{max} is the longest side of the actual size of the picture taken by the camera. The obtained T is the area of the pixel corresponding to 1 cm^2.

Fig. 5. Schematic diagram of reference object

3.2. Data set preparation

All experimental samples were randomly sampled from the underground stope and rock dump of Yanqianshan iron mine, as shown in Fig. 6. The main recognition part of this article will identify the ore, and the feature of the ore is relatively simple. The requirements for the pictures in the input data set are low, so grayscale images could be used as the training data set. The data set images are made from the video recorded by high-speed cameras, and the definition and blur degree are not much different from those taken by normal mobile phone cameras.

1) The inconsistent light conditions between the pictures collected on site and the pictures collected in the laboratory will cause the robustness of recognition to decrease. It will have an effect even under grayscale conditions. Therefore, 24 hours of work videos are continuously collected in the on-site production environment, and 360 photos are evenly selected as the data set in each hour of the video.

2) The resolution of the collected pictures is 1920×1080, and a total of 3600 pictures are selected from the video. After all data pictures are preprocessed, a data set of 7200 pictures is obtained. In the end, 85% of all data pictures are used as the training set, 10% as the test set, and 5% as the validation set.

a. Underground stope   **b.** Rock dump
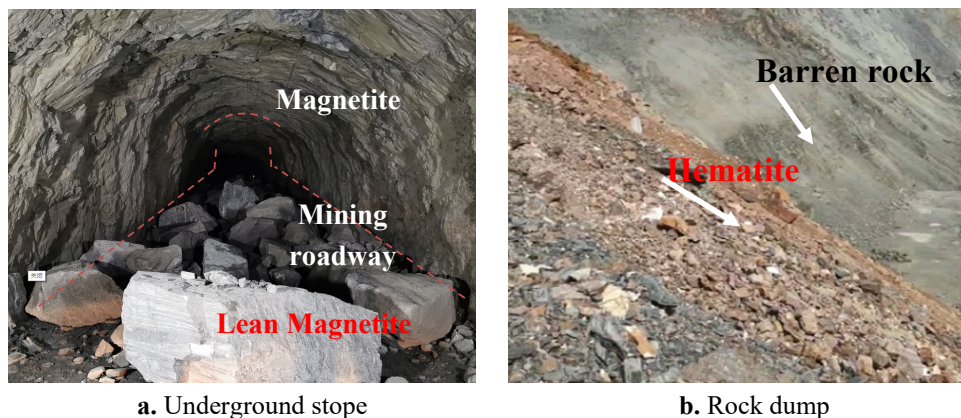
Fig. 6. On-site sampling at Yanqianshan Iron Mine

3.3. Model training

Transfer learning mode is to apply the trained model to target task learning. Because the similarity of shallow weight information can greatly reduce the difficulty of training, limited data usage can obtain a good training effect. However, after the model structure has been changed, the existing pre-training model is not applicable, so the voc2012 dataset is used for 20epoch pre-training to obtain the pre-training model of this algorithm. Then, the pre-training model obtained is used in the following formal training.

Neural network recognition and ore fragmentation judgement of the OIR method are based on Keras in Tensorflow under Python 3.7. In the training process, the network model is trained by GPU, and the early stopping is used to control training steps. The ore fragmentation judgement structure does not participate in the training.

The training set is used to train the Unet and OIR models, and the training parameters are shown in TABLE 2. Batch_size represents the number of input training pictures, epochs represent the number of training times required for the model, train_size represents the total number of training set pictures, and freeze_epoch represents the number of training times after freezing the first 50% of the layer.

Both models are controlled by early stopping; after three epochs, the training will be stopped if the LOSS value does not decrease or decreases less than 0.1. A cyclic cosine annealing learning rate schedule could be used to adjust the learning rate of OIR methods training.

TABLE 2

Model training parameters

Training model	Unet	OIR
Train_size	1200	3600
Batch_size	1	10
Epochs	10	30
Freeze_epoch	0	15

3.4. Ore fragmentation judgement

In the actual calculation process, the size of the anchor box is used to calculate the actual size of the ore. Since there is a certain error between the shape of the anchor box and the shape of the actual ore, a correction parameter is needed to make the algorithm more accurate. Under tensorflow2.2, semantic segmentation is used to mark the pictures as the correct values, and compared with the values obtained by the algorithm in this paper, the final correction error parameter is determined to be 0.86. The calculation method is shown in Formula 6.

$$S = T * W * H * 0.86 \tag{6}$$

Among them, S represents the extreme value of the ore fragmentation judgement, W and H are the length and width values obtained by the anchor box, and T represents the size of the measured scale. The 0.86 error correction parameter can be changed according to the actual situation, and the calculation example is shown in Fig. 7.
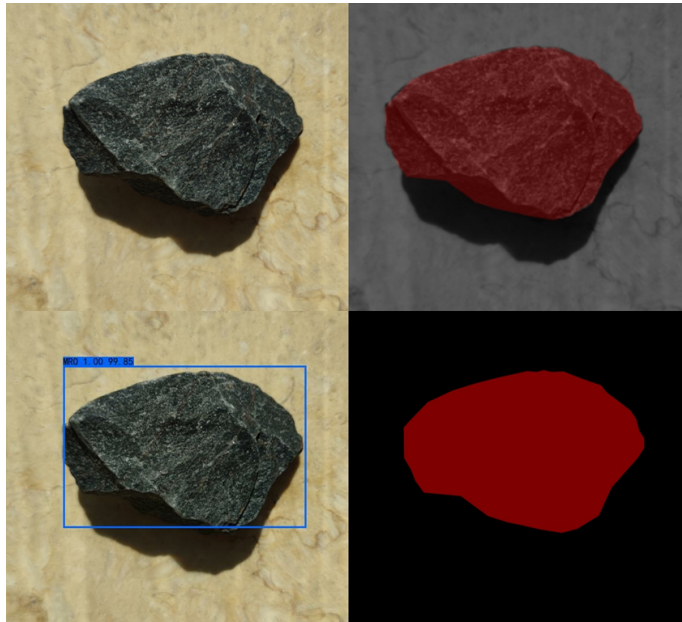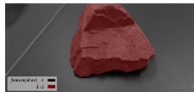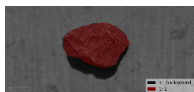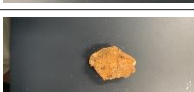
Fig. 7. Parameter selection

4. Result and discussion

In order to verify the processing speed and effectiveness of the OIR method, 10% of the total sample is selected as the test set to compare the OIR model and Unet semantic segmentation. The marked pixel area is used as the correct value for comparison. The top three values of the anchor box are the recognition category, the confidence coefficient, and the pixel area of the anchor box/1000. The specific comparison is shown in TABLE 3.

456

Comparison of Unet segmentation and OIR

Original image	Label image	Unet split image	OIR recognition image	Unet/OIR accuracy
				97.34/91.17
				99.77/92.10
				98.82/90.96
				94.13/90.23
				98.88/90.00
				97.13/92.18
				95.65/92.64
				96.41/91.21
				96.15/94.33

It can be seen from TABLE 3 that the accuracy of the OIR model is 6% lower than the accuracy of the Unet semantic segmentation model on average. In the training process, because the target detection algorithm structure is used in the OIR method, and only one feature size is used for prediction, the resources required for training are effectively reduced. In the case that the number of Unet data sets is 1/3 of the number of OIR method data sets, the training time of Unet for one epoch is 13 times that of the OIR method. When the accuracy is acceptable, the OIR method has a stronger practical application value. In the recognition process, the average time for Unet to recognize a picture is 297 ms, while the average recognition time for OIR is 14 ms. The Unet model takes 21.2 times the time of the OIR model. For the above reasons, the OIR method can also quickly recognize and judge the ore fragmentation in the video (for example, on a conveyor belt).

5. Conclusion

The accurate recognition of ore fragmentation is a critical challenge in the mining industry, affecting various aspects such as crushing efficiency, equipment safety, beneficiation efficiency, and overall mining costs. Traditional methods of ore fragmentation measurement mainly rely on manual labour and lack the accuracy required for modern mining operations. Moreover, existing algorithms, predominantly semantic segmentation algorithms like Unet or instance segmentation algorithms like PSPnet, while offering high accuracy, fall short in terms of dataset production efficiency and model prediction speed when compared to target recognition algorithms. To address these issues, this paper introduces the Ore Identification and Recognition (OIR) method, a deep learning-based approach designed to enhance the efficiency and accuracy of ore fragmentation detection. The OIR method utilises a deep learning model that not only identifies ore fragmentation but also captures size information about the target ore, providing a practical metric for evaluation. In summary, the main results of this study are as follows:

(1) This article proposes an ore fragmentation detection method (OIR) based on deep learning. The algorithm can determine the degree of fragmentation of the identified ores and calculate the actual area of the target ores.

(2) In the case of a 6% decrease in accuracy, the model training time was reduced by 13 times, and the model running efficiency was improved by 21.2 times, greatly shortening the model development time.

(3) The size of the OIR model is 11.3MB, which is 5.1MB after being converted to the tensorflow-lite model. The high degree of miniaturisation enables the model to be mounted on various embedded devices to realise convenient on-site applications.

List of abbreviations

TABLE 4

List of abbreviations

Symbol	Description
1	2
CSPDarkNet21	A convolutional neural network structure based on the Darknet framework, used for feature extraction.
PANet	A simplified two-way fusion feature network, used for feature extraction, reduced from three feature layers to one.
CSPnet	A network structure used for segmenting residual blocks.
DarknetConv2D_BN_Leaky	A network structure composed of a 2D convolutional layer and a regularisation layer under the Darknet framework, using LeakyRelu as the activation function.
LeakyRelu	An activation function used to speed up training efficiency.
IOU	Intersection over Union, a metric used to measure the overlap between the predicted and true bounding boxes.
CIOUS	Complete Intersection over Union, a loss function that considers scale, overlap rate, and penalty items of the objects and the anchor boxes.
Unet	A convolutional neural network structure used for image segmentation.
Residual block	A structure used to build deep network architectures.

458

1	2
P3, P4, P5	Different feature layers in the network, corresponding to different feature map sizes.
Conv	Convolution operation.
Max pool	Max pooling operation.
T	The area of a pixel corresponding to 1 square centimetre.
GPU	Graphics Processing Unit, used for accelerating the training and prediction of neural networks.
Batch_size	The number of training images input in each batch.
epochs	The total number of training iterations the model undergoes.
train_size	The total number of images in the training set.
freeze_epoch	The number of training iterations after freezing the first 50% of the layers.
voc2012	A commonly used image recognition dataset for model pre-training.

Authors' contributions

Hongdi Jing contributed to the conception of the study; Wenxuan He and Miao Yu performed the experiment; Xin Li and Xingfan Zhang contributed significantly to analysis and manuscript preparation; Xiaosong Liu, Zhijian Wang and Yang Cui performed the data analyses and wrote the manuscript.

Acknowledgement

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] M. Mohebbi, A.R. Yarahmadi Bafghi, M. Fatehi Marji, J. Gholamnejad, Rock mass structural data analysis using image processing techniques (Case study: Choghart iron ore mine northern slopes). Journal of Mining and Environment **8** (1), 61-74 (2017). DOI: https://doi.org/10.22044/JME.2016.629

[2] H. Haeri, K. Shahriar, M.F. Marji, P. Moarefvand, A coupled numerical – experimental study of the breakage process of brittle substances. Arabian Journal of Geosciences **8**, 809-825 (2015).
DOI: https://doi.org/10.1007/s12517-013-1165-1

[3] M. Fatehi Marji, Modelling of cracks in rock fragmentation with a higher order displacement discontinuity method. PhD Thesis 1 (1), METU, Ankara, Turkey (1996).

[4] M. Lak, M.F. Marji, A.Y. Bafghi, A. Abdollahipour, Analytical and numerical modelling of rock blasting operations using a two-dimensional elasto-dynamic Green's function. International Journal of Rock Mechanics and Mining Sciences **114**, 208-217 (2019). DOI: https://doi.org/10.1016/j.ijrmms.2018.12.022

[5] H.C. Bazame, J.P. Molin, D. Althoff, et al., Detection, classification, and mapping of coffee fruits during harvest with computer vision. Computers and Electronics in Agriculture **183** (2021).
DOI: https://doi.org/10.1016/j.compag.2021.106066

[6]   K. Chen, Z. Zeng, J. Yang, A deep region-based pyramid neural network for automatic detection and multi-classification of various surface defects of aluminium alloys. Journal of Building Engineering **43**, 102523 (2021).

[7]   W. Chen, X. Zhong, J. Zhang, Optimization Research and Defect Object Detection of Aeroengine Blade Boss Based on YOLOv4. Journal of Physics: Conference Series **1746** (1), 012076 (2021).
      DOI: https://doi.org/10.1088/1742-6596/1746/1/012076

[8]   R. Cheng, X. He, Z. Zheng, et al., Multi-Scale Safety Helmet Detection Based on SAS-YOLOv3-Tiny. Applied Sciences **11** (8), 3652 (2021). DOI: https://doi.org/10.3390/app11083652

[9]   X. Cui, C. Zou, Z. Wang, Remote sensing image recognition based on dual-channel deep learning network. Multimedia Tools and Applications **80** (18), 27683-27699 (2021). DOI: https://doi.org/10.1007/s11042-021-11079-5

[10] H. Deng, J. Cheng, T. Chen, et al., Research on Iron Surface Crack Detection Algorithm Based on Improved YOLOv4 Network. Journal of Physics: Conference Series (2020). DOI: https://doi.org/10.1088/1742-6596/1631/1/012081

[11] N.J. Francis, N.S. Francis, S.V. Axyonov, et al., Diagnostic of Cystic Fibrosis in Lung Computer Tomographic Images using Image Annotation and Improved PSPNet Modelling. Journal of Physics Conference Series **1611**, 012062 (2020). DOI: https://doi.org/10.1088/1742-6596/1611/1/012062

[12] D. Jiang, G. Li, C. Tan, et al., Semantic segmentation for multiscale targets based on object recognition using the improved Faster-RCNN model. Future Generation Computer Systems **123** (2021).
      DOI: https://doi.org/10.1016/j.future.2021.04.019

[13] C. Liu, M. Li, Y. Zhang, et al., An Enhanced Rock Mineral Recognition Method Integrating a Deep Learning Model and Clustering Algorithm. Minerals **9** (9), 516 (2019). DOI: https://doi.org/10.3390/min9090516

[14] H. Liu, K. Fan, Q. Ouyang, et al., Real-time small drones detection based on pruned yolov4. Sensors **21** (10), 3374 (2021). DOI: https://doi.org/10.3390/s21103374

[15] X. Liu, Y. Zhang, H. Jing, et al., Ore image segmentation method using U-Net and Res_Unet convolutional networks. The Royal Society of Chemistry **16**, (2020). DOI: https://doi.org/10.1039/C9RA05877J

[16] S.A. Magalhães, L. Castro, G. Moreira, et al., Evaluating the Single-Shot MultiBox Detector and YOLO Deep Learning Models for the Detection of Tomatoes in a Greenhouse **21** (10), 3569 (2021).
      DOI: https://doi.org/10.48550/arXiv.2109.00810

[17] M. Młynarczuk, A. Górszczyk, B. Ślipek, The application of pattern recognition in the automatic classification of microscopic rock images. Computers & Geosciences **60**, 126-133 (2013).
      DOI: https://doi.org/10.1016/j.cageo.2013.07.015

[18] B. Nigam, A. Nigam, R. Jain, et al., COVID-19: Automatic detection from X-ray images by utilising deep learning methods. Expert Systems with Applications **176**, 114883 (2021).

[19] S.T. Ishikawa, V.C. Gulick, An automated mineral classifier using Raman spectra. Computers & Geosciences **54**, 259-268 (2013). DOI: https://doi.org/10.1016/j.cageo.2013.01.011

[20] H. Wang, M. Lei, Y. Chen, et al., Intelligent Identification of Maceral Components of Coal Based on Image Segmentation and Classification. Applied Sciences **9** (16), 3245 (2019). DOI: https://doi.org/10.3390/app9163245

[21] Y. Zhang, M. Li, S. Han, et al., Intelligent Identification for Rock-Mineral Microscopic Images Using Ensemble Machine Learning Algorithms. Sensors **19** (18), 3914 (2019). DOI: https://doi.org/10.3390/s19183914