

ZIBELINE INTERNATIONAL™
PUBLISHING

ISSN: 2521-0831 (Print)

ISSN: 2521-084X(Online)

CODEN: MSMAD

Matrix Science Mathematic (MSMK)

DOI: <http://doi.org/10.26480/msmk.02.2024.56.61>

RESEARCH ARTICLE

THE VISUAL WEIGHING METHOD OF MINING DUMP ATRUCK BASED ON RESNET

Kai Bai^a, Likun Zhao^a, Wenqing Che^b, Beijun Guo^b, and Zhanlong Li^{a*}^aSchool of Vehicle and Transportation Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China^bEntrepreneur of Shanxi Jiacheng Hydraulic Technology Center, Xinzhou City, Shanxi Province, Xinzhou 034100, China*Corresponding Author Email: lizi@tyust.edu.cn

This is an open access article distributed under the Creative Commons Attribution License CC BY 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ARTICLE DETAILS

Article History:

Received 23 May 2024

Revised 15 June 2024

Accepted 10 July 2024

Available online 12 July 2024

ABSTRACT

Intelligent weighing systems play a significant role in guiding coal production and assisting production decisions. However, the traditional weighing method not only cannot guarantee efficiency, but also is prone to artificial fraud, causing economic losses to enterprises. In view of the needs of industrial production, this paper studies a visual weighing method for mining dump trucks with ResNet deep learning network as the core. Firstly, the collected dataset is processed by the Resize function and the Blend function, and then the ResNet neural network is trained with the processed dataset, and finally the cross-entropy loss function and Adam optimization strategy are used to improve the recognition accuracy. The results show that the final recognition accuracy is 60%.

KEYWORDS

Deep learning; ResNet; Dump truck; Intelligent weighing.

1. INTRODUCTION

Under the background of globalization, there is an urgent need for mining enterprises to carry out the reform of informatization and intelligence in order to adapt to the increasingly fierce market competition environment (Rylnikova et al., 2017; Clausen et al., 2022; Sunhare et al., 2022). To this end, intelligent mine based on the cloud computing technology, internet of things and big data technology emerge as the times require.

The weighing of dump trucks is an important link in the production system of intelligent mine. The traditional method of the weighing of dump trucks is manually inputting the data into the computer. This method is not only time-consuming and laborious, but also easy to make mistakes. Therefore, it is of great significance to design an efficient and intelligent weighing method.

ResNet is a widely used recognition method (Shafiq and Gu, 2022). ResNet adequately utilizes the information of the shallow network and solves the problem of gradient disappearance and gradient explosion (Kanai et al., 2017; Kag et al., 2019; Koonce, 2021). Based on this, ResNet is widely used in the field of image recognition. Wightman et al. investigated the performance of the vanilla ResNet-50 (Wightman R et al., 2021). It was found that the vanilla ResNet-50 could achieve 80.4% top-1 accuracy at resolution 224×224 on ImageNet-val without extra data or distillation. Cao et al. used an improved ResNet to investigate the recognition of the plant leaf disease and insect pest (Cao et al., 2021). They showed that the recognition accuracy rate using the improved ResNet was 92.45%. Xie et al. used a ResNet18 feature encoder to investigate the recognition of the rice disease and insect pest (Xie et al., 2022). It was shown that the recognition accuracy rate using the ResNet18 feature encoder could reach to 98.48%. Targ et al. found that the residual ResNet could achieve the state-of-the-art results on challenging the computer vision task (Targ et al., 2016). In order to improve the recognition accuracy rate of the rock particle size characteristics, Fu et al. proposed a rock recognition method based on multi-scale block convolutional ResNet34 network (Fu and Wang, 2022). For the multiclass classification of the lung infection, Bharati

et al. suggested a novel CO-ResNet algorithm (Bharati et al., 2021).

In this paper, in order to realize the intelligent weighing of mining dump truck, a weight recognition method based on ResNet deep learning network is proposed. Firstly, we adopt a multi-angle image acquisition method, combined with network retrieval means to enrich the number of samples, to solve the problem of complex sample library construction; secondly, on the basis of empirical analysis, we select four kinds of loss functions, L1, KL, cross-entropy loss function and multiclassification folding loss function as the test structure, compare and analyze the recognition effect of different network structures, with a view to finding the loss function that can ensure the best matching of recognition accuracy and operational efficiency. In order to find the best loss function that can ensure the matching of recognition accuracy and operation efficiency. The algorithm is designed to meet the needs of dynamic target precision weighing industry and provide algorithmic support for the upgrading of industrial intelligent technology.

2. WEIGHING SYSTEM

2.1 ResNet deep networks

Neural network is a model in machine learning, by simulating the neural network of the human brain to be able to realize machine learning techniques similar to artificial intelligence. And as the basis of neural network-neuron, its mode of action can be expressed by equation (1):

$$Y = f(wx + b) \quad (1)$$

where x is the input signal, w is the weight value, b is the bias, and f is the activation function. The corresponding connection strengths of the synapses are called the weight values denoted by W. W and b are adjustable parameters in the neural network. Its expression (2):

$$Y = f(W_{1l}x_1 + W_{2l}x_2 + \dots + W_{il}x_i + b) = f(Wx + b) \quad (2)$$

Quick Response Code



Access this article online

Website:

www.matrixsmathematic.com

DOI:

10.26480/msmk.02.2024.56.61

Based on this, a model with multiple neurons can be made, as shown in Figure.1:

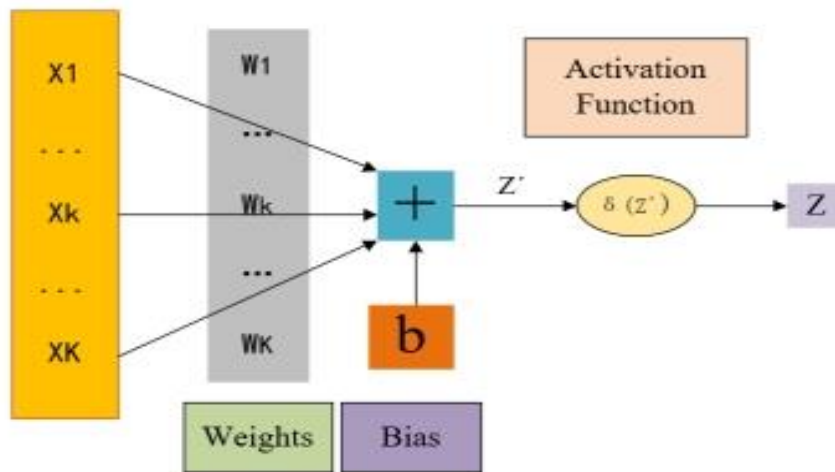


Figure 1: Schematic diagram of neuron structure

Experiments have shown that as the depth of the network continues to increase, the accuracy of the network decreases, making it difficult to train deep learning models. To solve this problem, Dr. Kaiming He proposed to solve the problem of model accuracy degradation by using residual learning. The principle is that through constant mapping, the mapping of X is equal to x , i.e., $H(X) = x$. Therefore, the problem can be transformed into the solution of the residual mapping function $F(x) = H(X) - x$. $F(x)$ is the learned residual. When the depth of the network is deeper and there is a degradation problem, it is only necessary to set the residual mapping $F(x)$ to 0, i.e., $H(x)$ is equal to the feature mapping x . At this point, the state of the network at that layer becomes a perfect state. The residual units used in this scheme are shown in Fig. 2:

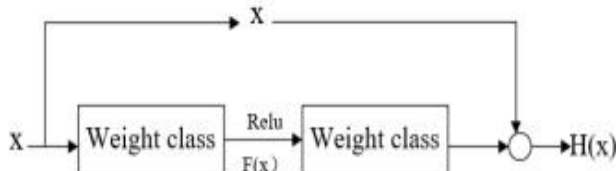


Figure 2 : Schematic diagram of residual unit

The residual unit equation 3 can be expressed as:

$$y_1 = h(x_i) + F(x_i, W_i) \tag{3}$$

$$x_{l+1} = f(y_1)$$

Where x_i and x_{i+1} denote the input and output of the l residual cell, respectively, $h(x_i) = x_i$ denotes the constant mapping, and f is the ReLU activation function. Since each residual unit generally contains multi-layer structure. Therefore, based on the above equation, we can know that the learning characteristics from l shallow to L deep layers are:

$$x_L = x_i + \sum_{i=1}^{L-1} F(x_i, W_i) \tag{4}$$

Finally, the gradient of the inverse process can be obtained by the chain rule:

$$\frac{\partial loss}{\partial x_i} = \frac{\partial loss}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_i} = \frac{\partial loss}{\partial x_L} \cdot [1 + \frac{\partial}{\partial x_i} \sum_{i=1}^{L-1} F(x_i, W_i)] \tag{5}$$

The loss function represented by the first factor of equation 5, with a loss component of x_i , reaches a gradient of 1.1 in parentheses, indicating that

the short-circuit mechanism can propagate the gradient without loss. Since the residual gradient does not go all the way to -1, it does not cause the gradient to vanish even if the value is relatively small, making it easier to learn the parameters in the residual unit.

The ResNet neural network used in this paper is shown in Table 1. This network does down sampling by directly utilizing convolution with 'stride=2' and the fully connected layer is replaced with 'global average pool' layer. And when 'feature map' size is reduced by 1/2, the number of 'feature map' is doubled and this operation maintains the complexity of the network layers. So, when the network level is deeper, it performs residual learning between three layers with three convolutional kernels of 1x1, 3x3 and 1x1, a noteworthy fact is that the number of 'feature map' in the hidden layer is smaller and is 1/4 of the number of outputs 'feature map'.

Table 1: Schematic diagram of resnet neural network structure			
Layer Name	Output Size	18-Layer	34-Layer
Conv1	112*112	7*7,64, stride2	
Conv2-x	56*56	3*3 max pool, stride2	
		0.55	0.31
Conv3-x	28*28	0.31	0.11
Conv4-x	14*14	0.51	0.17
Conv5-x	7*7	0.15	0.15
0:0:1	0.3±0.15	0.30	0.17

2.2 Loss function

In order to effectively reflect the gap between the model and the actual data, the deep learning network introduces the concept of loss function, which becomes one of the heaviest factors in the network. The loss function used in this paper is the cross-entropy loss function. Its formula 6 is shown below.

$$L = -[y \log y + (1 - y) \log(1 - y)] \tag{6}$$

In the dichotomous classification problem model, the real samples are labeled [0, 1], denoting the negative and positive classes, respectively. The model usually ends up going through a Sigmoid function, Equation 7 for this function is shown below.

$$g(s) = \frac{1}{1 + e^{-s}} \tag{7}$$

The concept is that when a probability value is output, this probability value reflects the likelihood of the prediction being a positive class, i.e., the higher the probability, the higher the likelihood. Where s is the output of the previous layer of the model, the horizontal coordinate is the predicted

output, and the vertical coordinate is the cross-entropy loss function L . The advantage of the Sigmoid function is that regardless of whether the true sample label y is 0 or 1, L represents the difference between the predicted output and y . The more the predicted output differs from y , the better the value of L . The more the predicted output differs from y , the larger the value of L . The advantage of this is that the model will tend to bring the predicted output closer to the true sample label.

3. EXPERIMENTAL PROCEDURE

3.1 Experimental Platform Introduction

In this paper, experiments were conducted on TensorFlow platform using Python language, the networks were trained and tested using NVIDIA GTX 1080 Ti 8G GPU and I5-4440 3.10 GHz CPU. Since the ResNet network can only process one mode of data, it is necessary to process three images and then input them into the ResNet network. All networks in this paper were trained with 5500 iterations and the size of the training test image was 500×500 pixels. The method used in this paper is to fuse the three images, use the end function to set different weights for the three images, and find a set of weight combinations with the best prediction effect through experiments.

3.2 Collection Of Data Sets

In order to evaluate the performance of the visual weighing system, group collected 1000 sets of photos as dataset through scientific method. Finally, 85% of the dataset is selected as the training set, 10% as the validation set and the remaining 5% as the test set. After that by analyzing the dataset, it is found that the weight distribution of the target in the pictures is between 220g-400g, so 10 classes are set up with a difference of 20g between each class. The labels of the samples are redefined, and the labels are converted from real weights to numbers 1-10.

3.3 Relationship Between Learning Rate And Loss

Experiments on some parameters in the network found that: The learning rate controls the rate of loss reduction in the network, the larger the learning rate, the faster the loss reduction in the network. At the same time, the larger the learning rate, the better the network training, although a larger learning rate can be faster to reduce the network loss, but also lead to the training of the late loss cannot be converged, and smaller learning rate may lead to the completion of the training loss cannot be minimized. Figure 4 shows the effect of different learning rates on the change of loss within the network, due to the unique structure of ResNet, the loss will not fluctuate greatly when the learning rate is larger. From Figure. 3, the loss rate is minimized when the learning rate is $10e-3$.

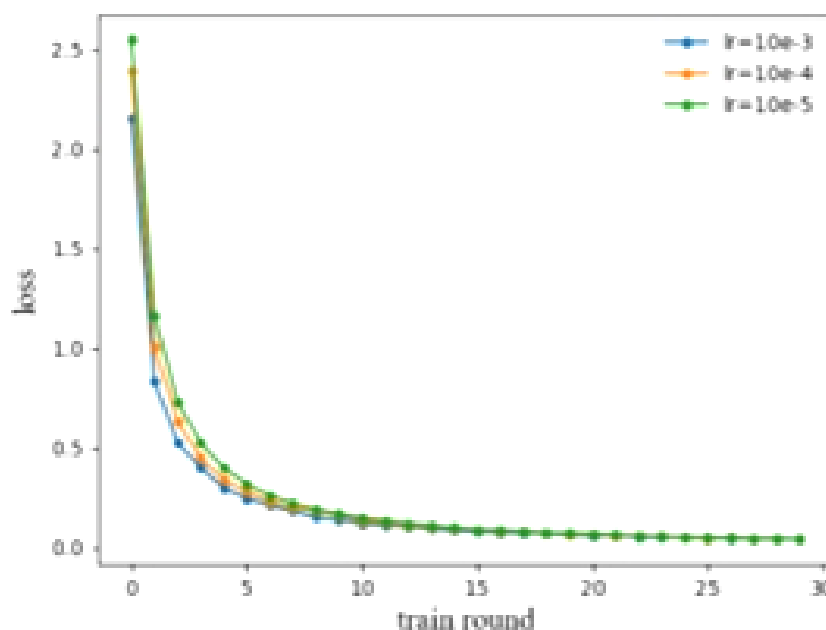


Figure 3: Relationship between learning rate and loss

3.4 Loss Function

Effect of different loss functions on loss and accuracy: Different loss functions have their own characteristics, and different loss functions need to be designed for different tasks. Fig. 4 and Table 2 show the images using

four commonly used loss functions. Where the circle size represents the current training loss value. It can be seen that the L1 loss and the KL scatter loss, perform poorly on this task, while the cross-entropy loss is better, but the accuracy is unstable. The multi-categorization page fold loss function performs well overall during training.

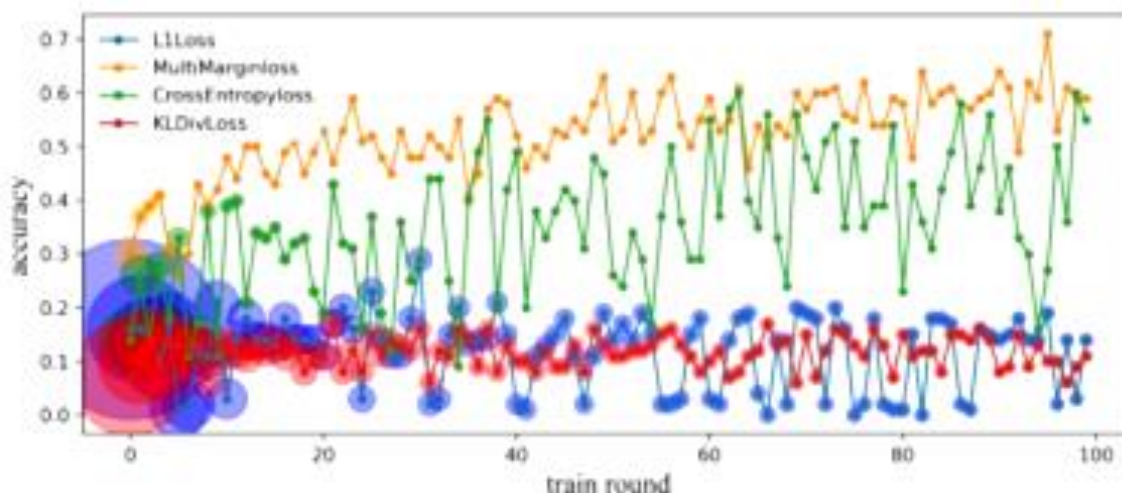


Figure 4: Loss images of four commonly used loss functions

Table 2: Influence of different loss functions on results

Loss	Accuracy	Macro-P	Macro-R	Macro-F1
KL Loss	0.13±0.2	0.06	0.19	0.08
L1 Loss	0.15±0.05	0.09	0.19	0.11
Multimarginal Loss	0.47±0.08	0.34	0.48	0.33
Cross Entropy Loss	0.65±0.07	0.44	0.57	0.49
KL Loss	0.13±0.2	0.06	0.19	0.08

4. EXPERIMENTAL PROCEDURE

4.1 Effect of Different Fusion Ratios on Results

4.1.1 Accuracy

Accuracy, i.e., accuracy rate, indicates the ratio of the number of correct judgments to the number of all judgments. The number of judgments is $(TP + TN + FP + FN)$ and the number of correct judgments is $(TP + TN)$. The specific formula 8 is shown below.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

In the formula.

TP-Number of times the judgment was positive, and the judgment was correct.

FP-The number of times the judgment is positive but incorrect.

FN - the number of times the judgment is negative, and the judgment is correct.

FN - the number of times the judgment is negative, but the judgment is wrong.

4.1.2 Recall

Recall is the recall rate, which represents the ratio of the number of correctly determined positive cases to the number of all actual positive cases. The number of all actual positive cases is $(TP + FN)$ and the number of correctly determined positive cases is TP. Specific equation 9 is shown below.

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

When the classification task is changed from binary classification to multiple classification, the macro-averaging algorithm is used to evaluate the performance of the classification model as a whole. The macro-averaging algorithm is an arithmetic mean based on the accuracy and recall of each category and F1 and is often used to evaluate the overall performance of a classification model on multiple sets of data. The formula is shown below:

$$Macro - P = \frac{1}{k} \sum_{i=1}^k P_i \quad (10)$$

$$Macro - R = \frac{1}{k} \sum_{i=1}^k R_i \quad (11)$$

$$Macro - F1 = \frac{1}{k} \sum_{i=1}^k F1_i \quad (12)$$

4.2 Results of Different Fusion Ratios In Three Views

Since the ResNet network can only deal with one mode of data, it is necessary to process the detected pictures to a certain extent before feeding them into the ResNet network. In order to improve the accuracy of the recognition system, this paper designs an image noise reduction network, the network first cuts the original picture through the Resize function, as shown in Figure. 5, and then mixes the weights through the blete function picture to meet the training requirements.

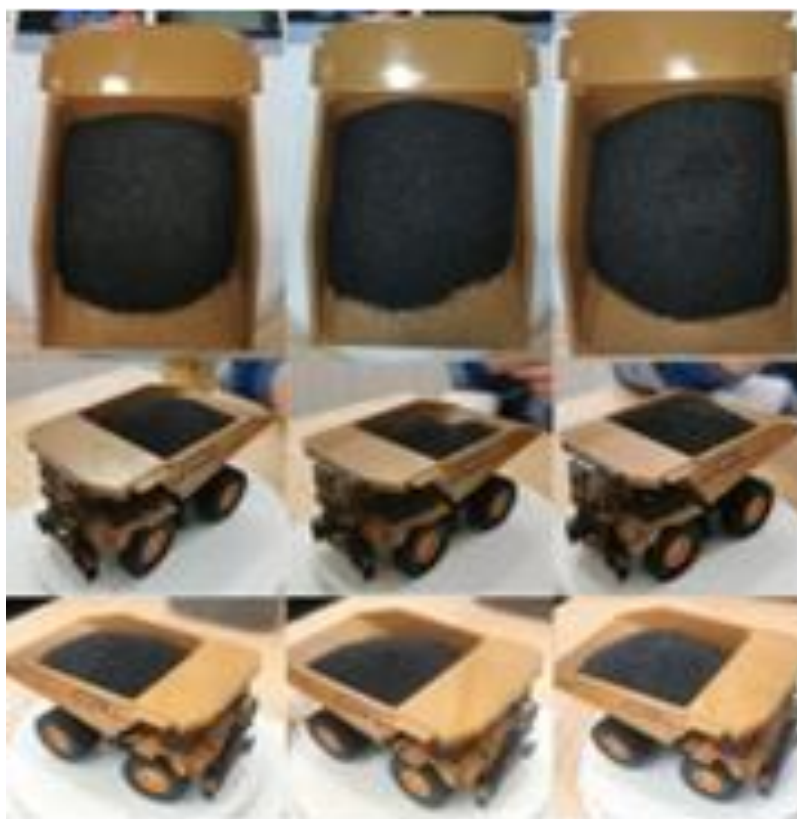


Figure 5: Image before fusion images (top view, right view, left view) (top-down weight 7:1.5:1.5,1.5:7:1.5,1.5:1.5:7,1:1:1)

Table 3 shows the results for three views with different fusion ratios. Where rate is the weight of the three views in the final fused image and order is top view, left view and right view. From the last three rows of data in the table, it can be concluded that the top view contains the most information and better results can be achieved by using only the top view

for training. However, in this task, the top view lacks information about the thickness of the mound, so the left and right views are needed to supplement the information. The final fusion ratio selected for this program is 7:1.5:1.5.

Table 3: The influence of different fusion ratio on the results

Rate	Accuracy	Macro-P	Macro-R	MacroF1
7:1.5:1.5	0.65±0.07	0.44	0.57	0.49
1.5:7:1.5	0.55±0.05	0.31	0.43	0.42
1.5:1.5:7	0.31±0.1	0.11	0.21	0.15
1:0:0	0.52±0.3	0.17	0.31	0.24
0:1:0	0.2±0.1	0.15	0.19	0.11
0:0:1	0.3±0.15	0.17	0.29	0.24

4.3 Effect of different weights on accuracy for top, left and right views

Figure. 6 shows the effect of different weight combinations and accuracy of the images during training. Although about 60% correctness can be

achieved in the end, combining the above results shows that the use of other weight combinations does not learn the true feature distribution, so only the validation set used in the training process works better, and the unlearned test set shows poorer results.

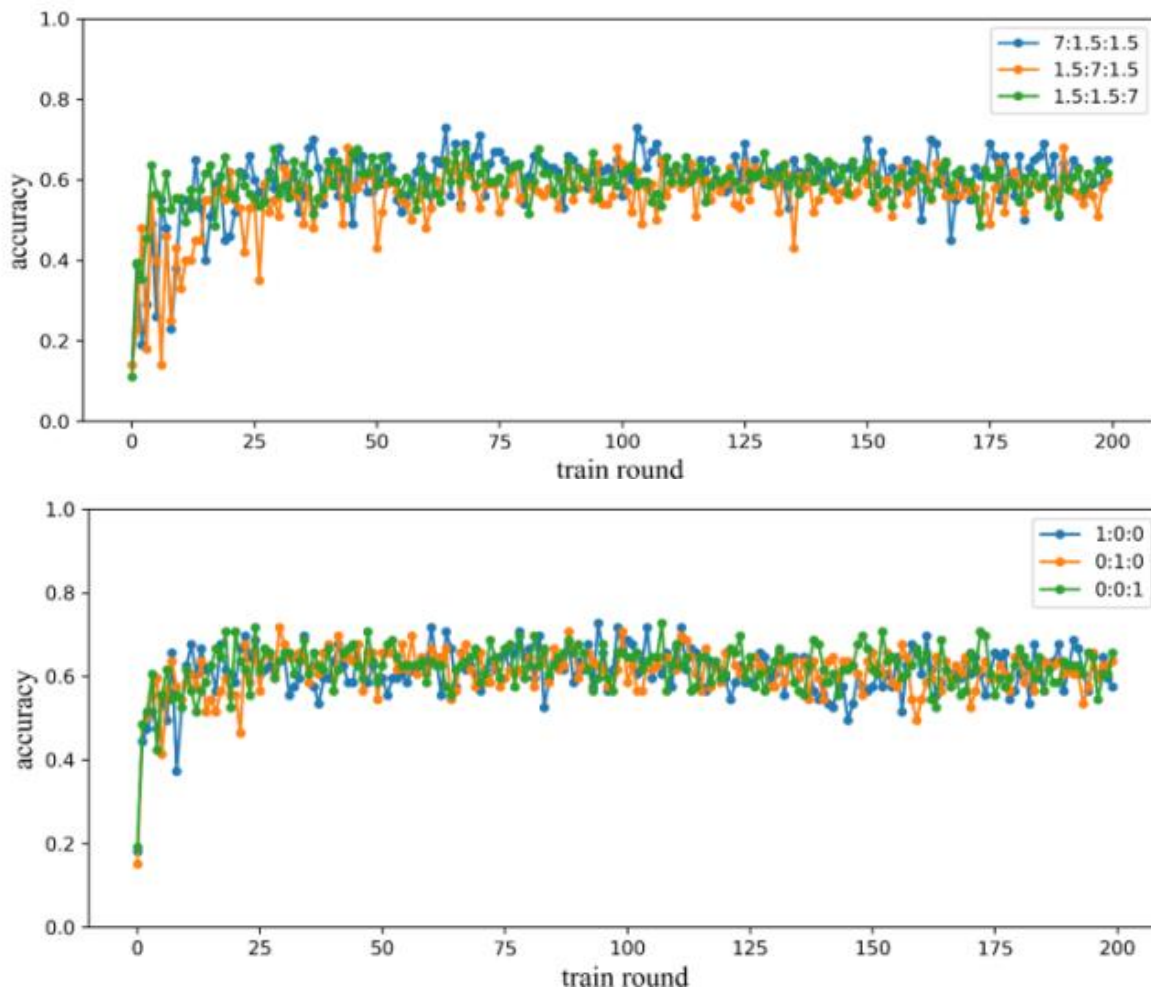


Figure 6: Influence of different weights on accuracy in fused images

5. CONCLUSIONS

In this paper, a visual weighing method for mining dump trucks based on the ResNet deep learning network model is proposed, and a cross-entropy loss function is introduced to constrain the optimization process of the network parameters to ensure the estimation accuracy of the network. In addition, this paper proposes an image processing network based on Resize function and Blend function, which utilizes convolution kernels of different sizes to extract multi-scale features while increasing the sensory field of the network to further improve the recognition performance of the network. Although this paper solves the problem of how to successfully measure the weight of the measured object to a certain extent, how to design a ResNet network with better performance to improve the accuracy of image detection is still a valuable research topic.

REFERENCES

- Rylnikova, M., Radchenko, D., and Klebanov, D., 2017. Intelligent mining engineering systems in the structure of Industry 4.0. *E3S Web of Conferences*. 21: 010-32. <https://doi.org/10.1051/e3sconf/20172101032>.
- Clausen, et al., 2022. Creating Human-centered Smart Mines of Tomorrow. *Engineering and Mining Tomorrow Journal*. 223(3):36-39. <https://doi.org/10.3390/app12188972>.
- Sunhare, P., Chowdhary, R. R., and Chattopadhyay, M. K., 2022. Internet of things and data mining: An application oriented survey. *Journal of King Saud University-Computer and Information Sciences*. 34(6):35693590. <https://doi.org/10.1016/j.jksuci.2020.07.002>.
- Shafiq M., and Gu, Z.Q., 2022. Deep Residual Learning for Image Recognition: A Survey. *Applied Sciences*. 12(18):8972-8972. <https://doi.org/10.3390/app12188972>.
- Kanai, S., Fujiwara, Y., and Iwamura, S., 2017. Preventing gradient explosions in gated recurrent units. *Advances in neural information processing systems*.
- Kag, A., Zhang, Z., and Saligrama, V., 2019. Rnns incrementally evolving on an equilibrium manifold: A panacea for vanishing and exploding gradients? *International Conference on Learning Representations*.

- Koonce B., 2021. ResNet 50. In: Convolutional Neural Networks with Swift for Tensorflow. Apress, Berkeley, CA. <https://doi.org/10.1007/978-1-4842-6168-26>.
- Wightman, R., Touvron, H., and Jégou, H., 2021. Resnet strikes back: An improved training procedure in timm. arXiv. <https://doi.org/10.48550/arXiv.2110.00476>.
- Yueteng, C., Xueyan, Z., Yandong, Z., and Fengjun, C., 2021. Recognition of plant leaf diseases and insect pests based on improved ResNet. Journal of Chinese Agricultural Mechanization. 42(12): Pp. 175-181. <https://doi.org/10.13733/j.jcam.issn.2095-5553.2021.12.26>.
- Xie, Z., Feng, Y., Hu, Y., and Liu, H., 2022. Generating image description of rice pests and diseases using a ResNet18 feature encoder. Transactions of the Chinese Society of Agricultural Engineering. 38(12): Pp. 197-206. <https://doi.org/10.11975/j.issn.1002-6819.2022.12.023>.
- Targ, S., Almeida, D., and Lyman, K., 2016. Resnet in resnet: Generalizing residual architectures. arXiv. <https://doi.org/10.48550/arXiv.1603.08029>.
- Fu, J.X., and Wang, Q., 2022. Rock recognition based on multi-scale grouped convolution ResNet34. Journal of Shaanxi University of Science and Technology. 40 (01): 167-173. <https://doi.org/10.19481/j.cnki.issn2096-398x.2022.01.025>.
- Bharati, S., Podder, P., Mondal, M., and Prasath, V. B., 2021. CO-ResNet: Optimized ResNet model for COVID-19 diagnosis from X-ray images. International Journal of Hybrid Intelligent Systems. 17(1-2): Pp. 71-85. <https://doi.org/10.3233/HIS-210008>.

