ResDAU-Net AE signal pickup

X.B.Qin*, Y.Gao and W.Tang

College of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China

* Corresponding Author Email: gaoyuanlaile2000@163.com

Abstract. In order to improve the pickup efficiency of AE signals, the sampling part was replaced with residual units based on the U-Net network structure, and the PAM and CAM dual attention modules were embedded in the jump structure to form the ResDA-UNet network. Input the AE signal waveform data, output the probability distribution of P wave, S wave and noise, and pick it up at the maximum probability point. By comparing the AR-AIC and reference threshold STA / LTA picking method, the proportion of the absolute errors reached 5ms, 15ms, 25ms and 50ms, and the proportion of ResDAU-Net network reached 93.5%. Compared with traditional ResDAU-Net methods, the picking accuracy and picking speed are higher, and the network has higher adaptability to feature extraction of acoustic emission signal waveform data.

Keywords: AE signal; initial pickup; U-Net model; neural network.

1. Introduction

With the development of society and economy, the demand for coal mine, oil and other resources is increasing day by day. Due to the exhaustion of shallow resources caused by several years of development, most coal industries have to turn the mining direction from shallow to deep layer. However, the geological inspection of underground mines cannot accurately, truly and comprehensively reflect the mining conditions, and the deep mining is easy to produce roof collapse, the safety factor of mining work is low and it is difficult, which seriously affects the efficiency of resource acquisition and threatens the safety of workers.

The illing mining method is an underground mining method, mainly used for mining deposits located in the deep ore body or with high stability demand. The basic principle is to fill the af by the filling material to maintain the structural stability of the ore body. In the working process, the ore is first extracted from the underground ore body by drilling and blasting or mechanical mining means, and then the filling materials (such as tailings, sand, cement, etc.) are prepared and transported to the goaf for filling. The illing material is evenly injected into the cavity through pipes or other equipment to provide support and prevent ore body collapse and ground subsidence. The filling goaf is stable, the safety of the mine is guaranteed, the mining work can continue, and the recovery rate of ore resources is also improved.

Loaf filling body structures to a certain extent provides the protection for deep mining, but in recent years filling body rupture and roof falling accident frequently, filling body rupture timely warning problems to be solved, filling body and external physical cracks are accompanied by the emergence of acoustic source, accurate and real-time pick up is the premise of locate the acoustic source.

2. The U-Net network structure

Convolutional neural network is one of the representative algorithms of deep learning. It is mainly composed of convolutional layer and is mostly used for image processing. The core feature of the network is weight sharing and local connection, and the most reasonable value of convolution kernel parameters is obtained through the training process feedback. Through these two core operations, the network structure can be simplified, the model convergence speed is improved, which ensures that the network can adapt to the changing model parameters, improves the training speed and accuracy, reduces the connection between the network layers, and effectively reduces the risk of overfitting[1].

The U-Net algorithm model is an extension of a convolutional neural network introduced by Olaf et al. in 2015, which is widely applied in medical image processing and other fields. The U-net network includes encoding and decoding structure. The coding part gradually reduces the resolution by extracting the top features of the image through the convolution layer and pooling layer, restores the image resolution to the original size in the decoding part, and recovers the details with jumping links. Therefore, the U-net classification of images can reach the pixel level, and the U-net has become one of the current research hotspots in the field of image processing[2].

Net mainly consists of U channel and short channel . The function of the left part is mainly feature extraction and capture of context information, while the right part is the decoding process, which mainly uses the decoding of feature map to predict pixel labels. The middle part is the short connected channel, which is mainly used to improve the model accuracy and solve the gradient vanishing. The structure diagram of the U-net network is shown[3].

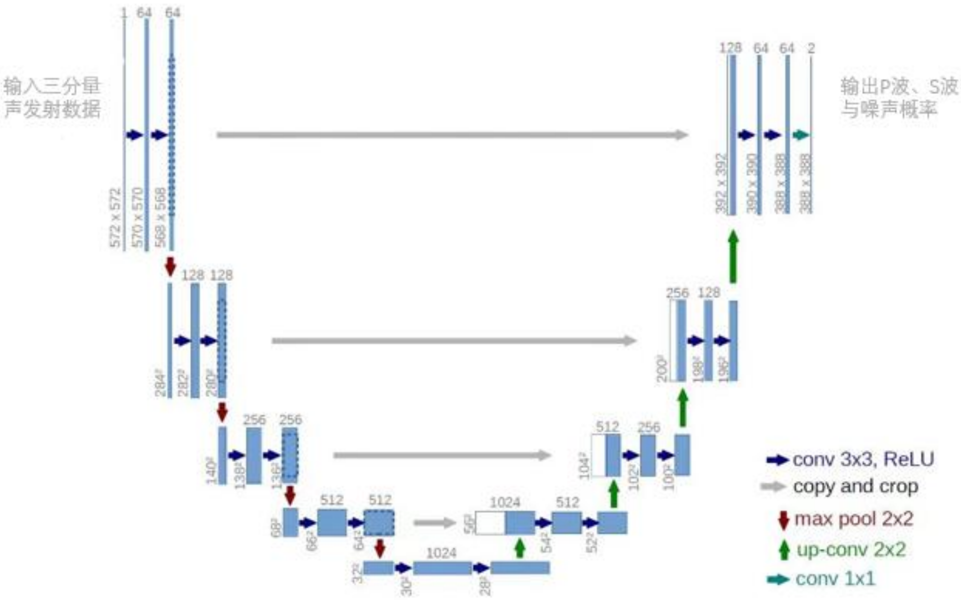

Fig. 1 The U-Net network structure

The convolution structures in the figure are all convolution kernels of 33, with filling 0 and step size 1;

$$n_{out} = \left[\frac{n_{in}+2p-k}{s}\right] + 1 \quad (1)$$

By (1):

$$n_{out} = n_{in} - 2 \quad (2)$$

According to the formula, the first to fifth convolutional layers consists of three 33 convolutions, and each time the 33 convolution size decreases by 2.

The first four convolutional layers all enter the next layer through Max-pooling. The core size of each pooling layer is k=2, filling p=0, and the step size is s=2:

$$n_{out} = n_{in}/2 \quad (3)$$

The fifth layer does not maximize the pooling, and directly transfers the obtained feature graph to the decoding part. The main function of the right part is to decode the feature graph to restore the original size, which consists of convolution, upsampling and jump structure. In the U-net network structure, the main role of skip connection is to achieve feature fusion, and this operation can effectively solve the problem of image information loss and inaccurate segmentation in the segmentation process.

The advantage of U-Net network in the field of waveform data processing is that it can automatically extract multi-scale features from complex waveforms, and combine the details of low-level time domain with high-level semantic features with jump connection, so as to realize the accurate identification and segmentation of features such as P wave and S wave in the signal. In addition, the end-to-end training method of U-Net avoids the tedious feature engineering, and can directly realize the pattern recognition and segmentation of waveform signals without manually extracting features, which greatly improves the automation level of waveform picking.

However, U-Net network also has a series of obvious limitations and disadvantages in waveform data processing. Because its core is based on convolution operation, it is difficult for U-Net to effectively capture the long-range dependence in long-time waveform data, especially in the long time interval of P wave and S wave in time identification, the model cannot accurately establish remote correlation, which affects the picking accuracy. At the same time, although the jump connection improves the retention of low-level details, it is also easy to pass the low-level noise together, resulting in the poor robustness of the network for signal processing with high noise and waveform distortion. U-Net network also lacks an adaptive global feature extraction mechanism. In the face of waveform signals with complex wave velocity change and non-uniform propagation path, it is difficult to dynamically adjust the receptive field to adapt to waveform events at different scales, which eventually leads to the instability of waveforms with large differences in signal duration and amplitude. In the face of complex scenarios such as strong noise interference, signal superposition, and multi-source aliasing, the accuracy and reliability of picking up decreased significantly, resulting in the deviation between the final positioning and identification effect. Although U-Net network has certain advantages in waveform data processing, it urgently needs to be improved for complex waveform, long time dependence and strong noise.

2.1 ResDAU-Net network

The main content of acoustic emission signal pickup for P wave and S wave early to then, due to the traditional U-Net network structure in the coding stage through continuous convolution and pooling capture context information, easy to cause the high frequency part of acoustic emission signal missing, at the same time, after experiment found that the original U-Net network structure of acoustic transmission signal pickup still exist characteristic parameters loss and semantic mismatch problems, seriously affect the pickup accuracy and efficiency. For the above problems, based on the network U-type structure, the residual learning unit instead of the sampling stage in the U-Net network, add the double attention mechanism in the jump structure (Skip), form the structure of ResDAU-Net (Residual and Dual-Attention U-Net) network structure, the advantage is that the integrity of the acoustic emission signal picking is higher, and make the low frequency signal and high frequency signal more closely, and pickup efficiency is higher.

2.1.1 2.1.1 ResU-Net network

The ResU-Net structure is shown in Fig, consisting of encoding, decoding, and hopping. In the downsampling stage, four groups of Residual block are added. The pooling operation affects the pickup accuracy of acoustic emission signal, especially the spatial and temporal resolution and feature retention of signals, and the pickup of acoustic emission signal needs to accurately locate the arrival time of the signal, so the pooling layer is not set in this model. At the same time, the convolution mode with step size of 4 is selected. In the last layer decoding stage of the network, the form of 11 convolution is adopted, and the Softmax function is used to output the probability distribution of P wave, S wave and noise[4].

$$p_i(x) = \frac{e^{z_i(x)}}{\sum_{k=1}^{3} e^{z_k(x)}} \tag{4}$$

Where: i=1,2 and 3 represent P wave, S wave and noise respectively; Z(x) is the final output parameter.

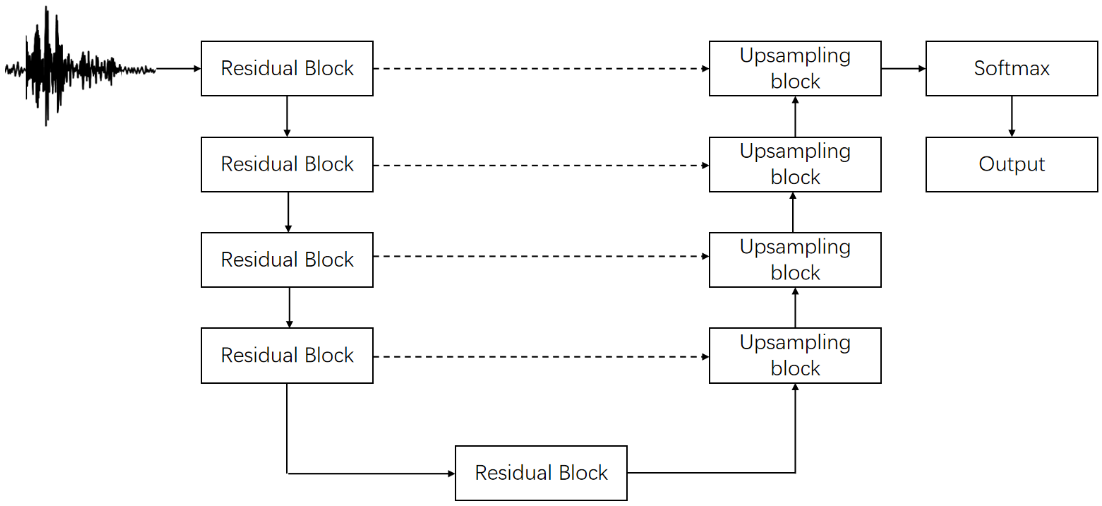

Fig. 2 The ResU-Net network structure

Residual Block (Residual block) was originally proposed in ResNet (Residual Network) in 2015. The core idea is to bypass the input and add the output to one or more convolution layers to learn the "residual" between the input and the target output. The output of the mathematical angle Residual Block can be expressed as $y = F(x) + x$, where x is the input and F (x) is the output of the convolutional layer (after a series of normalization and activation functions). The advantage is that even if the deep network is difficult to directly learn the complex mapping relationship, it can solve the problem by learning the residual functions that are easier to optimize, so as to alleviate the gradient disappearance problem and improve the training stability and feature expression ability of the network.

Upsampling Block (Upsampling block) is a basic module used in deep learning to increase the spatial resolution of feature maps. It is often used in the segmentation network (U-Net) and other tasks that need to reconstruct high-resolution output. The main goal is to recover low-resolution feature maps to higher resolution by methods such as interpolation or inverse convolution, while retaining key features. Upsampling Block Combine convolution, batch normalization and activation function to further extract fine-grained features, so as to improve the detailed expression ability of the output.

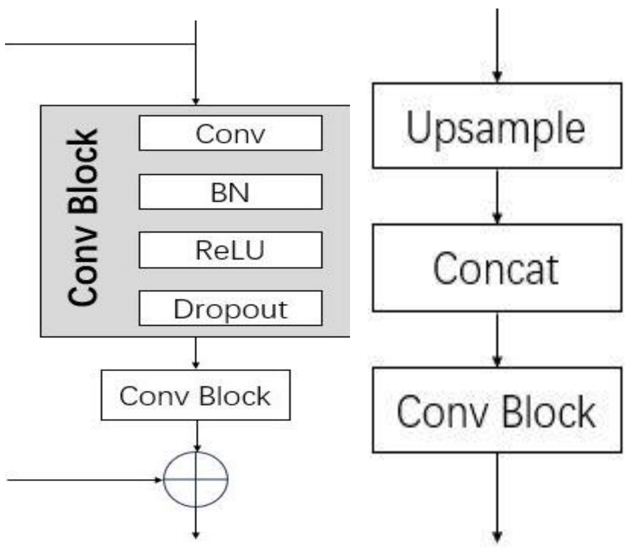

Fig. 3 The residual unit

2.1.2 2.1.2 The DA-Skip structure

Because the traditional U-net network structure is directly linked between encoding and decoding by skipping (Skip), it may lead to information redundancy. The hop connection directly passes the low-level features from the encoder to the decoder, which usually contains information about the edges, texture and so on in the image. These features are necessary to recover the image details, but at the same time may overlap with the high-level semantic features learned during the decoding stage. Redundant feature transfer leads to decoder dependence on low-level features, which affects the learning of high-level information. Redundancy not only affects the effectiveness of the model, but also may increase the computational burden during training, affecting the efficiency and convergence rate of the network. Simultaneous jump connections may pose the risk of overfitting. By directly passing low-level features to the decoder, the jump connection makes the network can directly use the details of the training data, the direct features in the training data or data quality is small, may lead to the network overfitting to the training set, makes the model generalization ability on the test set is poor, affect its performance in practical application[5]. Decoding with a direct connection to encoding may also lead to increased computational quantities. Since skip connectivity requires the splicing or merging of multiple feature maps in the encoder and decoder, this can significantly increase the computational burden of the network. Each jump connection produces an additional computational amount, requiring not only more memory to store these intermediate feature maps, but also more computing resources for the synthesis and processing of the feature maps. As network depth increases, the computational overhead of jumping connections doubles. Even when the model performance is improved, the training time and hardware resources requirements will increase accordingly.

To further solve the above problems, this paper adds DA-block dual attention module to the skip link, and the network can autonomously select the low-level features that are most helpful to the segmentation task. By adaptively adjusting the weight of each channel, the attention module can reduce the interference of irrelevant features, avoid information redundancy, and improve the representation ability of the network. At the same time, DA-block can weight the features of different scales when jumping the connection, so that the low-level features and high-level features can be more effectively integrated.

The dual attention block (DA-Block) is used as a feature extraction module that integrates the image-specific features of the location and the channel. In particular, in the environment of U-Net shape architecture, the dedicated feature extraction function of DA-Block is crucial. DA-Block performs well in both location-based and channel-based feature extraction, enabling more detailed and accurate feature sets. Therefore, it was incorporated into the Skip connection to improve the segmentation performance of the model. DA-Block consists of two main components: one with location attention module (PAM) and the other with channel attention module (CAM), both of which learn from the dual attention network for scene segmentation.

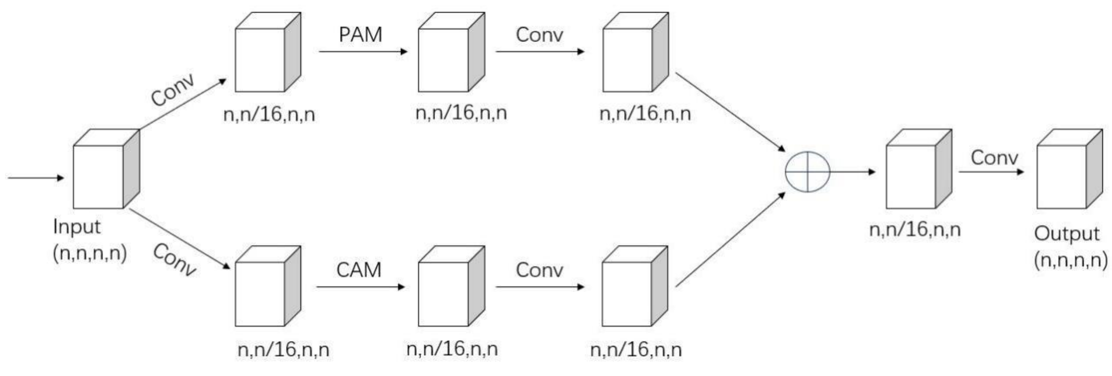

Fig. 3 Dual attention module

Skip, DA-Block in the connection optimizes the features of the encoder transmission to facilitate the decoder reconstruction of more accurate feature maps. Thus, the proposed architecture incorporates the advantages of the two basic techniques and alleviates their shortcomings, resulting in a robust system capable of image-specific feature extraction.

DA-block architecture as shown in the figure, which mainly integrates location attention module (PAM) and positional attention module (CAM), combining the extraction advantages of the two: location feature extraction and channel feature extraction ability. At the same time, when DA is combined with traditional convolution, DA-block shows relatively strong feature extraction ability. After performing a convolution calculation in PAM, the number of drop channels is scaled to 1 / 16 to get α^1, and after PAM feature extraction, the second convolution calculation gets $\hat{\alpha}^1$:

$$\alpha^1 = Conv(input) \tag{5}$$

$$\hat{\alpha}^1 = Conv(PAM(\alpha^1)) \tag{6}$$

Similar to PAM, the CAM process is as follows:

$$\alpha^2 = Conv(input) \tag{7}$$

$$\hat{\alpha}^2 = Conv(CAM(\alpha^2)) \tag{8}$$

Both perform the sum operation and a single convolution to calculate the number of recovery channels to obtain the final output:

$$output = Conv(\hat{\alpha}^1 + \hat{\alpha}^2) \tag{9}$$

2.2 The ResDAU-Net network pickup

ResDAU-Net network structure is shown in the following figure, including 9 layers of network, composed of DA-Skip, encoding, decoding, the P wave, S wave characteristic parameters can be extracted from the acoustic emission signal, decoding extraction results to the original data, DA-Skip structure combines PAM and CAM dual attention mechanism, while focusing on the spatial and semantic features, significantly reduce the redundancy and noise caused by the original Skip connection.

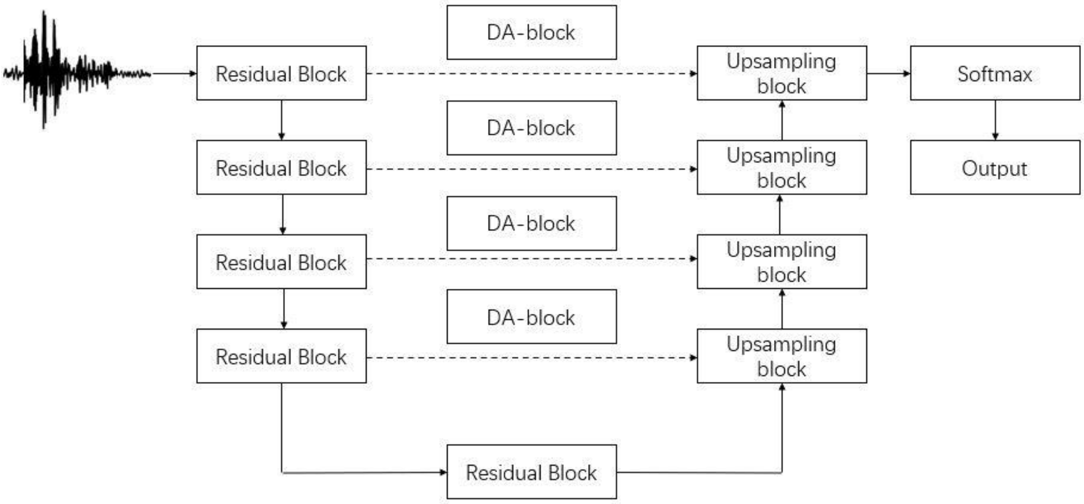

Fig. 4 The ResDAU-Net Network structure

Input the AE signal waveform data and output the probability distribution of P wave, S wave and noise. That is, when the P wave and the S wave arrive at a certain moment:

$$\mathrm{P(t)} \approx 1, \mathrm{S(t)} \approx 0, \mathrm{N(t)} \approx 0 \tag{10}$$

$$\mathrm{P(t)} \approx 0, \mathrm{S(t)} \approx 1, \mathrm{N(t)} \approx 0 \tag{11}$$

Other background noise moments:

$$P(t) \approx 0, S(t) \approx 0, N(t) \approx 1 \tag{12}$$

Finally, according to the maximum probability point of P wave and S wave.

3. Pick up the results

Using ResDAU-net acoustic emission signal to pick up, to the original signal classification, except for the initial wave P wave and S wave called background, with "1", P wave with "2", "S wave represented by" 3 ", initial wave picking problem is actually three classification problem, in the experiment, the input for known P wave S wave to three component acoustic transmission signal, output for the probability distribution of P wave, S wave and noise. AR-AIC algorithm, long and short window energy ratio method based on reference threshold, ResDAU-net initial pickup method comparison;

Table 1. Pick up algorithm comparison

Pick type	Data	seismic phase	Absolute error mean value/10-2	The proportion of data quantity from beginning to time when the absolute error meets the following range%			
				≤5ms	≤15ms	≤25ms	≤50ms
AR-AIC	Training (35700)	P波	2.48	32.1	51.6	65.3	85.6
		S波	2.85	30.8	49.7	69.1	83.0
	Test (8505)	P波	2.52	35.5	50.8	62.8	81.6
		S波	2.88	36.0	48.8	59.2	82.5
STA/LTA	Training (35700)	P波	2.36	34.3	55.0	69.2	87.5
		S波	2.55	32.8	52.8	67.4	86.0
	Test (8505)	P波	2.60	36.2	53.3	67.0	85.1
		S波	2.73	35.7	51.7	63.8	84.6
ResDAU-net	Training (35700)	P波	1.34	52.0	64.1	83.1	93.5
		S波	1.48	51.4	62.0	76.5	91.4
	Test (8505)	P波	1.59	51.6	61.3	80.2	91.5
		S波	1.67	50.3	60.8	75.6	90.3

According to Table 1, compared with the ResDAU-Net network, when the absolute error of 5ms, the amount of data is less than 40%, while the ResDAU-Net network is almost more than 50%, and the absolute error of the ResDAU-Net network is over 90%. The results show that the accuracy of the pickup method is high.

4. Summary

This paper introduces the process of ResDAU-Net network structure. First, based on the U-Net network structure, the sampling module is replaced with the residual learning unit, and the CAM and PAM dual attention mechanism is embedded in the Skip structure. By comparing AR-AIC, the reference threshold STA / LTA and the picking method, the algorithm is verified with good integrity and high accuracy.

References

[1] He Bin, Zhou Yunyao, Lu Yongqing. Seismic primary-to-wave picking algorithm combining U-net and FPN [J]. Mapping and geographic information, 2024,49(01):82-87.DOI:10.14188/j.2095-6045.2022180.

[2] Couper K N, Blount D G, Hafalla J C, et al. Macrophage-Mediated but Gamma Interferon-Independent Innate Immune Responses Control the Primary Wave of Plasmodium Yoelii Parasitemia[J]. Infection and Immunity, 2007, 75(12):5 806-5 818

[3] Hu Ting, Xu Bin, Wang Yongfa, et al. Start-up study based on U-Net [J]. Science, Technology and Engineering, 2023,23 (16): 6802-6809.

[4] Si Wenxue, Li Chao. Start-to time pickup analysis of seismic waves based on an improved U-Net model [J]. Microcomputer applications, 2024,40 (07): 1-4.

[5] Cha Min. Research on bus passenger flow prediction method based on Spark platform [D]. University of Electronic Science and Technology, 2017.