MICROSEISMIC EVENT DETECTION BASED ON MULTISCALE DETECTION CONVOLUTIONAL NEURAL NETWORK

YAN ZHANG[1,2,3], XIAO-QIU LIU[1,2,3], LI-WEI SONG[1,2,3] and HONG-LI DONG[4,5]

[1] School of Physics and Electronic Engineering, Northeast Petroleum University, Da Qing, Heilongjiang 163318, P.R. China. zhangyan1999@nepu.edu.cn
[2] Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Da Qing, Heilongjiang 163318, P.R. China.
[3] Key Laboratory of Networking and Intelligent Control of Heilongjiang Province, Da Qing, Heilongjiang 163000, P.R. China.
[4] School of Physics and Electronic Engineering, Northeast Petroleum University, Da Qing, Heilongjiang 163318; P.R. China.
[5] Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Da Qing, Heilongjiang 163318, P.R. China.

(Received August 5, 2023; accepted September 15, 2023)

ABSTRACT

Zhang, Y., Liu, X.Q., Song, L.W. and Dong, H.L., 2023. Microseismic event detection based on multiscale detection convolutional neural network. *Journal of Seismic Exploration*, 32: 455-477.

The traditional microseismic event detection method is mainly based on the characteristic calculation of microseismic signals. Its accuracy is greatly affected by the empirical parameter setting of the algorithm, characteristics selection of signal, and signal-to-noise ratio of microseismic signals. Furthermore, it also takes a long computation time when dealing with massive microseismic data. Therefore, this paper presents a method of microseismic event detection based on the multiscale neural network. Firstly, according to the characteristics of microseismic signals, one-dimensional convolutional neural network is built to extract the fine-grained features of the shallow layers and the semantic features of the deep layers. Then, the credibility factor model is established for the detection results of the different scale feature expressions, and the final recognition results are obtained by uncertainty reasoning. Compared with wavelet analysis, BP neural network, and traditional convolution neural network, the experimental results show that the proposed model is superior to other methods, and has better anti-noise and generalization ability. In addition, this method also provides a new strategy for processing other monitoring signals with large interference.

KEY WORDS: microseismic event detection, neural network, certainty factor, multiscale.

0963-0651/23/$5.00

INTRODUCTION

Microseismic monitoring technology is a method to reveal the geometry and spatial distribution of underground hydraulic fractures, evaluate reservoir reconstruction measures, and optimize post-treatment, by identifying and locating the microseismic caused by rock fracture in the process of hydraulic fracturing. As indispensable processing means for microseismic monitoring, microseismic event identification is of great significance for focal position monitoring, focal mechanism analysis, and fracture interpretation. In recent years, microseismic monitoring technology has been widely used in oil and gas field exploration and development, coal mine safety, rock fracture monitoring, and other fields. For example, dynamic detection of oil and gas reservoir (Li et al., 2016), monitoring of hydraulic fracturing in oil field (Ge et al., 2005), front tracking of fluid driving in oil field industrial production (Maxwell et al, , 2010), water inrush from mine, landslide, tunnel safety monitoring (Sun et al., 2012), etc.

As an important basis of microseismic data processing, microseismic event detection plays a very significant role in subsequent microseismic source location (Xie, 2015). At present, the detection of microseismic events mainly refers to the traditional seismic detection methods, including STA/LTA (Yi, 1988), Akaike information criterion (AIC) (Sleeman and van Eck, 1999), and wavelet analysis (Holschneider, 2008), etc. Many scholars have proposed many improvement methods on this basis. Zhang (2003) developed an automatic P-wave arrival detection and picking algorithm based on the wavelet transform and AIC picker. In each time window, they evaluated the consistency of threshold absolute wavelet coefficients at different scales by using AIC automatic classifier, to determine whether P-wave arrival is detected. Shi-Cha et al. (2013) used LTA/STA method, combined with polarization analysis, and designed the polar-energy ratio method to detect microseismic events based on the characteristics of multi-channel signals and noise. Sheng (2019) used wavelet transform to separate the effective signal from the noise, and then carried on the high-order statistics to the effective signal. The first arrival time picking precision of low signal-to-noise ratio (SNR) data is effectively improved by analyzing the abnormal points of signals with characteristic curves. Effectively improve the first arrival time pick-up accuracy of low SNR data. Traditional algorithms have limited recognition accuracy and are sensitive to noise. To improve the efficiency and accuracy of data processing, it is necessary to study algorithms suitable for the automatic identification of effective microseismic events with massive data.

Deep learning algorithms are particularly appropriate for processing large amounts of data and can extract feature maps layer by layer to achieve automatic recognition. Velis et al. (2015) used pattern recognition technology to detect waveform from microseismic data and used rank reduction filtering methods to improve the SNR. First, pattern recognition is used to seek plausible hyperbolic phase arrivals, and then the identified event is denoised and reconstructed, which improves the reliability of

single-wave arrival detection. Shang (2017) established discrimination between microseismic events and blasting based on different source parameters, using principal component analysis and the artificial neural network. While these source parameter-based methods have achieved excellent performance in microseismic events, these techniques always require post-processing by experienced analysts. Xu et al. (2021) proposes the INN model that combined the genetic algorithm and neural network, it used to identify rock blasting vibration waveform in metal mines. Six features are selected as the input of the waveform classification model, and the genetic algorithm is used to optimize the number of nodes in the hidden layer. The trained model can effectively identify microseismic waveform. Such practices often lead to incomplete identification results because of the neglect of low-scale signal features. The above-mentioned deep learning methods have certain effects, but there are still some problems: (1) The difficulty of data acquisition and the small number of effective signals in microseismic monitoring lead to the limitation of multi-trace property design algorithms. (2) When using deep learning for microseismic event detection, it often enhances the extraction of semantic features by deepening the network layers because microseismic event energy is weak. Such practices often lead to incomplete detection results because they ignore the features of low-scale.

Multiscale feature analysis was first proposed in the field of image processing, and commonly used methods to construct multiscale features are as follows: (1) Building image pyramids based on image resolution (Singh, 2018); (2) Based on different levels of feature maps within neural networks feature pyramid (Cai, 2016; Liu., 2016); (3) Space pyramid based on different receptive fields (Zhao, 2017). The fusion methods of multiscale feature maps include (1) Early fusion: First fusion of multiscale feature maps, and train predictors on the fused feature maps. For example, VGGNet (Zhang, 2017), R-FCN (Dai, 2016), ParseNet (Liu, 2015), etc. (2) Late fusion: Improve recognition performance by combining detection results of different scales. For example, SSD (Liu, 2016), YOLO2 (Redmon and Farhadi, 2017) YOLO9000: better, faster, stronger, YOLO3 (Redmon and Farhadi, 2018), etc. The features of microseismic signals at different scales

have different meanings, synthetic multiscale features for detection will improve the accuracy of the results.

Given the above analysis, we propose a multi-scale detection convolutional neural network (MSD-CNN) for microseismic event detection technology, builds a convolutional neural network to extract microseismic data characteristics to identify and introduces the Certainty Factor model (Certainty-Factor, C-F) into the neural network; the fine-grained information and semantic information of the microseismic signal obtained by the scale feature map is used for credibility modeling; the final result is obtained through the uncertainty algorithm of the detection results on multiple scales, and the constructed model is more suitable for the processing of microseismic data, accurately identify diverse microseismic data without deepening the network. In the process of model processing, the sliding

window size and step size are used to segment the data to form the sample input model, and the convolutional neural network is used to extract the features of different scales of microseismic events, learn the credibility of each scale and the dynamic strength of knowledge, and avoid manual design. The dynamic strength on each scale is obtained by the detector on that scale, and the detection result is integrated with the uncertainty of the trust degree through the credibility model, and the recognition result obtained has a higher degree of confidence. Experimental results prove that the proposed method has higher efficiency and can better automatically identify microseismic signals.

Compared to what has been done in microseismic detection in the past, in this study: (1) Dividing samples through a sliding window not only increases the diversity of samples but also avoids the damage caused by microseismic events due to data segmentation. (2) The MSD-CNN we built pays attention to the multiscale features of microseismic signals. (3) In the multiscale detection, combined with the C-F model in the uncertainty reasoning theory, the uncertainty of the neural network is noticed

THEORY

Multiscale features

The detector is usually connected to the last layer of the convolutional neural network (CNN) for prediction in the traditional microseismic recognition network structure. Deep feature maps cannot provide fine-grained information. Therefore, multiscale features expression is sought as an effective way to solve fine-grained loss. The feature maps with different depths within the CNN form multiscale expressions of the feature map because of the layered structure of the network, and the deeper the feature map, the larger the receptive field. SSD algorithm and MS-CNN algorithm proposed the idea of directly detecting targets on feature maps of different

scales, and finally integrating them. The shallow feature map is responsible for detecting and recognizing detailed targets, and the deep feature map is responsible for detecting subject targets.

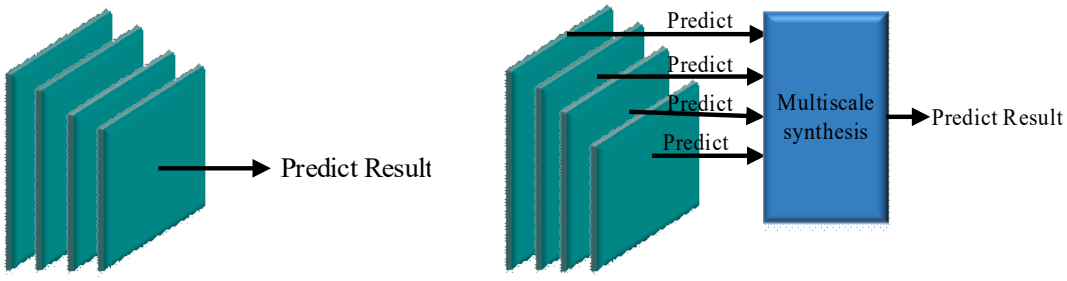
Since the surrounding environment of fracturing operations has a great influence on the waveform characteristics of microseismic events, it is not comprehensive to perform feature detection of microseismic waveform on only one scale. Observations and judgments on multiple scales should be integrated. As shown in Fig. 1a, the traditional network makes predictions in the last layer after obtaining high-scale feature maps. As shown in Fig. 1b, the multiscale detection network can not only obtain feature maps of different scales but also predict results from feature maps of different scales. When the feature maps of different scales are assigned weights for weighted fusion processing, the overall loss function E we define:

$$E = \sum_{l=1}^{L} \omega^{l} E^{l}(\alpha^{l}) \quad , \tag{1}$$

where l is a number of different scales; ω^{l} is the weighting factor of the scale l; $E^{l}(\alpha^{l})$ is the loss of scale l; α^{l} is the feature map of scale l. The scale loss is affected by the scale feature map, which is related to the corresponding convolution kernel and bias. Using the expression weighting method, formula (1) can be rewritten as:

$$E = \sum_{l=1}^{L} f(E^{l}(\alpha^{l}, K^{l}, b^{l})) = \sum_{l=1}^{L} f(\alpha^{l}) = \sum_{l=1}^{L} f(K^{l}, b^{l}) \quad , \tag{2}$$

where K^{l} is the convolution kernel of the scale l; b^{l} is the bias term of the scale l.



(a) The traditional method (b) The multiscale detection method

Fig. 1. Method of prediction.

Certainty-Factor model

Neural networks usually judge the prediction results by thresholding the confidence, and the confidence score is output during classification. This style of discriminant methods is often not conducive to the uncertainty of knowledge expressed on the network. The C-F model is an effective uncertainty reasoning method. It combines the confidence scores of various scales and obtained the fusion results through uncertainty inference, which has the advantages of intuitive, simplicity, and good effect. The neural network combined with the C-F model takes the detection results of each scale as evidence and the final recognition result as the conclusion. The relationship between the detection result of each scale and the final result is knowledge. The uncertainty of knowledge can be expressed as a production rule:

$$if\ E_i\ then\ H\ (CF(H,E_i)),\ i = 1,2,3\ldots n. \tag{3}$$

where $CF(H,E_i)$ is the credibility factor (static strength) that indicates the strength of knowledge, that is, the degree of influence on H when the evidence corresponding to $\boldsymbol{E}_i$ is true. The meaning of the above production formula is: when a microseismic event is detected at scale i, the credibility of a microseismic event is $CF(H,E_i)$. The appearance of relevant evidence increases the credibility of the conclusion that H is true, then $CF(H,E_i)>0$, which means that the emergence of evidence supports the fact that H is true and will increase the value of $CF(H,E_i)$; otherwise, $CF(H,E_i)<0$ means that the appearance of evidence supports that H is false, and the value of $CF(H,E_i)$ will become smaller; $CF(H,E_i) = 0$ means that the appearance of evidence has nothing to do with the conclusion. The static strength is obtained by the prior probability $P(H)$ and the posterior probability $P(H|E_i)$, but in practical applications, the prior probability and the posterior probability are difficult to obtain. In general, the static strength is directly given by the expert. The principle is that the more corresponding evidence appears to support H as true, the larger $CF(H,E_i)$ should be; otherwise, the smaller.

The model transfers the uncertainty of knowledge and evidence to the conclusion through step-by-step reasoning. Conclusion The credibility of H from the evidence inference corresponding to E_i is expressed as $CF_i(H)$, which is calculated by the following formula:

$$CF_i(H) = CF(H,E_i) \times MAX\{0, CF(E_i)\}\ , \tag{4}$$

where $CF_i(H)$ is the dynamic strength. it represents the current degree of uncertainty of the evidence E_i.

The credibility of the microseismic event detected on each scale is calculated respectively, and then the credibility of the final conclusion is

synthesized by the uncertainty algorithm of the conclusion. The credibility of the final conclusion combines the uncertainty of knowledge on multiple scales.$CF_{i,j}(H)$ is used to represent the credibility formed by the combined effects of E_i and E_j. Uncertainty algorithm of conclusions is defined as follows:

when $CF_i(H)<0$, $CF_j(H)<0$:

$$CF_{i,j}(H) = CF_i(H) + CF_i(H) + CF_i(H)CF_j(H) \tag{5}$$

when $CF_i(H)\geq 0$, $CF_j(H)\geq 0$:

$$CF_{i,j}(H) = CF_i(H) + CF_j(H) - CF_i(H)CF_j(H) \tag{6}$$

when $CF_i(H)$ and $CF_j(H)$ have different signs

$$CF_{i,j}(H) = \frac{CF_i(H) + CF_j(H)}{1\text{-}\min\{|CF_i(H)|,|CF_j(H)|\}} \quad . \tag{7}$$

CONSTRUCTION OF MULTISCALE DETECTION NEURAL NETWORK

Model working principle

Forward propagation of multiscale detection network

During the convolution operation, the feature map of the previous layer is convolved through a learnable convolution kernel, and then the output feature map of this layer can be obtained through the normalization function and the activation function. The output of the normalization operation is defined as follows:

$$\boldsymbol{a}^{l} = BN(\boldsymbol{a}^{l-1}) = \frac{\boldsymbol{a}^{l-1} - E(\boldsymbol{a}^{l-1})}{\sqrt{\mathrm{var}(\boldsymbol{a}^{l-1}) + \varepsilon}} \quad , \tag{8}$$

where $\boldsymbol{a}^{l-1}$ is the output of layer $l-1$, which is the input of layer l; $E(\boldsymbol{a}^{l-1})$ is the mean of input; $\mathrm{var}(\boldsymbol{a}^{l-1})$ is the variance of the input; ε is a random minimal positive value to avoid the special case where the denominator is zero; $BN(\bullet)$ is batch normalization function, Therefore, the output of the convolution operation after normalization is defined as follows:

$$\boldsymbol{a}^{l} = BN(\sigma(\boldsymbol{a}^{l-1} * \boldsymbol{K}^{l} + \boldsymbol{b}^{l})) \quad , \tag{9}$$

where $\sigma(\bullet)$ is the nonlinear activation function; There are many definitions of $\sigma(\bullet)$, generally *softmax* function, *sigmoid* function or *tanh* function.

Generation of down-sampling pooling layer of input feature map. The number of feature maps remains unchanged, and the size becomes smaller after down sampling. The output of the pooling operation is defined as follows:

$$\alpha_j^l = pool(\quad^{l-1}) \quad , \tag{10}$$

where $pool(\bullet)$ is the down sampling function, generally, the down sampling function has methods such as taking the average value, maximum value of the data in the sampling window. The error of feature extraction mainly comes from two aspects: (1) Due to the limitation of the size of the neighborhood, the estimated variance increases; (2) The parameter error of the convolutional layer will also cause the mean shift. Average pooling emphasizes the sampling of the overall feature information, which can reduce the first error; the maximum pooling selects better classified features and retains more texture information, which can reduce the second error (Boureau, 2010). In this paper, to better extract attributes for accurate identification, we choose maximum pooling.

For the fully connected layer, the operation output of the fully connected layer is defined as:

$$\boldsymbol{a}^l = \sigma(\boldsymbol{z}^l) = \sigma(\boldsymbol{a}^{l-1}\boldsymbol{K}^l + \boldsymbol{b}^l) \quad . \tag{11}$$

Considering the layered structure of the convolutional neural network, the deeper the feature map, the larger the receptive field, so the feature maps of different depths in the network form a multiscale expression. The neural network learns the credibility of each scale and the static strength of the detection result. The dynamic strength $\boldsymbol{y}^l$ of scale l is:

$$\boldsymbol{y}^l = sigmoid(\boldsymbol{a}^{L-1}\boldsymbol{K}^L + \boldsymbol{b}^L) \tag{12}$$

where $sigmoid(\bullet)$ is the activate function of detectors; $\boldsymbol{K}^L$ is the convolution kernel of the final layer of detectors; $\boldsymbol{b}^L$ is the bias of the final layer of the detectors.

The C-F model is equivalent to the non-linear weighting of the detection results of the feature map of each scale. Credibility is calculated as:

$$\boldsymbol{u}^l = H(\boldsymbol{y}^l, \omega^l) = \omega^l \times \max(0, \boldsymbol{y}^l) \quad , \tag{13}$$

where $\boldsymbol{u}^l$ is credibility of scale l; $H(\bullet)$ is credibility reasoning function; ω^l is static strength of scale l. The synthesis result of scale i, j is:

$$\begin{aligned} \boldsymbol{u}^{i,j} &= \boldsymbol{u}^{i} + \boldsymbol{u}^{j} - \boldsymbol{u}^{i} \cdot \boldsymbol{u}^{j} \\ &= H(\boldsymbol{y}^{i}, \omega^{i}) + H(\boldsymbol{y}^{j}, \omega^{j}) - H(\boldsymbol{y}^{i}, \omega^{i}) H(\boldsymbol{y}^{j}, \omega^{j}) \end{aligned} \tag{14}$$

Back propagation of multiscale detection network

In the process of neural network training, the parameters of the convolution kernel are adjusted with loss as the objective of the optimization. The binary cross entropy (BCE) loss function is used in this paper, which is defined as follows:

$$E = -\frac{1}{C}\sum_{i=1}^{C} y_i^* \ln(y_i) + (1 - y_i^*) \ln(1 - y_i) \quad , \tag{15}$$

where C is the number of samples; y_i^* is a label of the sample i; $\boldsymbol{y}_i$ is a predicted value of the sample i. The model parameters are updated by the gradient descent algorithm, and the loss information is fed back through back propagation. The back propagation error can be regarded as the sensitivity δ of the basis of each neuron. Sensitivity δ in the C-F model is defined as follows:

$$\begin{aligned} \delta &= \frac{\partial E}{\partial \boldsymbol{z}^{l}} = \frac{\partial E}{\partial y} \times \frac{\partial y}{\partial y^{l}} \times \frac{\partial y^{l}}{\partial \boldsymbol{z}^{l}} \\ &= \frac{\partial E}{\partial y} \times \frac{\partial y}{\partial y^{l}} \times y^{l} \times (1 - y^{l}) \end{aligned} \tag{16}$$

The sensitivity δ^{L} of the final detector layer is defined as:

$$\boldsymbol{\delta}_{\boldsymbol{z}}^{L} = \frac{\partial E}{\partial \boldsymbol{y}^{l}} \frac{\partial \boldsymbol{y}^{l}}{\partial \boldsymbol{z}^{L}} = \delta \, \mathrm{e} \, \sigma'(\,^{L}) \quad , \tag{17}$$

where e is the Hadamard product.

According to the sensitivity of layer $l+1$, the expression of the sensitivity of layer l is:

$$\delta^{l} = \frac{\partial E}{\partial \boldsymbol{z}^{l}} = \frac{\partial E}{\partial \boldsymbol{z}^{l+1}} \frac{\partial \boldsymbol{z}^{l+1}}{\partial \boldsymbol{z}^{l}} = \delta^{l+1} \frac{\partial \boldsymbol{z}^{l+1}}{\partial \boldsymbol{z}^{l}} \quad . \tag{18}$$

When the backpropagation, the gradient calculations of different

layers are different. If the layer $l+1$ is fully-connected, the sensitivity of the layer l is:

$$\delta^{l}=(\boldsymbol{K}^{l+1})^{T}\delta^{l+1}\circ\sigma'(z^{l}) \tag{19}$$

If the layer $l+1$ is a pooling layer, the sensitivity of the layer l is:

$$\delta^{l}=upsample(\delta^{l+1})\circ\sigma'(z^{l}) \quad , \tag{20}$$

where *upsample* function completes the magnification and error distribution logic of the pooled error matrix.

If the layer $l+1$ is a convolutional layer, it passes through convolution, activation function, and batch normalization layers in the forward propagation. Then the sensitivity of the layer l is:

$$\begin{aligned}\delta^{l} = &(1/\sqrt{\mathrm{var}(\boldsymbol{a}^{l-1})+\varepsilon}\ )\times\delta^{l+1}* \\ &rat180(\boldsymbol{K}^{l+1})\circ\sigma'(z^{l})\end{aligned} \quad , \tag{21}$$

where $rat180(\bullet)$ means rotating the loss matrix 180 degrees. Rotate the feature map and calculate the cross-correlation matrix firstly, then inverts the output. In this way, convolution kernels used for convolution in forwarding propagation can be obtained.

After determining the sensitivity of each layer, the gradient of the loss function to the convolution kernel K and the bias $\boldsymbol{b}$ is:

$$\frac{\partial E}{\partial \boldsymbol{K}^{l}}=\frac{\partial E}{\partial z^{l}}\frac{\partial z^{l}}{\partial \boldsymbol{K}^{l}}=\delta^{l}\boldsymbol{a}^{l-1} \tag{22}$$

$$\frac{\partial E}{\partial \boldsymbol{b}^{l}}=\frac{\partial E}{\partial z^{l}}\frac{\partial z^{l}}{\partial \boldsymbol{b}^{l}}=\delta^{l} \tag{23}$$

When the learning rate is η, the parameters of each layer of a neural network are updated as follows:

$$\boldsymbol{K}^{l}=\boldsymbol{K}^{l}-\eta\,\boldsymbol{a}^{l}\boldsymbol{\delta}^{1\ \ l} \tag{24}$$

$$\boldsymbol{b}^{l}=\boldsymbol{b}^{l}\boldsymbol{\delta}-\eta\sum^{\ \ l} \quad . \tag{25}$$

Network architecture

Fig. 2 shows the MSD-CNN architecture that predicts the noise or microseismic events. There are 6 detection scales in the model, and each

detection scale is composed of convolution, activation function, and batch normalization. After each scale is connected to the corresponding detector that is composed of three pooling layers and three fully connected layers alternately. The result is input into the C-F model by the detector to make a decision. With the deepening of a convolutional neural network, the size of the convolution kernel is gradually reduced, and the fine-grain multiscale expression is improved. The three maximum pooled layers of the detector have a window of 2 × 1 and a step of 2, and the output of the fully connected layer is 1/2 of the input.

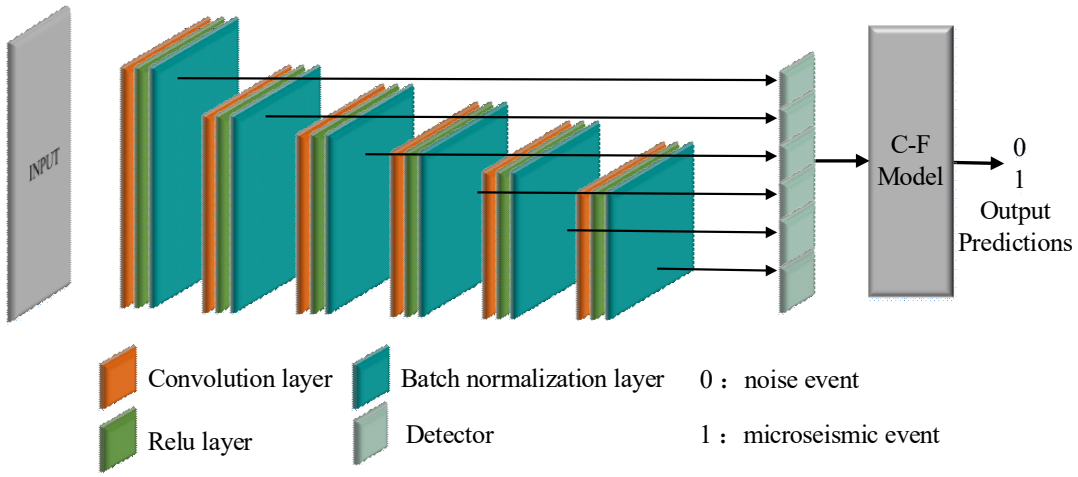


Fig. 2. MSD-CNN architecture for microseismic events detection.

4

The multiscale network detection we propose is based on the independent recognition of multiscale feature maps, and the final detection results of samples are supervised trained through labels. The detailed network data processing is shown in Fig. 3. The original data is preprocessed to obtain samples and labels. The preprocessed samples are extracted through a multiscale network for multiscale feature maps. Then, multiscale feature maps are sent to detectors for detection, and the accuracy of each scale is calculated. The confidence level detected by each scale is used as the dynamic strength and the accuracy as the static strength into the credibility model. The final composite result is calculated with the label calculation loss

error and various evaluation indicators.

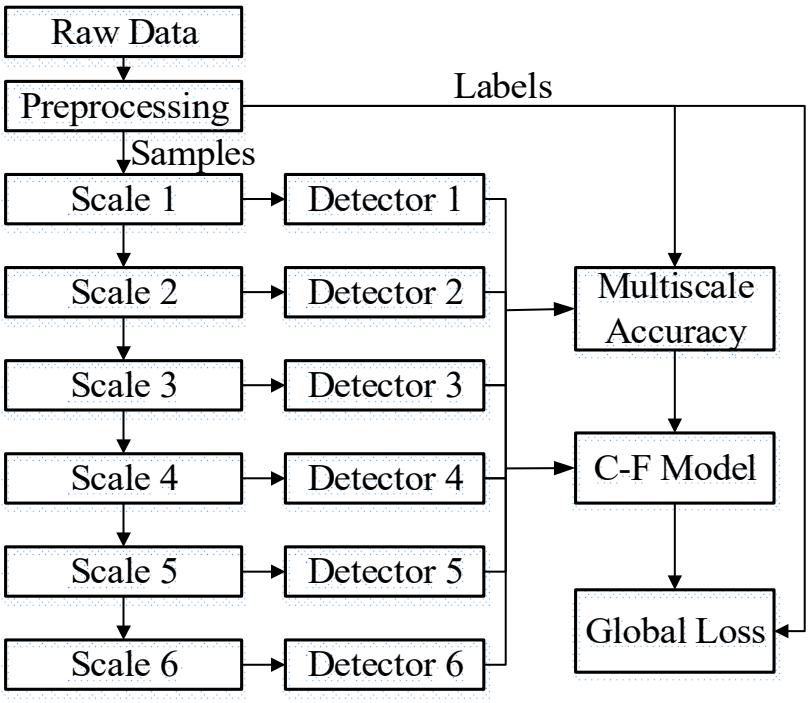


Fig. 3. MSD-CNN data flow.

The network constructs a global loss function and obtains the accuracy of the detection results on the shallow fine-grained features and the deep semantic features through verification sets. The accuracy is defined as follows:

$$ACC = (TP + TN) / (TP + TN + FP + FN) \quad , \tag{26}$$

where TP and FP represent true positives; false positives, respectively; FN represents false negatives; TN represents true negatives.

The static strength $\omega^{l} (l = 1,2,3\cdots 6.)$ is set to 1 on each scale when the model is initialized, and is updated to the accuracy of the detection on this scale after each round of verification. The C-F model has 6 uncertain pieces of knowledge: if the detection contains micro-seismic waveform at scale i, then the sample does contain microseismic waveforms at that scale with the credibility of $\omega^{l} (l = 1,2,3\cdots 6.)$. In the credibility calculation process of the conclusion H, the premise condition E_i is to detect the dynamic intensity of the microseismic waveform at the scale i, which is learned by the neural network detector. Finally, the conclusion uncertainty algorithm is used to combine two by two to finally obtain the credibility formed by the comprehensive influence of the six pieces of evidence on the final conclusion H. If the comprehensive credibility exceeds 0.5, the sample is judged to contain a microseismic event.

Model Training

Data preprocessing

The experimental data include synthetic data and field data. The theoretical underground medium model is used in the synthetic experiment to simulate the detected microseismic data. The field data is the monitoring data of four fracturing wells in a certain area. Since the amplitude of the original waveform data fluctuates relatively large, normalization is performed before the data is input to the network. The sliding window method is used to amplify the amount of data and solve the problem of sample balance in the training process to meet the requirements of the neural network for samples. As shown in Fig. 4, the samples use one-hot encoding, the positive samples containing microseismic events are marked as 1, and the negative samples containing only background noise are marked as 0.

If the sliding window length is a, the step length is s, and the total data length is L, then the total number n of samples obtained is:

$$n = (L - a)/s \tag{27}$$

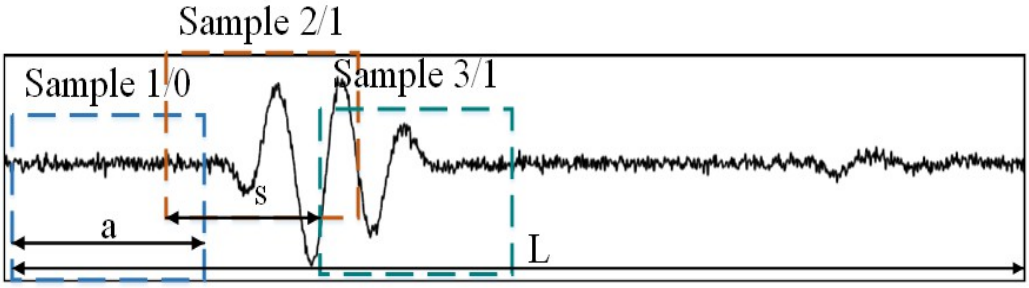


Fig. 4. Sliding window processing of micro-seismic data.

Most of the microseismic data received by the geophone are background noise, and the effective waveform only occupies a small part. Some negative samples are randomly removed to make the number of positive and negative samples equal. The single trace microseismic signal

sampling point in the synthetic data is 10000, the sliding window length is set to 100, and the step size is set to 20. In this way, 9995 samples of data can be obtained, including 245 samples containing microseismic events. The amount of data is increased by 4.9 times compared to the method of simply setting the length of 100 to divide the data.

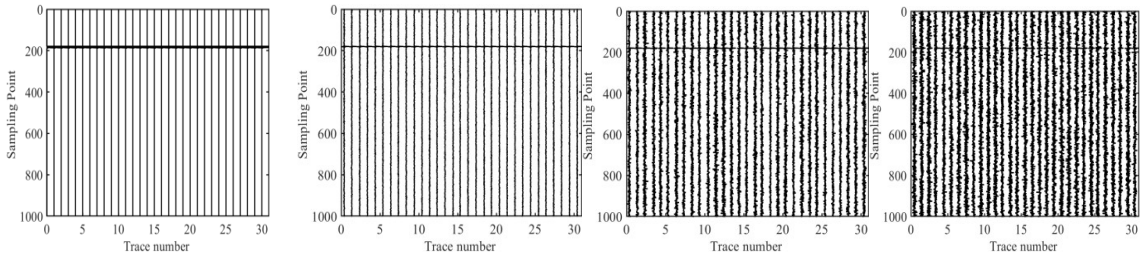
We simulated synthetic microseismic monitoring records, synthesized

the records using ricker wavelet convolution time series, and added different degrees of random noise to match the low SNR characteristics of the field data. The SNR of synthesized records are calculated by the following equation:

$$SNR = 10 \times \log_{10}\left(\| X \|_F / \| N \|_F\right) \tag{28}$$

where X is one trace data; N is the random noise; $\|\cdot\|_F$ is the Frobenius norm.

Figs. 5a-5d show the synthesized signals with different SNR. As the SNR decreases, noise pollution becomes more serious, and even almost effective signal is overwhelmed. It can be seen from the figure that when there is no noise (Fig. 5a) and when SNR = 21dB (Fig. 5b), the microseismic event is clearly visible; when SNR = 11dB (Fig. 5c), the noise begins to seriously influence the effect; When SNR = 6dB (Fig. 5d), the effective signal is completely submerged in the noise, and it is difficult for the human eye to distinguish.



(a) noise-free (b) *SNR* = 21dB (c) *SNR* = 11dB (d) *SNR* = 6dB

Fig. 5. Synthetic signals with different *SNR*.

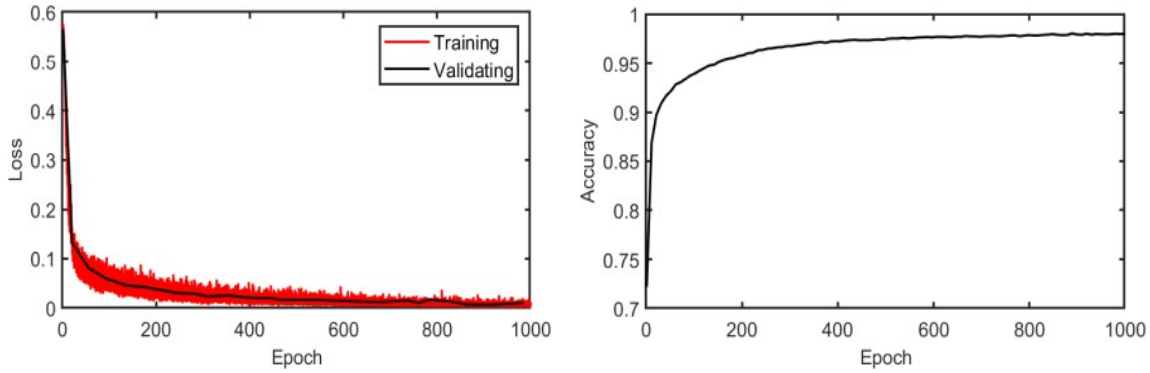
Model training

The training set contains 4000 samples. During training, the batch size is set to 200, the epoch is set to 1000, and the learning rate is set to 0.01. The model uses Python language and Pytorch framework as the programming environment. The experiment was run on a computer equipped with an Intel Core i7 9700k processor and 32G memory, and the model performance was verified after every 10 training sessions.

The loss function is shown in eq. (15). We backpropagate the loss to

train the network and optimize the predicted distribution that describes the true one by a stochastic gradient descent (SEG) optimizer. Other popular optimizes such as the Adam can also be used in the algorithm. From our experience, SEG shows reasonably good convergence for this task, and thus, we use this optimizer in our algorithm.

Figs. 6a-6b show the loss and accuracy with iterations of MSD-CNN. The accuracy reflects the probability of accurate detection of waveform, and the loss indicates the learning effect of the model. The results verify that the accuracy and loss of the model has not changed much in the last 600 epochs, indicating the model eventually approaches the fitting. The final accuracy of verification is 97.98%, respectively, and the losses are 0.014 and 0.042 in train and validation, respectively. It can be seen that the MSD-CNN model performs well in both training and verification.



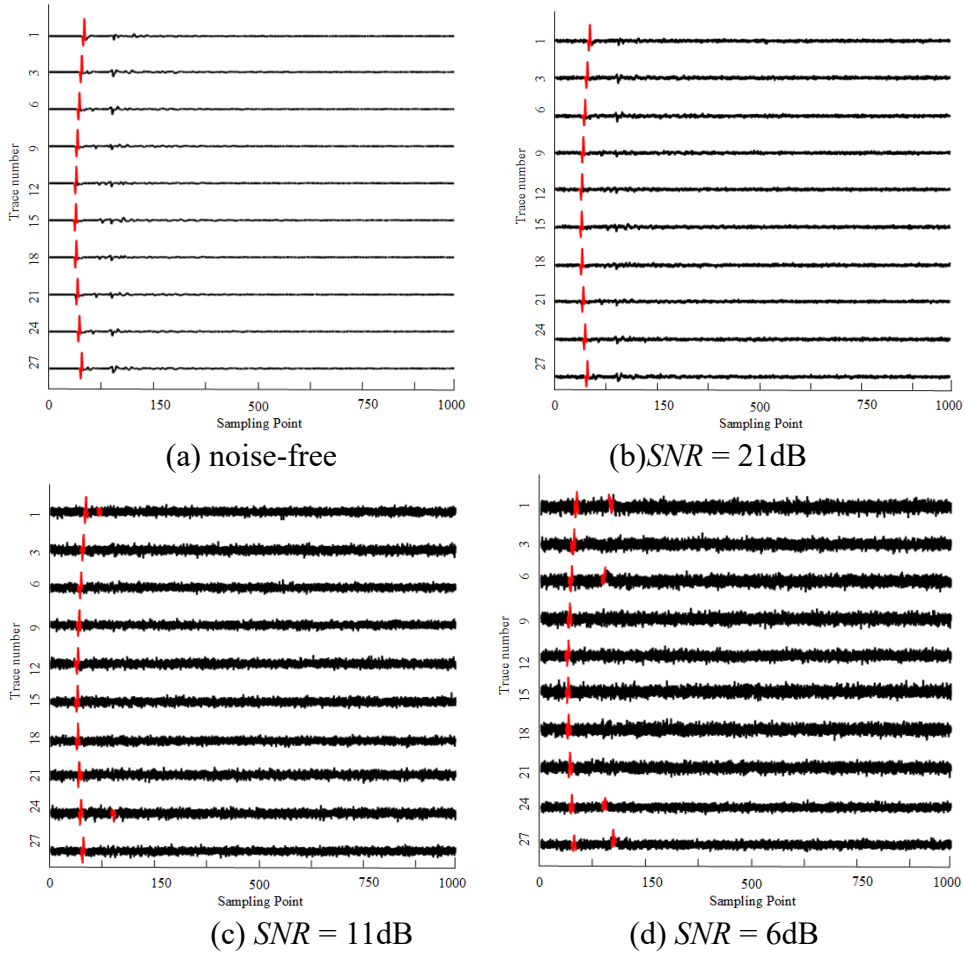
(a) The training and validation loss (b) The validation accuracy

Fig. 6. Synthetic data train results.

Input the test set samples into the network with a good performance model, and the detection results are shown in Table 1. Among the 800 test samples, half are positive samples. When noise free, both TP and FN are 400, which means that samples containing microseismic waveform and samples that do not contain microseismic waveforms are all correctly identified. As the SNR decrease, the more positive samples are submerged by noise, the accuracy of the recognition results gradually decreases. As shown in the table, the accuracy is 100% when there is no noise interference; the accuracy is 97.38% when SNR = 21dB; the accuracy is 93.75% when SNR =11dB; the accuracy is 91.13% when SNR = 6dB.

Table 1. The result of model identification.

SNR (dB)	*TP*	*FN*	*TN*	*FP*
Noise-free	400	0	400	0
21	390	10	389	11
11	373	27	377	23
6	360	40	369	31



(a) noise-free (b)*SNR* = 21dB

(c) *SNR* = 11dB (d) *SNR* = 6dB

Fig. 7. Detection results of synthetic data.

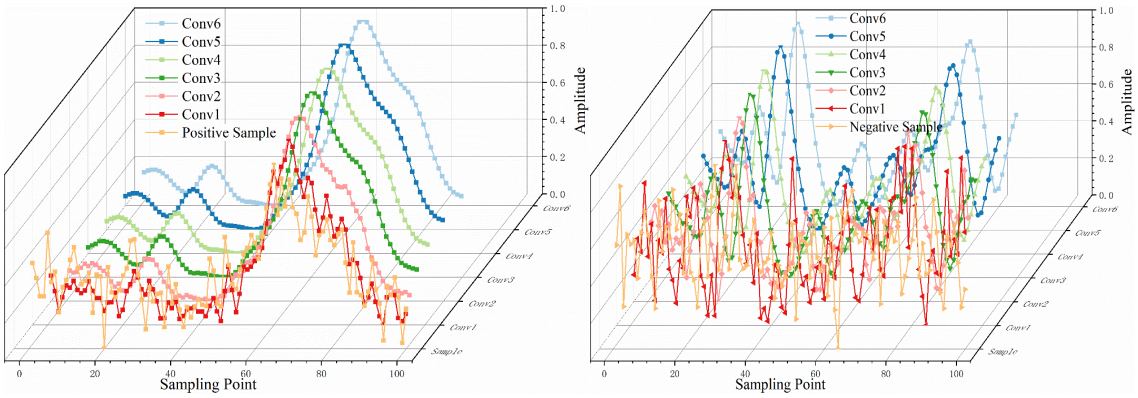
Figs. 7a-7d show the partial detection results of synthetic data. It can be seen from the figure that when noise free (Fig. 7a) and SNR = 21dB (Fig. 7b), the detection results obtained are all accurate. The 10 waveform have identified microseismic event waveforms at samples 71, 65, 57, 54, 49, 48, 51, 53, 59, and 64, respectively. When SNR = 11dB (Fig. 7c), the detection

results include the first trace (sample 105) and the ninth trace (sample 139) respectively has a misdetection. When SNR = 6dB (Fig. 7d), four noise samples are misdetection as containing microseismic event waveforms.

EXPERIMENTAL RESULTS

Network structure analysis

Compared with the traditional method, the neural network has the advantage of nonlinear mapping ability. The sample input network extracts the features through the convolution layer, activation function, batch normalization layer to get the feature map. Fig. 8a is a feature map extracted from a positive sample the noise is suppressed and the main trend of change is strengthened, through the network. Fig. 8b is a feature map extracted from the negative sample. The samples are smoothed through the network, the main frequency information is enlarged and the trend of change is strengthened. The neural network cannot distinguish the amplitude difference between the microseismic waveform and the noise because of the function of the batch normalization layer in the network, but the network can distinguish the microseismic waveform and the noise by amplifying the variation trend of the samples and the difference of frequency.



(a) Feature map of a positive sample (b) Feature map of a negative sample

Fig. 8. Feature map.

Low-scale features contain more detailed information, but have lower semantic and more noise; high-level features have stronger semantic information, but their ability to perceive details is poor. In the process of a continuous deepening of the CNN, the ability to express abstract features is becoming stronger, but the loss of low-scale spatial information is serious, and the integration of the two can improve the overall performance of the network.

Multiscale analysis

The C-F model tends to favor a more reliable detector, which makes the final result more reliable when selecting the detection result. In the training process, each detection result of the detector will affect its credibility, and the detector will adjust the parameters in the next round of training to improve its credibility as much as possible.

471

To verify the validity of the multiscale, the operational status of the model in the middle part of the microseismic waveform detection is analyzed. For the convenience of analysis, select the following ten typical samples (the first half waveform with noise-free (sample 1), the complete waveform with noise-free (sample 4), the second half waveform with noise-free (sample 7), and the SNR = 21dB samples (samples 2, 5, 8), SNR = 11 dB (samples 3, 6, 9) and noise background with SNR = 6dB (sample 10)), as shown in Fig. 9. In the lower-scale feature maps (scale 1, scale 2, scale 3), it is clearly observed that the network still has a certain degree of attention to noise interference. As the network deepens, more attention is paid to the characteristics of microseismic waveform.

Noise samples are still paid attention to low scales (scale 1, scale 2, scale3), but high scales no longer follow; positive samples receive high attention at low scales (the color is very dark), at high scales, it will distract a little attention, marked by the red box in Fig. 9.

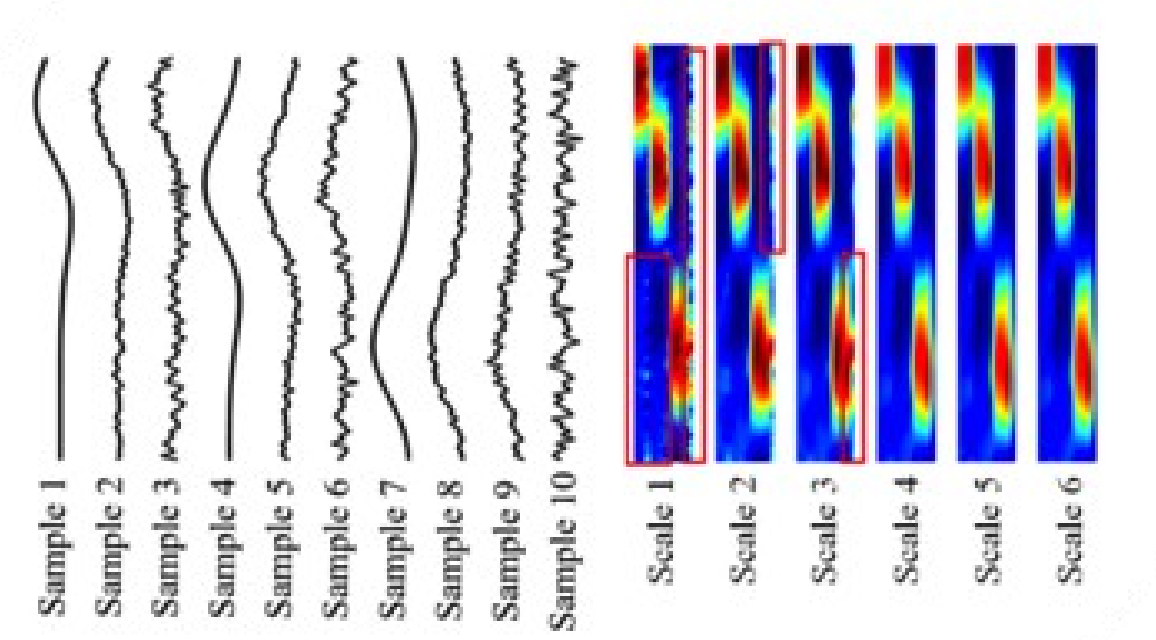


Fig. 9. Analysis about Multiscale feature map.

Because the depth of the feature layer is different, the characterization ability is also different. At low scales, when the sample SNR is high, the dynamic intensity is high, but when the noise interference is large, the dynamic intensity is still high, and even large background noise is misjudged as a microseismic event; at high scales, the dynamic intensity of negative samples is very small, but the dynamic intensity of positive samples with high SNR or even clean samples will be slightly reduced. Therefore, the comprehensive calculation of the confidence factor of each scale by the C-F model can improve the network detection accuracy.

472

Comparison

To verify the performance of the model, the detection ability of different methods for microseismic waveform is compared, wavelet analysis, BP network, CNN network, and multiscale detection neural network are used to identify microseismic data with different SNR.

Fig. 10 shows the accuracy of the detection results of the four methods under the different noise disturbances. It can be seen that the model in this paper combines multiscale features, and the detection effect is the best. Even when the noise-free (shown as SNR = 30dB) pollution, the accuracy can reach 99.99%; the accuracy can reach 97.38% under light noise pollution (SNR = 21dB).

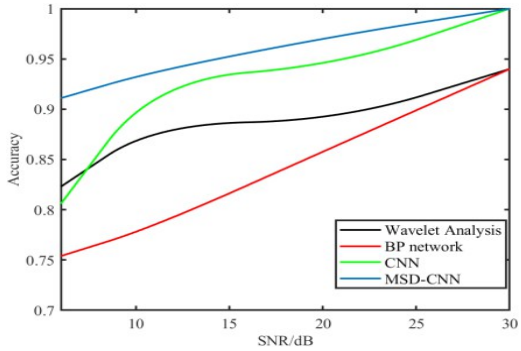


Fig. 10. Accuracy of different methods with different SNR.

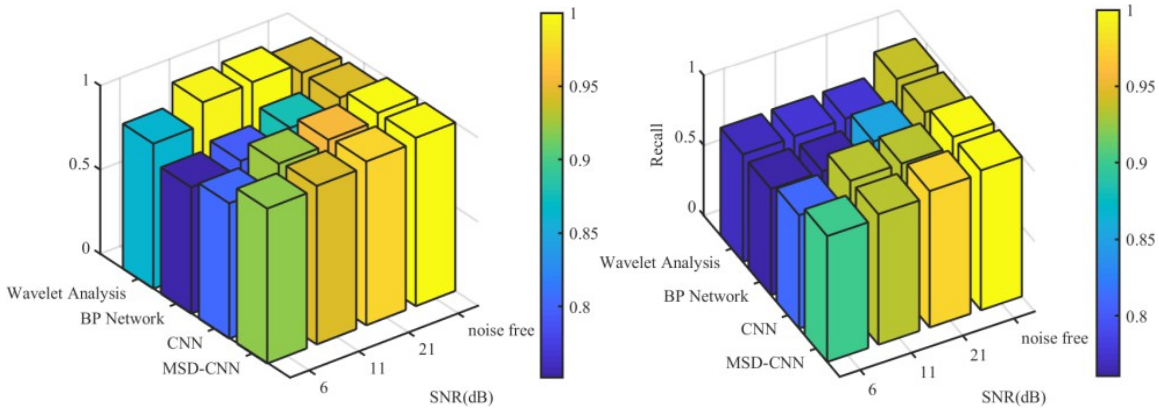
To further evaluate the performance of the model, we introduce two parameters, namely precision and recall.

$$Precision = TP/(TP + FP \tag{29}$$

$$Recall = TP/(TP + FN \tag{30}$$

Precision is defined as the ratio of predictions that are correctly positive and predictions that are classed as positive (both true and false positive). Recall is defined as the ratio of predictions that are correctly positive and actual positive samples (true positive and false negative).

Fig. 11a show the accuracy of different methods under different noise intensities. Fig. 11b show the recall of different methods under different noise intensities. It can be seen from the figure: in the case of no noise pollution, the accuracy and recall of the four methods are all above 0.94. In general, the four methods show that the accuracy and recall decrease with the increase of noise interference. Two parameters of the three deep learning methods are gradually increased by the BP model, the CNN model, and MSD-CNN. The accuracy and recall of wavelet analysis is affected by the threshold value. The larger the threshold, the more positive samples cannot be identified, and the smaller the threshold, the more negative samples can be misdetected.



(a) P of different methods under different SNR (b) R of different methods under different SNR

Fig. 11. Precision and recall.

Randomly select 100 positive samples and negative samples with SNR = 21dB, SNR = 11dB, SNR = 6dB, and noise-free from test sets, and use wavelet analysis, BP model, CNN model, and MSD-CNN to obtain the values of precision and recall for each method. As shown in Fig. 12, the precision and recall of MSD-CNN is higher than the other two deep learning methods. The precision of the wavelet analysis method is slightly higher than MSD-CNN, but its recall and accuracy is lower than MSD-CNN.

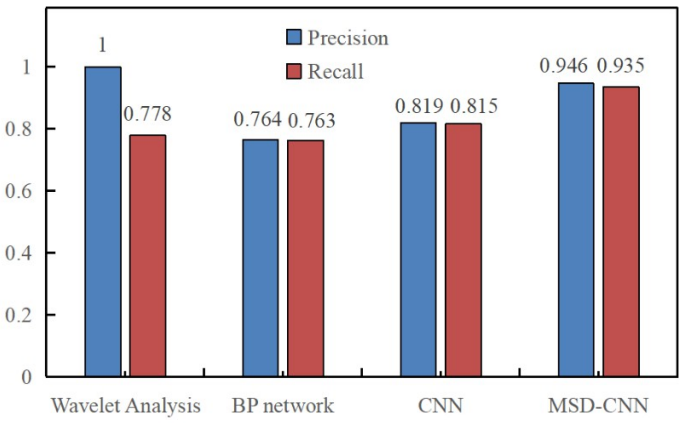


Fig. 12. Precision and Recall of different.

The time consuming of identifying 4000 samples by wavelet analysis, BP network, CNN network, and MSD-CNN is shown in Fig. 13. Wavelet analysis is much more time-consuming than deep learning. Among the deep learning methods, the detection of the BP model consumes the least time, followed by CNN. MSD-CNN is slower than the other two methods because the increase in the number of detectors and parameter calculation. But to obtain increased accuracy, it's worthwhile to increase the average time-consuming.

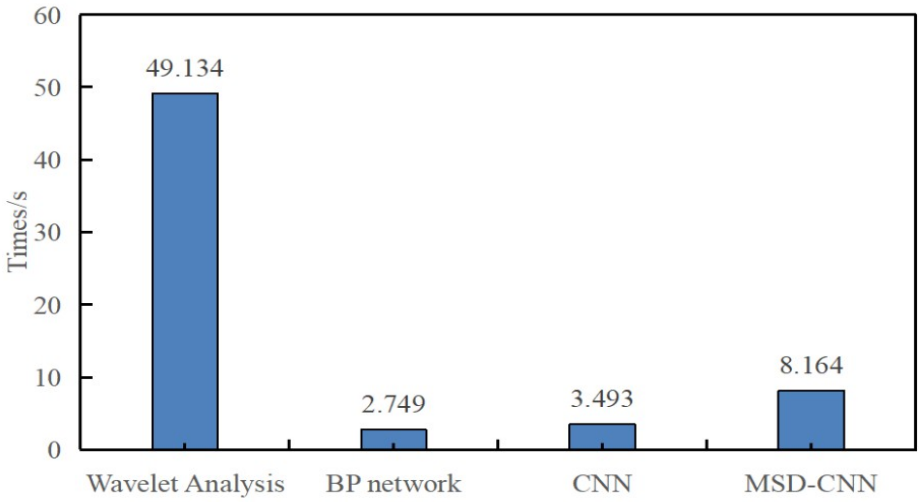


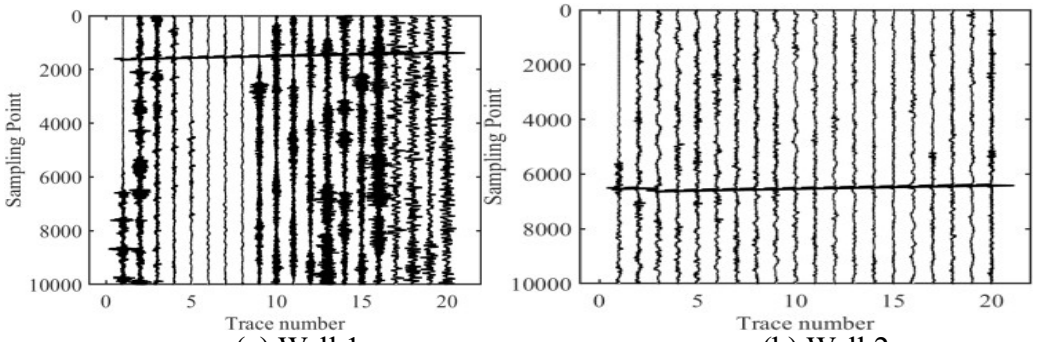
Fig.13. Time consuming of different methods

Field data event detection

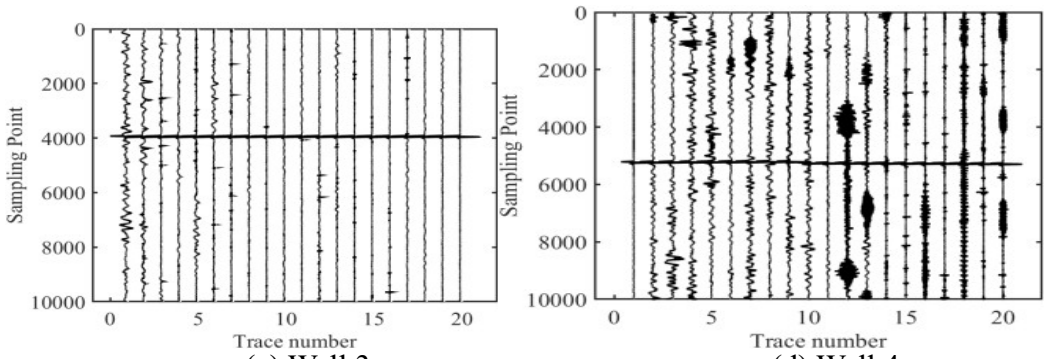
Due to differences in lithological characteristics, natural cracks, and

soil thickness, the noise contained in the microseismic waveform in different regions is various, and the microseismic data collected in the same region due to the equipment and operating environment will also be very different. The filed microseismic data is 100 traces from four fractured wells. The monitoring data of each well is randomly selected 60 traces. After preprocessing, the network is trained again on the model obtained from the simulation data. The remaining data of each fractured well were used for testing, and the results prove that the model can effectively identify field microseismic events.

Figs. 14a-14d shows the partial monitoring waveform data of four fractured wells. The data of Well 1 suffers a lot of noise interference, which almost floods the effective microseismic signal from a single trace; The data of Wells 2 and 3 suffer from less noise pollution and the amplitude characteristics of the microseismic signal are more obvious; Well 4 data had strong intermittent interference and macroshock events; Figs. 15a-15d show the partial detection results of the four wells. It can be seen from the recognition results that was misdetected noise samples and the microseismic events have a great similarity in the waveform, and it is extremely difficult for the human eye to distinguish. The recognition results show that our proposed method has good generalization ability and can effectively identify microseismic events from field data.

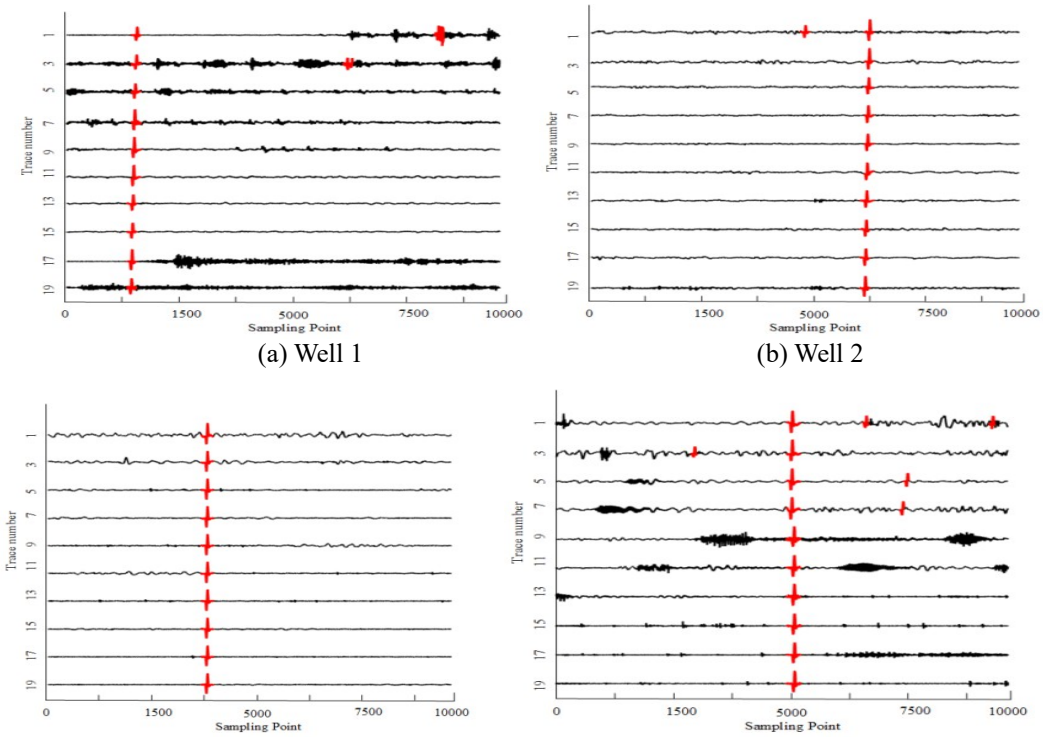


(a) Well 1 (b) Well 2



(c) Well 3 (d) Well 4

Fig. 14. Waveform data of different wells.



(a) Well 1 (b) Well 2

(c) Well 3 (d) Well 4

Fig. 15. Detection results of different wells.

DISCUSSION AND CONCLUSIONS

In this study, we proposed a method of microseismic event detection based on a multiscale neural network. The MSD-CNN model is built according to the natural multiscale feature extracted from the neural network, detectors are trained to detect microseismic events at each scale, and the contribution of each scale detection results to the final detection result is determined by the C-F model. The experimental results show that the model has certain anti-noise ability and generalization ability.

However, the waveform of microseismic data is diverse and the generalization ability of the network is limited, which can not guarantee the good detection ability of microseismic data in all regions. Further research is needed to find an automatic microseismic detection method suitable for multiple regions. Moreover, there are many types of microseisms, such as oil and gas production microseisms, marine microseisms, landslide microseisms, volcanic microseisms, tunnel collapse microseisms, etc. our method can not directly distinguish which type of microseisms belong to. Therefore, looking for a deep learning method that can not only identify microseismic events, but also distinguish different types of microseismic events is also the focus of further research.

ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of Heilongjiang Province under Grant No. LH2023D009.

REFERENCES

Boureau, Y.L, Bach, F., LeCun, Y. and Ponce, J., 2010. Learning mid-level features for recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE,: 2559-2566.

Cai, Z., Fan, Q., Feris, R.S. and Vasconcelos, N., 2016. A unified multiscale deep convolutional neural network for fast object detection. Proc. Europ. Conf. Comput. Vis., Springer Internat. Publish.: 354-370.

Dai, J., Li, Y., He, K. and Sun, 2016. R-FCN: Object detection via region-based fully convolutional networks. Proc. 30th Ann Conf. Neur. Inform. Process. Syst., Barcelona: 379-387.

Kulesh, M., Holschneider, M. and Diallo, M.S., 2008. Geophysical wavelet library: Applications of the continuous wavelet transform to the polarization and dispersion analysis of signals. Comput. Geosci., 34: 1732-1752.

Li, Y., Yang, T.H. and Liu, H.L., 2016. Real-time microseismic monitoring and its characteristic analysis in working face with high-intensity mining. J. Appl. Geophys., 132: 152-163.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y. and Berg. A.C., 2016. SSD: Single shot multibox detector. Europ. Conf. Comput. Vision, Springer International Publishing: 21-37.

Liu, W, Rabinovich, A., Berg, A.C. and Parsenet, B., 2015. Looking wider to see better. arXiv preprint arXiv:1506.04579.

Maochen, G., 2005. Efficient mine microseismic monitoring. Internat. J. Coal Geol., 64,:

44-56.

Maxwell, S.C., Rutledge, J., Jones, R. and Fehler, M., 2010. Petroleum reservoir characterization using downhole microseismic monitoring. Geophysics, 75(5): A129-A137.

Redmon, J. and Farhadi, A., 2017. YOLO9000: better, faster, stronger. Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Piscataway: 7263-7271.

Redmon, J. and Farhadi, A., 2018. YOLOv3: An incremental improvement. Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Piscataway: 1-6.

Shang, X., Li, X., Morales-Esteban, A. and Chen, G.H., 2017. Improving microseismic event and quarry blast classification using Artificial Neural Networks based on Principal Component Analysis. Earthq. Engineer., 99: 42-149.

Sheng, G.Q., Tang, X.G., Xie, K. and Jie, X., 2019. Hydraulic fracturing microseismic first arrival picking method based on non-subsampled shearlet transform and higher-order-statistics. J. Seismic Explor., 28: 593-618.

Shi-Chao, L.V., Song, W.Q., Liu, Y.M., Guo, X.Z. and Zhang, H.F., 2013. The polarization constrained LTA/STA method for automatic detection of microseismic. Geophys. Geochem. Explor., 37: 488-493.

Singh, B., Najibi, M., Davis, L.S., 2018. SNIPER: Efficient multiscale training. Proc. 32nd Conf. Neural Informat. Process. Syst., Montreal: 9310-9320.

Sleeman, R. and Eck, T.V., 1999. Robust automatic P-phase picking: an on-line implementation in the analysis of broadBand seismogram recordings. Phys. Earth Planet. Inter., 113: 265-275.

Sun, J., Wang, L. and Hou, H., 2012. Application of micro-seismic monitoring technology in mining engineering. Internat. J. Mining Sci. Technol., 22: 79-83.

Tan, Y. and Nava, F.A., 1988. Automatic seismic wave detection and autoregressive model method. Mathemat. Geol., 20: 37-48.

Velis, D., Sabbione, J.I. and Sacchi, M.D., 2015. Fast and automatic microseismic phase-arrival detection and denoising by pattern recognition and reduced-rank filtering. Geophysics, 80(6): WC25-WC38.

Xie, J., Yang, C., Gupta, N., King, M.J. and Datta-Gupta, A., 2015. Integration of shale gas production data and micro-seismic for fracture and reservoir properties using fast marching method. SPE Journal, 20: 347-359.

Xu, S., Zhang, C.R., Chen, Z.Y., Li, Y.H. and Liu, J.P., 2021. Accurate identification of microseismic waveforms based on an improved neural network model. J. Appl. Geophys., 190: 926-951.

Li, Y., Yang, T.H., Liu, H.L., Wang, H., Hou, X.-G., Zhang, P.-H. and Wang, P.-T., 2016. Real-time microseismic monitoring and its characteristic analysis in working face with high-intensity mining. J. Appl. Geophys., 132: 152-163.

Zhang, H., 2003. Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings. Bull. Seismol. Soc. Am., 93: 1904-1912.

Zhang, Y., Chan, W. and Jaitly, N., 2017. Very deep convolutional networks for end-toend speech recognition. IEEE Internat. Conf. Acoust., Speech Signal Process.: 4845-4849.

Zhao, H., Shi, J., Qi, X., Wang, X. and Jia, J., 2017. Pyramid scene parsing network.

IEEE Conf. Comput. Vis. Pattern Recognit., Piscataway: 2881-2890.