Geophysical Journal International

Geophys. J. Int. (2021) **226,** 2145–2157
Advance Access publication 2021 March 15
GJI General Geophysical Methods

https://doi.org/https://doi.org/10.1093/gji/ggab099

Integrated processing method for microseismic signal based on deep neural network

Hang Zhang [1,2,3] Chunchi Ma [1,2] Yupeng Jiang,[4] Tianbin Li,[1,2] Veronica Pazzi[3] and Nicola Casagli[3]

[1] *State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, Sichuan* 610059*, China. E-mail: machunchi17@cdut.edu.cn*
[2] *College of Environment and Civil Engineering, Chengdu University of Technology, Chengdu, Sichuan* 610059*, China*
[3] *Department of Earth Sciences, University of Florence, 50121 Florence, Italy*
[4] *School of Civil Engineering, University of Sydney, Sydney, NSW* 2006*, Australia*

Accepted 2021 March 5. Received 2021 February 9; in original form 2020 June 30

SUMMARY

Denoising and onset time picking of signals are essential before extracting source information from collected seismic/microseismic data. We proposed an advanced deep dual-tasking network (DDTN) that integrates these two procedures sequentially to achieve the optimal performance. Two homo-structured encoder–decoder networks with specially designed structures and parameters are connected in series for handling the denoising and detection of microseismic signals. Based on the similarity of data types, the output of the denoising network will be imported into the detection network to obtain labels for the signal duration. The procedures of denoising and duration detection can be completed in an integrated way, where the denoised signals can improve the accuracy of onset time picking. Results show that the method has a good performance for the denoising of microseismic signals that contain various types and intensities of noise. Compared with existing methods, DDTN removes the noise with a minor waveform distortion. It is ideal for recovering the microseismic signal while maintaining a good capacity for onset time picking when the signal-to-noise ratio is low. Based on that, the second network can detect a more accurate duration of microseismic signals and thus obtain more accurate onset time. The method has great potential to be extended to the study of exploration seismology and earthquakes.

Key words: Neural networks, fuzzy logic; Time-series analysis; Induced seismicity; Seismic noise.

1 INTRODUCTION

Recorded microseismic data are often polluted by various types of noise due to complicated environmental situations, such as building construction, mechanical, electrical, or traffic noise. Signal denoising and onset time picking are essential for the interpretation of microseismic events (location, moment magnitude, and energy release). The accuracy and efficiency of these two procedures are deeply affected by the properties of the noise, signal waveform, and sampling rate. For example, denoising becomes challenging if the frequency band of signal and noises overlap, every frequency band of signal wavelet decomposition may contain noise. In such a case, it is difficult to distinguish the signal from the noises in the frequency domain. The filtering of the noises often causes a slight or strong distortion of the signal waveforms, which further reduces the accuracy of the onset time picking of signals. A severe loss of sampling points during the time–frequency transformation also affects the accuracy of onset time picking. Data flow explosion is another challenge for practicing the signal denoising and onset time picking efficiently in real-time microseismic monitoring that is demanded by modern tunnel engineering.

Many studies have been conducted to suppress the noise in microseismic/seismic data, which include the wave-packet transform (WPT; Galiana-Merinoet *et al.* 2003; Liu & Xun 2014), S-transform (Tselentis *et al.* 2012), short time Fourier transform (STFT; Mousavi *et al.* 2016a), the continuous wavelet transform (CWT; Mousavi & Langston 2016b, 2016c), empirical mode decomposition (EMD; Huang *et al.* 1998; Chen *et al.* 2017; Bekara & Baan 2009), variational mode decomposition (VMD; Ma *et al.* 2020), fuzzy methods (Hashemi *et al.* 2008), singular spectrum analysis (Oropeza & Sacchi 2011), sparse transform-based denoising (Chen *et al.* 2016), mathematical morphology based denoising approach (Li *et al.* 2016) and the non-local means (NLM) algorithm (Bonar & Sacchi 2012). Although these methods can, in a

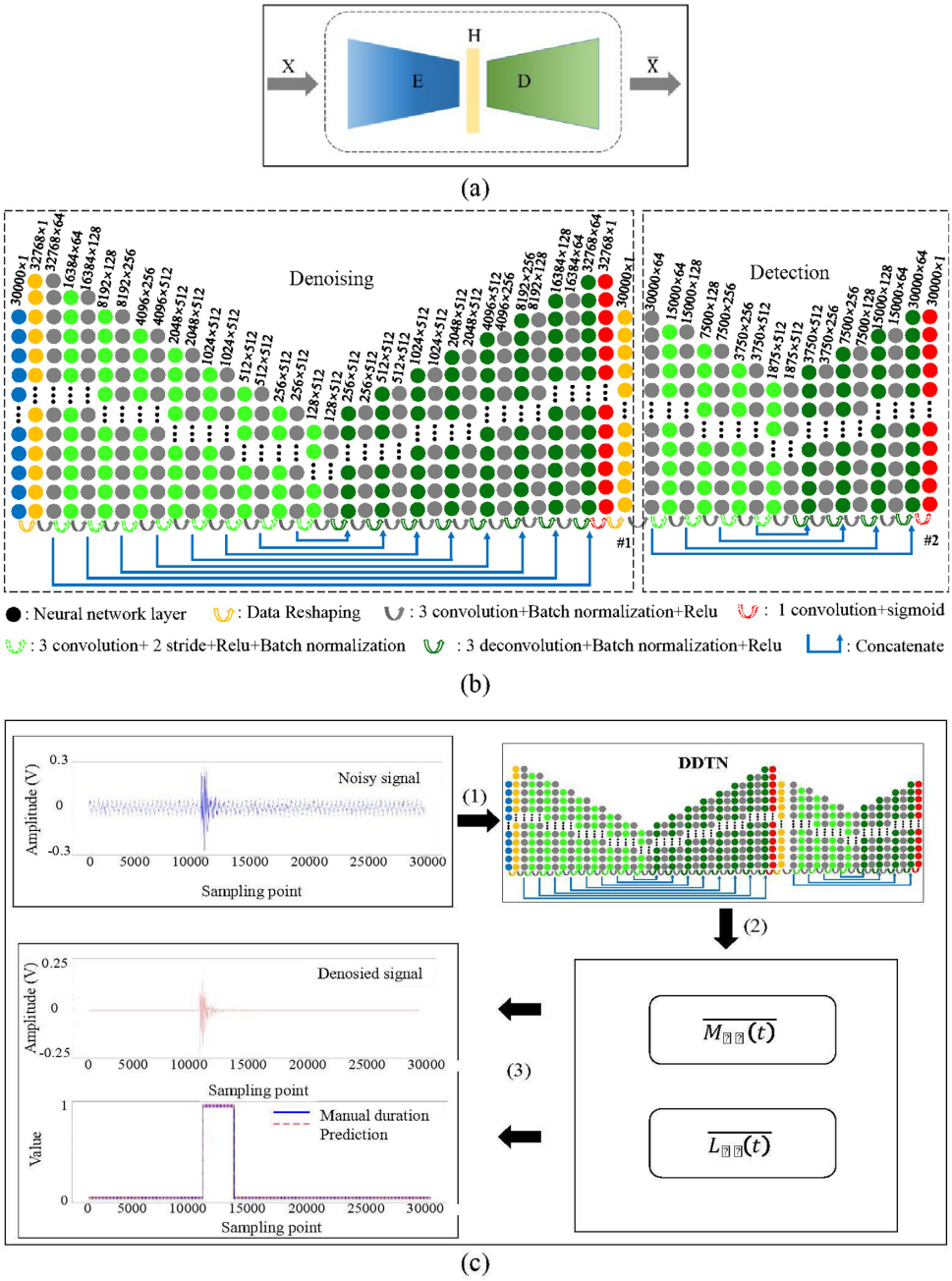

Figure 1. (a) Scheme of a convolutional encoder–decoder architecture. (b) The architecture of DDTN. Two dotted boxes represent the denoising and detection network, respectively. The circles of different colours represent neural network layers. Arrows represent different operations applied between two adjacent layers. Batch normalization and concatenate are used to improve convergence during training. The dimension of each layer is presented above, which contains 'features × filters'. The DDTN includes 24 convolution and 24 deconvolution layers, resulting in trainable parameters of about 2.04E + 07. Layers #1 and #2 correspond to the outputs of DDTN: the estimated mask $\overline{M_{\mathrm{Dn}}(t)}$ and label $\overline{L_{\mathrm{Dt}}(t)}$, respectively. (c) Flow diagram of the DDTN: (1) The noisy signals in the time domain are input into the DDTN; (2) the estimated mask $\overline{M_{\mathrm{Dn}}(t)}$ and label $\overline{L_{\mathrm{Dt}}(t)}$ are produced as outputs by the network; (3) The $\overline{M_{\mathrm{Dn}}(t)}$ and $\overline{L_{\mathrm{Dt}}(t)}$ are applied for denoising and duration detection.

certain extension, suppress noise components in the data, choosing the optimal mapping function between the sets of the noisy and denoised signal is still difficult.

To replace the manual operation, various adaptive algorithms have also been proposed for signal onset time picking. Especially the short-term/long-term averaging (STA/LTA) algorithm (Allen 1978) is the most popular method. However, it tends to misjudge, or even totally miss, the consecutive events in the data if the signal-to-noise ratio (SNR) is low. The Akaike information criterion (AIC) algorithm (Takanami & Kitagawa 1991), alternatively, can better distinguish the boundary of a series of events based on autoregressive model assumptions. But the method is error-prone since the local maximum and global maximum may not be separated. The fractal dimension method (Boschetti *et al.* 1996) identifies the onset time based on the change of the characteristics in the fractal dimension of signals, which also avoids the adjacent

Table 1. Performance comparison of different depth of network structures on the validation data set. The number in the name of the model indicates the total numbers of convolution and deconvolution layers of the denoising network. The indeces Val_loss and Val_accuracy are defined as the loss and accuracy of the validation data set, respectively.

Structure	Variables ($\times 10^7$)	FLOPs ($\times 10^7$)	Val_loss		Val_accuracy	
			Denoising	Detection	Denoising	Detection
DDTN-24	1.25	4.01	0.082	0.015	0.895	0.842
DDTN-32	2.04	6.23	0.051	0.009	0.945	0.895
DDTN-48	2.67	8.15	0.042	0.005	0.952	0.913

channel interference. However, these methods use the interpolation of the signal, which will generate large errors with the high sampling and seismic trace frequency. Methods based on higher-order statistics have been developed to identify the transition from Gaussianity to non-Gaussianity., which is consistent with the onset of the seismic event (Yung & Ikeller 1997; Saragiotis *et al.* 2004). Studies have tested the onset time picking of traditional shallow neural networks based on the definition of different features (e.g. variance, the absolute value of skewness, kurtosis, and a combination of skewness and kurtosis predicted based on sliding windows) (Mccomack *et al.* 1993; Gentili & Michelini 2006). Also, some hybrid methods have been proposed to evaluate the performance of onset time picking (Zhang 2003; Diehl *et al.* 2009; Zhao *et al.* 2012; Jia *et al.* 2015). Generally, studies have revealed that the accuracy of these methods has strongly dependent on the selection of parameters, which are usually set empirically.

Deep-learning-based methods have been rapidly developed to deal with the existing deficiencies of signal processing methods. It can learn extremely complex functions through the neural network. It has been proved to be a powerful tool for data processing (Goodfellow *et al.* 2016; Perol *et al.* 2018; Ross *et al.* 2018; Yuan *et al.* 2018; Zheng *et al.* 2018; Zhu & Beroza 2018; Dong *et al.* 2019; Dong *et al.* 2020). Inspired by the ability of the encoder–decoder network for image/signal processing (Mao *et al.* 2016), the goal of this work was approached as a supervised learning problem in which a deep dual-tasking network (DDTN) for denoising and duration detection of microseismic signals is built. Two optimized encoder–decoder networks in series, defined as denoising and detection networks, are integrated through a convolutional operator layer to construct our DDTN. Given a noisy input signal, the DDTN can produce two individual outputs (sparse representations). One is defined as a mask and map input to a clean signal; the other is defined as a label for labelling the duration of microseismic signals. Microseismic data recorded in the Micang Mountain tunnel in Sichuan, China is used for network training, validation, and testing. The performance is rigorously compared with existing methods on denoising and onset time picking of semisynthetic signals, which is generated by superimposing microseismic signals and real noise. The proposed method has been also tested in other projects that have different engineering situations.

2 METHODOLOGY AND TRAINING DATA

The noisy signal is defined as $NS(t)$, which represents a superposition of microseismic signal $MS(t)$ and noise $N(t)$ as written:

$$NS(t) = MS(t) + N(t) \ , \tag{1}$$

where t represents the sampling point and noise $N(t)$ is instrumental or unknown noises. The purpose of denoising in DDTN is to minimize the expected *Error* between the actual signal $MS(t)$ and the calculated signal $\overline{MS(t)}$:

$$\text{Error} = \frac{1}{n}\sum_{i=1}^{n}\left(MS(t)_i - \overline{MS(t)_i}\right)^2 , \tag{2}$$

where $\overline{MS(t)} = f(t)NS(t)$, $f(t)$ is the function that maps $NS(t)$ to the $\overline{MS(t)}$, and n is the number of samples. We regard the process of finding this mapping function as a supervised learning problem in which a deep neural network learns to extract a sparse representation of input noisy waveform $NS(t)$ and map it to the clean signal. The mapping function $f(t)$ is defined as a masking vector $M_{\text{Dn}}(t)$ for signal denoising, which contains a series of values between 0 and 1 to attenuate the noisy signal:

$$f(t) = M_{\text{Dn}}(t) = \frac{\frac{|MS(t)|}{|N(t)|}}{1 + \frac{|MS(t)|}{|N(t)|}}. \tag{3}$$

A binary label vector $L_{\text{Dt}}(t)$ is manually generated for labelling the signal duration, where the values are set equal to 1 corresponding to the duration of microseismic signal and the rest to 0:

$$L_{\text{Dt}}(t) = \begin{cases} 1 & t \in \text{Duration of signal} \\ 0 & t \notin \text{Duration of signal} \end{cases} . \tag{4}$$

Therefore, the onset time is the first sampling point with a value of 1 in $L_{\text{Dt}}(t)$. Both of these two vectors have the same sizes as the noisy signal N$S(t)$, and they are the targets for optimizing the performance of the neural network during training.

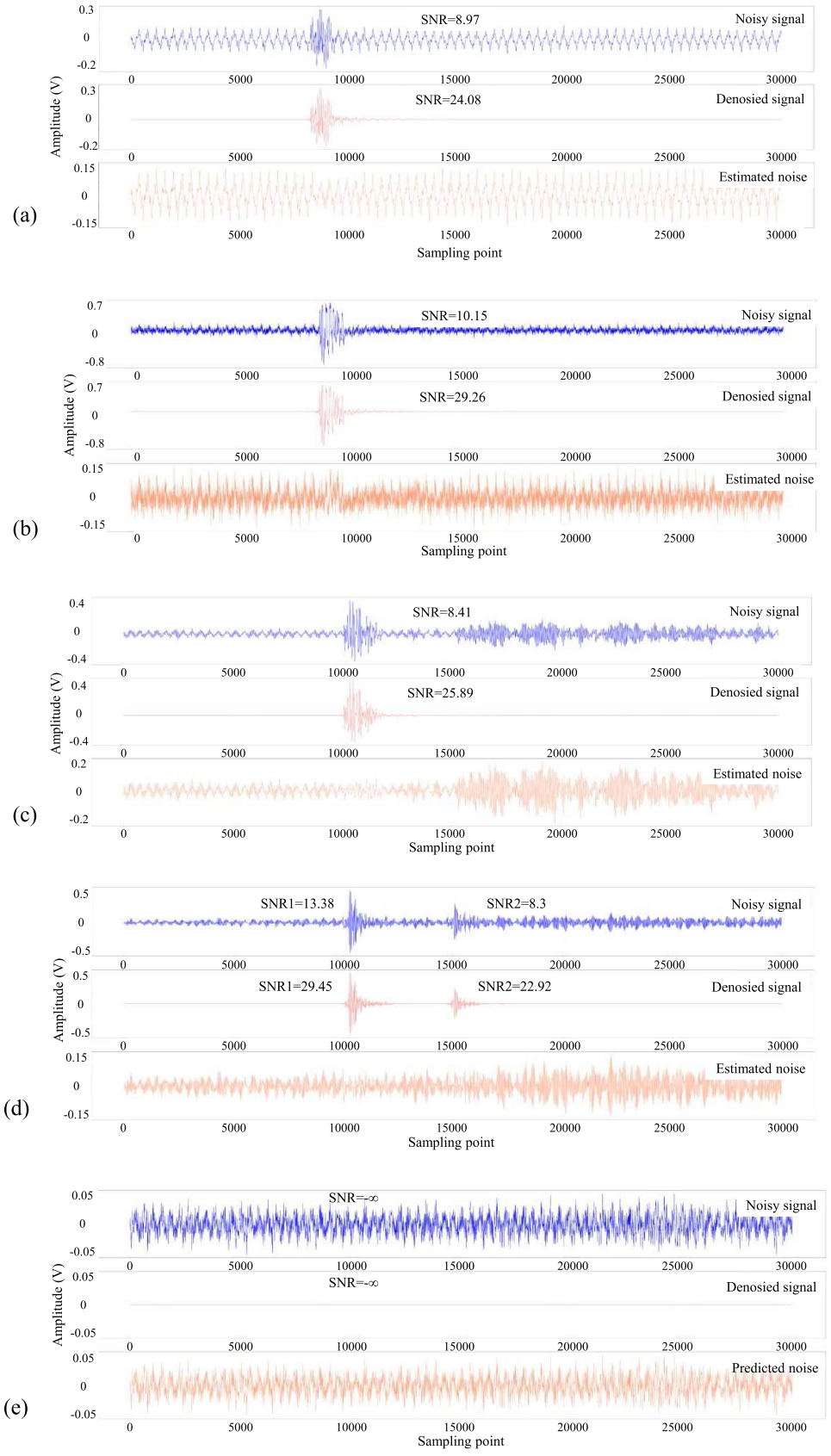

Figure 2. The denoising performance of DDTN for different noisy signals in the synthetic test set. The noises in (a) and (b) are generated using cyclic noises; (c) and (d) are further mixed with real-field noise; (e) is the real-field noise-only signal set. The first two rows of signals in each case are a noisy signal and a denoised signal, respectively; the third signal is the estimated noise.

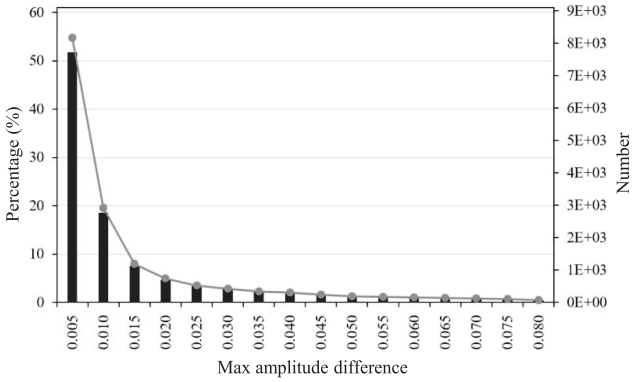

Figure 3. Distribution of max amplitude difference between real noise and estimated noise.

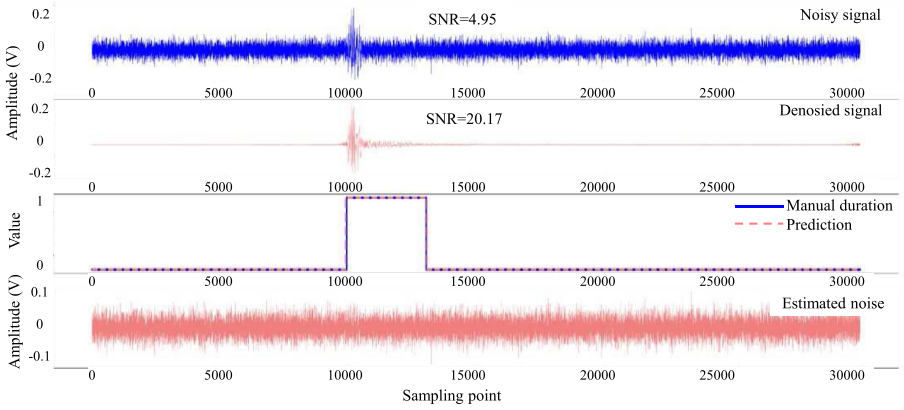

Figure 4. The combined performance of denoising and duration detection for noisy signals, which is formed with a clean microseismic signal and Gauss noise. The Gauss noise is generated by a standard normal distribution and the semisynthetic noisy signal with the SNR of 5 dB. The first two rows of signals are a noisy signal and a denoised signal, respectively, the third row of the signal includes the manual and predicted duration of microseismic signals, and the fourth row of the signal represents the estimated Gauss noise.

The convolutional neural network is applied for both signal denoising and duration detection. Its basic structure can be logically split into three individual parts: *encoder* and *decoder* and *representation* (Fig. 1a). The encoder maps the input data X into a hidden representation H through the function E (i.e. $H = E(X)$), which is an extraction process for the features of input data; the decoder transforms H into an estimation of the input by the function D, i.e. $\overline{X} = D(H)$, which generates outputs with a high-dimensional nonlinear mapping of the sparse representation. The convolutional neural network has been proved to be a powerful tool for denoising and onset time picking of seismic signals (Zhao *et al.* 2019; Zhu *et al.* 2019). To solve the denoising and detection of microseismic signals sequentially and effectively, the DDTN is constructed by connecting two homo-structured encoder–decoder networks through convolutional operators.

The noisy signal (see Fig. 1b), as the input vector of the denoising network, is first reshaped from (30 000, 1) into (32 768, 1) by a zero-padding procedure for the encoding process. It avoids a lack of noisy signal information in the encoding and decoding process. The reshaped vector has then been transformed into new layers through a series of encoding operators that consist of one convolution, one batch normalization (Ioffe & Szegedy 2015), and one ReLU (rectified linear unit) activation; a stride of 2 also performed alternatively to shrink the feature space gradually and improve the computational efficiency during the encoding process. For the convolution calculation, a larger kernel has a wilder receptive field, which obtains more non-local features of the signal. However, the large convolution kernel leads to a dramatic increase in computational time. It would severely limit the depth of the neural network. Therefore, the kernel size of convolution layers is set to be a constant three. The decoding operators also use a kernel size of three for the deconvolution to generate the estimated $\overline{M_{\mathrm{Dn}}(t)}$ and $\overline{L_{\mathrm{Dt}}(t)}$ in the decoding process.

The corresponding feature maps in the encoding and decoding process are concatenated, which improve the convergence of training and perform good reconstruction information of signal (Ronneberger *et al.* 2015). In the penultimate layer of the denoising network, a sigmoid binary activation function is used to produce $\overline{M_{\mathrm{Dn}}(t)}$ for signal denoising, and in the last layer the mask is reshaped into (30 000, 1). Apart from the reshaping step at the beginning, the encoding and decoding operators in the detection network are similar to the denoising network; the $\overline{L_{\mathrm{Dt}}(t)}$ for labelling signal duration is also produced by the sigmoid binary activation function. Fig. 1(c) shows the modulus flow for DDTN. The noisy signal in the time domain inputs into the DDTN, which produce the mask $\overline{M_{\mathrm{Dn}}(t)}$ and label $\overline{L_{\mathrm{Dt}}(t)}$ consecutively. The estimated mask $\overline{M_{\mathrm{Dn}}(t)}$ is then applied to the noisy signal $NS(t)$ for obtaining the denoised signal, and the estimated $\overline{L_{\mathrm{Dt}}(t)}$ labels the duration of the microseismic signal. In the label vector $\overline{L_{\mathrm{Dt}}(t)}$, the value for the real signal duration is set equal to 1 and the rest to 0. The duration denoted

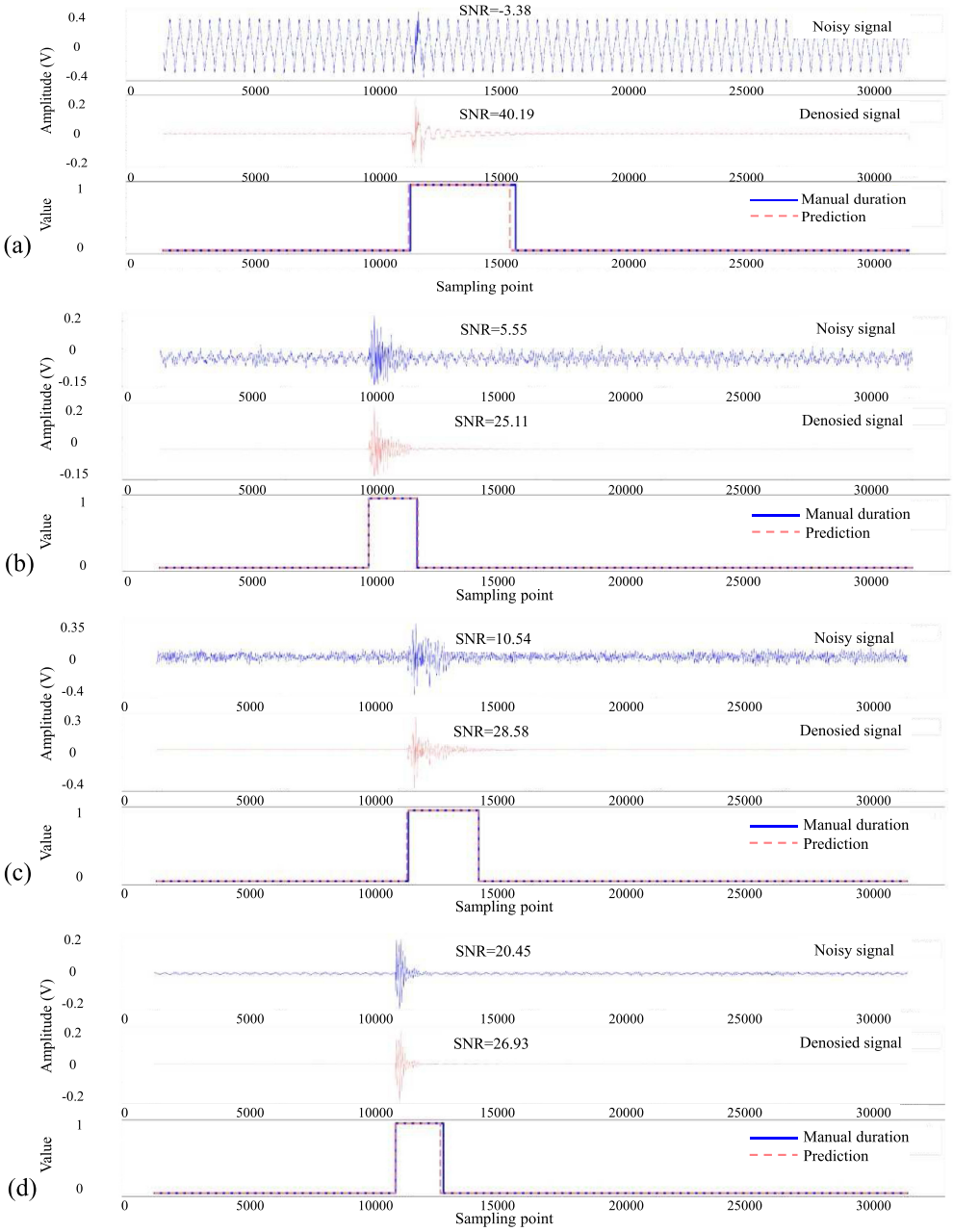

Figure 5. The combined performance of denoising and duration detection for various noisy signals in the test data set. The first two signals are a noisy signal and a denoised signal, respectively; the third signal includes the manual and predicted duration of microseismic signals. (a) SNR = −3.38; (b) SNR = 5.5; (c) SNR = 10.54; (d) SNR = 20.45.

by the blue curve in Fig. 1(c) is determined manually based on sampling point difference d between the first arrival and maximum amplitude of signals. The length from the first arrival point to the endpoint of the microseismic signal is set to $8d$ to represent the signal duration, which can well cover most of the microseismic waveforms and capture the main characteristics of signals.

One of the biggest advantages of DDTN is that instead of manually defining different features and thresholds to improve the SNR in microseismic signals, DDTN can automatically learn the richer features from the semisynthetic noisy signals to obtain the denoised signals and the duration of the microseismic signals. Two homo-structured encoder–decoder networks (denoising and detection networks) are integrated to form the structure of the DDTN. An input can produce two outputs in succession by DDTN. The structure is designed sequentially because the denoising output, as the input of the second training network, can facilitate the learning process and benefit the accuracy of the subsequent detection output. We can see that the first output $\overline{M_{\mathrm{Dn}}(t)}$ comes from the last layer of the denoising network, and it is seamlessly connected for feature extraction for the next detection network. Eventually, the second output $\overline{L_{\mathrm{Dt}}(t)}$ produced by the last layer of the detection network. This structure ensures that it is possible to train the two outputs in an integrative way. Thus, DDTN has a great potential to provide more efficient and accurate performance on denoising and duration detection of microseismic signals after sufficient training. Unlike existing

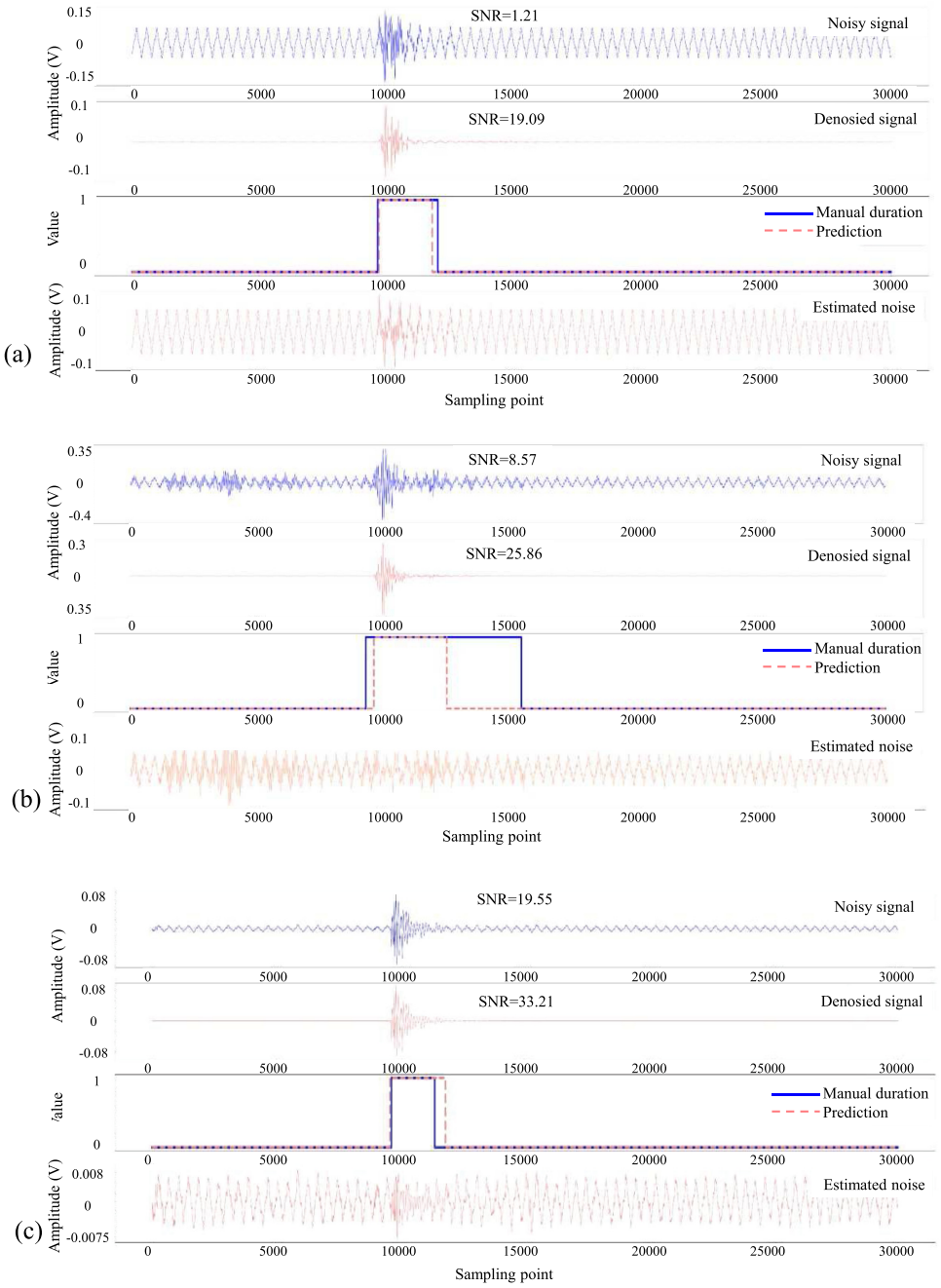

Figure 6. The combined performance of denoising and duration detection for noisy signals in Zijing tunnel (China). The first two signals are a noisy signal and a denoised signal, respectively, the third signal includes the manual and predicted duration of microseismic signals, and the fourth signal represents the estimated noise. (a) SNR = 1.21; (b) SNR = 8.57; (c) SNR = 19.55.

denoising methods, which involve the conversion between the time domain and other domains, the DDTN can directly perform the denoising and duration detection of microseismic signals in the time domain.

In this paper, the amplitude of the recorded microseismic signal is in voltage, and the response frequency ranges from 50 Hz to 5 kHz. The data acquisition station worked at a sampling frequency of 20 kHz and a sampling window of 1.5 s, which results in all signals having the size of 30 000 sampling points. 7500 microseismic signals with high SNR levels and 15 000 noises are selected to form a data set of 35 000 semisynthetic samples, which is randomly split into training (80 per cent), validation (10 per cent), and test (10 per cent) data sets. The validation set acts to fine-tune the hyperparameters and prevent overfitting of the network to achieve the best results, and the test set is primarily used to document network performance. To generate the input 'noisy signal' for training, we iterate through the training data set and generate the noisy signal with different SNR levels by superimposing the selected microseismic signals with randomly selected noise

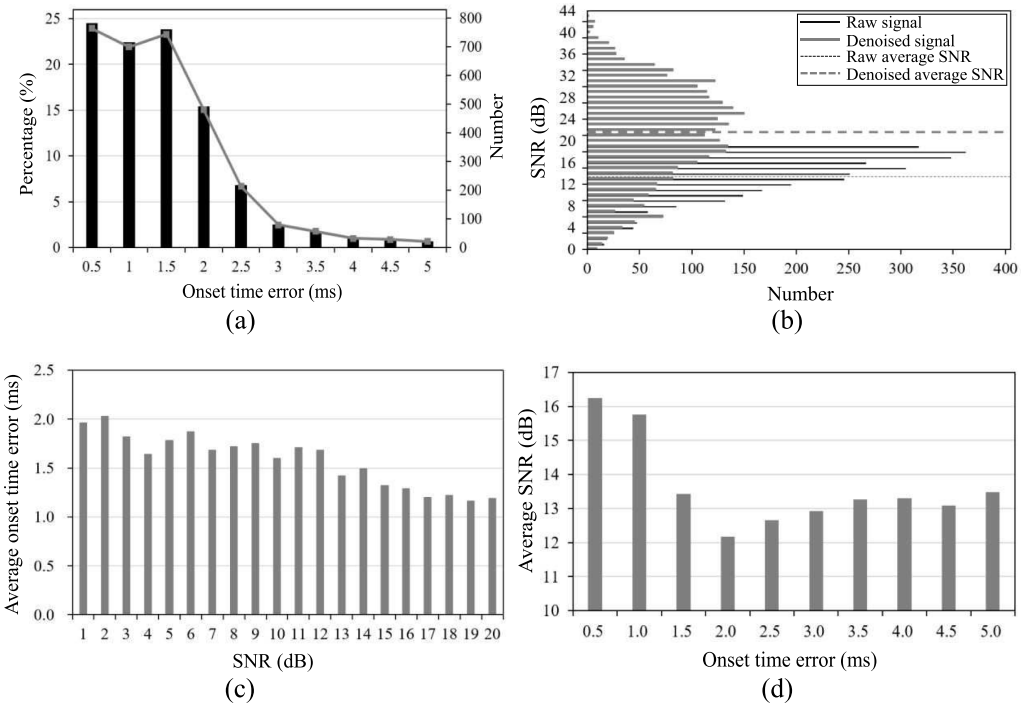

Figure 7. (a) Error distribution of onset time picking between the DDTN and manual. (b) Distribution of SNRs of noisy signals by DDTN. (c and d) Relationship between SNR and onset time picking error of noisy signals.

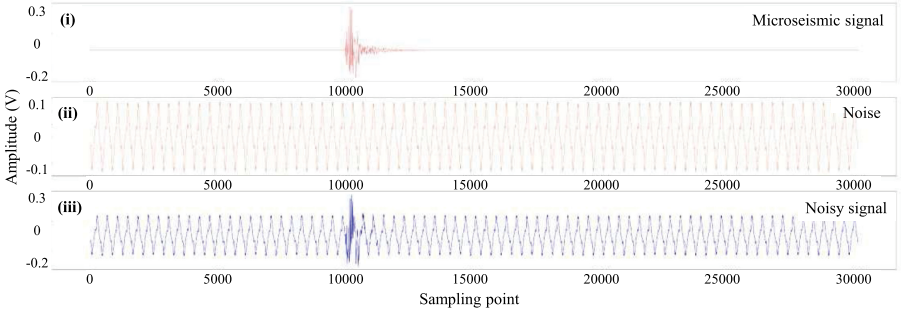

Figure 8. Noisy signals (iii) are formed with clean microseismic signals in (i) and noise in (ii).

samples on each iteration. The network is trained using NVIDIA GTX 2080Ti GPU, and Adam optimizer with the learning rate of 0.001, in batches of 32 noisy signals.

To obtain the optimal neural network structure, the influence of network depth (24, 32 and 48) is comprehensively examined using the validation data set and the results are presented in Table 1. Since it is the core of the whole integrated network, the changes of depth are only made in the denoising network for simplicity, which would help to clarify the variation of its performance and the influence on the final result after processing of the detection network. Results show that the increase of depth of layers and complexity will enhance the performance of the denoising and detection network yet bring more training variables and computational cost (i.e. high FLOPS, Floating-point-operations-per-second). It can be observed in Table 1 that the numbers of variables and FLOPs almost linearly increased from depth 24 to 32 and 32 to 48. However, the increase of accuracy is relatively large from 24 to 32 (0.050 in denoising) but becomes marginal from 32 to 48 (0.007 in denoising). Considering the computational complexity, memory consumption, and the performance of the two components (denoising and detection networks), the depth of 32 for DDTN (i.e. DDTN-32, as shown in Fig. 1) is adopted in this study.

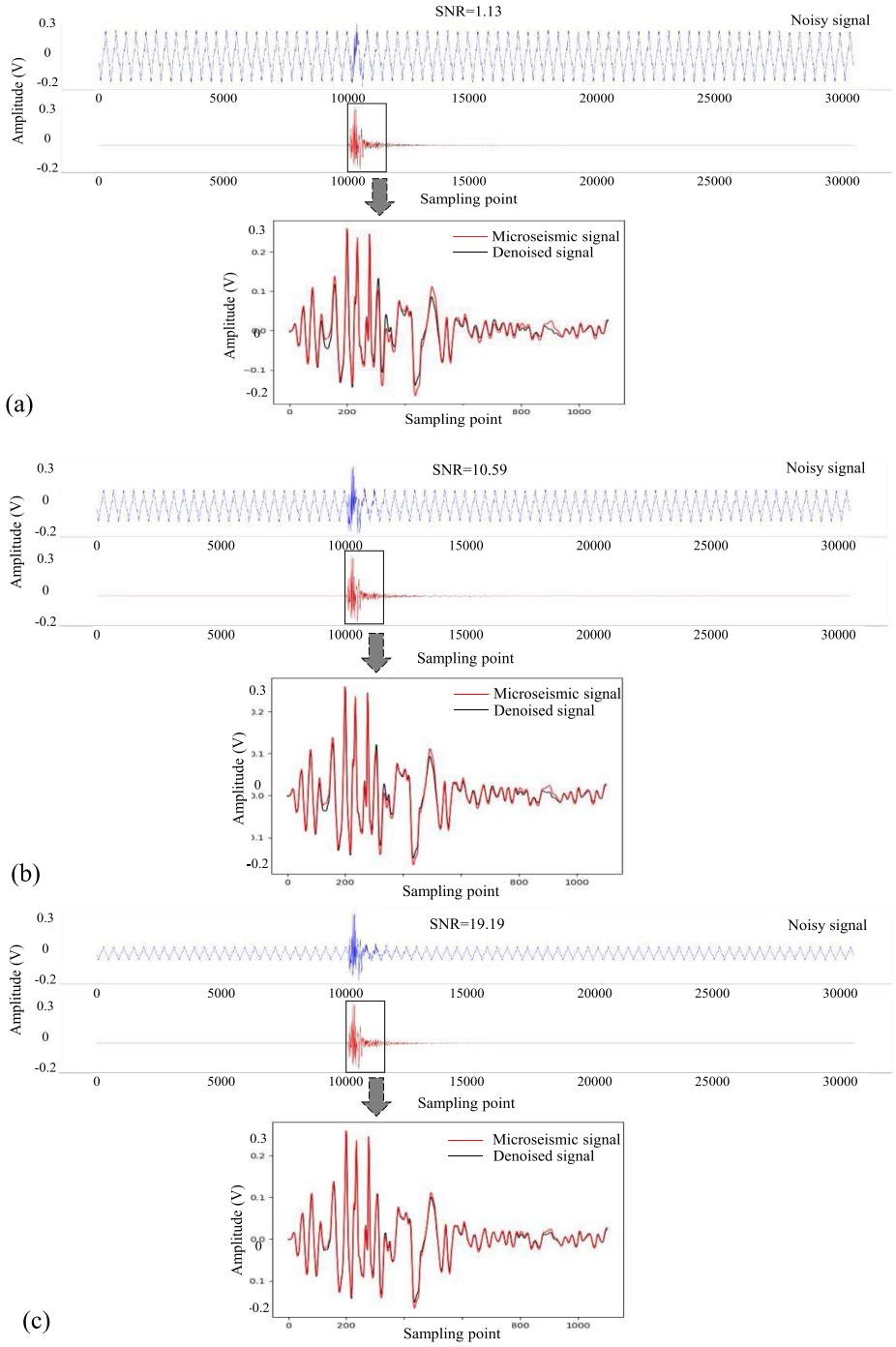

Figure 9. Denoising performance comparison between denoised signal by DDTN and microseismic signal of Fig. 8(i). SNR value represents the SNR of the noisy signal before denoising in (a)–(c).

3 RESULTS

3.1 Test results

Denoising results produced by the output mask $\overline{M_{\mathrm{Dn}}(t)}$ on signals with an increased SNR are shown in Fig. 2. The value of SNR is calculated by eq. (5) (Mousavi *et al.* 2019):

$$\mathrm{SNR} = 20 * \log_{10}(S_A/N_A), \tag{5}$$

where S_A and N_A are peak amplitudes of microseismic signal and noise, respectively. The noise components used to generate the synthetic signal in each case are collected from the real field. In Figs 2(a) and (b) are microseismic signals with a different frequency of cyclic noise;

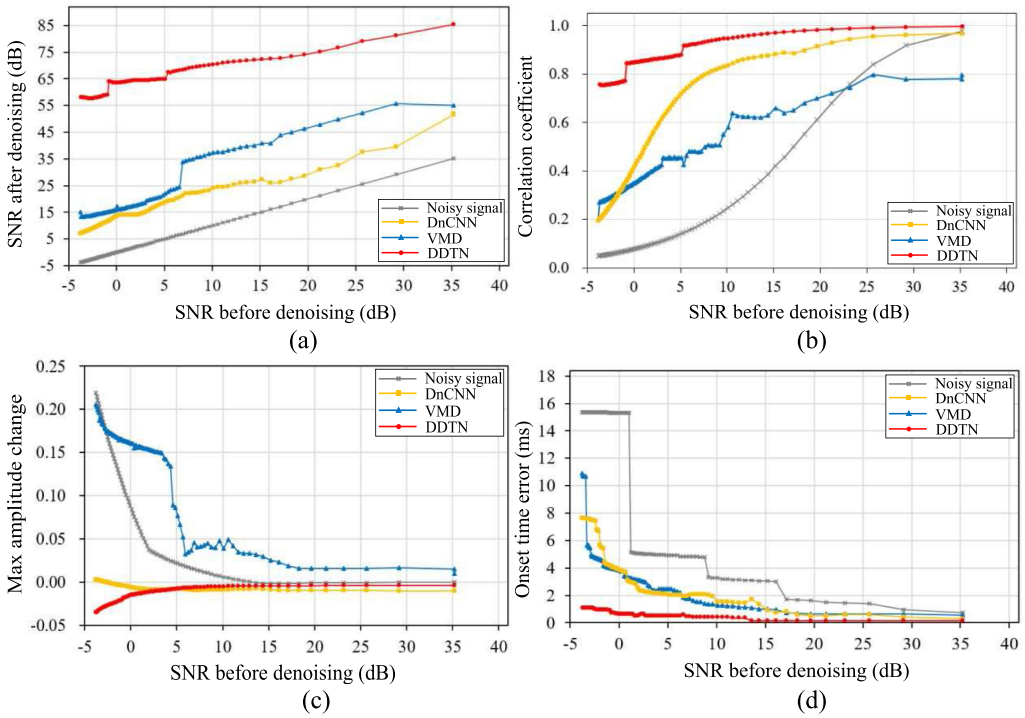

Figure 10. Performance comparison between VMD and DnCNN combined with STA/LTA and DDTN. (a–c) Comparison of SNR improvements, correlation coefficient and max amplitude change between the noisy microseismic signal and that processed by VMD, DnCNN and DDTN. (d) Onset time errors between STA/LTA (applied to noisy signal and that after denoising by VMD and DnCNN) and DDTN. Values in (b) are calculated by the Pearson product-moment correlation coefficient (Mantovani *et al.* 1986) between the signals (noisy signal and that denoised by VMD, DnCNN and DDTN) and the microseismic signal in Fig. 8(i). Values in (c) and (d) are differences between signals (noisy signal and that processed by VMD, DnCNN and DDTN) and the microseismic signal in Fig. 8(i).

Table 2. The comparison of onset time picking between different methods.

Methods	Precision	Recall	F1-score
VMD	0.444	1.000	0.615
DnCNN	0.522	1.000	0.686
DDTN	1.000	1.000	1.000

Figs 2(c) and (d) contain a mixture of different cyclic and other noises caused by unknown sources. Especially, the noisy signal in Fig. 2(d) includes more than one microseismic waveform to represent a more challenging situation for the denoising network of DDTN. These types of synthetic signals well represent most of the polluted data collected from the real engineering field. Results show that the improvement of the SNR and the performance of denoising are significant for all four tests; even noisy signal in Fig. 2(d) includes multiple microseismic waveforms. The noisy signals with different characteristics can be successfully separated into denoised signals and estimated noise. Moreover, the denoised signal leakage is minimal, and the shape and amplitude characteristics of the denoised signal are well preserved. These above characteristics are also useful for estimating the noise.

To further examine the limitation of the integrated network, DDTN is applied to denoise 15 776 real noise-only samples. Fig. 2(e) fully demonstrates the robustness of the denoising network when a pure periodic noise presents in the data. No denoised signals are predicted; and the estimated noise is almost equivalent to the input noise sample. Therefore, the DDTN has good denoising performance for microseismic signals containing various types of noise.

After processing the total noise samples of various types (including cyclic noise, unknown noise, and their combination), Fig. 3 shows the distribution of max amplitude difference between the noise sample and estimated noise. The results show that the max amplitude difference of more than 50 per cent noise samples is smaller than 0.005, and 90.59 per cent is smaller than 0.035, which means the proposed method only causes minor waveform distortion.

DDTN can be well applied in actual microseismic signals with non-Gauss noise (including cyclic noise, unknown noise, and their mixture) based on the previous test results. To further examine the ability of the algorithm for Gauss noise, DDTN is tested with noisy signals generated by superpositioning the clean microseismic signal and Gauss noise in Fig. 4. The results show that the improvement of the SNRs, shape recovery, the amplitude characteristics of the denoised signal, and the high accuracy of duration detection are well performed. And, the shape and amplitude characteristics of the estimated Gauss noise are also properly preserved.

The full performance produced by the integrated network of DDTN is provided in Fig. 5, which includes the results of denoising and duration detection of microseismic signals with different values of SNR and intensities of noise. The full result further proves that the duration of the microseismic signals can be detected with high accuracy. It removes the need for manual picking from the workflow of the microseismic data processing. This high accuracy is guaranteed by the denoising network, which provides the input of the second part of the DDTN (duration packing). Therefore, this integration design of the structure maximizes the advantages of using the neural network.

3.2 Application in real projects

The DDTN is tested with 3118 real-field noisy signals recorded in Zijing tunnel (China). These noisy signals included various types and intensities of noise so that the SNR levels ranged from 0 to 20 dB. The test results in Section 3.1 show that the proposed method only causes less waveform distortion, which means the peak amplitude of microseismic signal after denoising is approximately equal to the clean microseismic signal. Therefore, the value of the S_A in eq. (5) is set to the peak amplitude of the real microseismic signal after denoising, and the N_A is the peak amplitudes of the noise in SNR calculation. Figs 6(a)–(c) show that the DDTN has a much better performance on denoising and duration detection of noisy microseismic signals, which is reflected in the improvement of the SNRs, good shape recovery, preserved the amplitude of the denoised signal, and the high accuracy of duration detection. The shape and amplitude characteristics of the estimated noise are also well preserved. Also, the DDTN has a good ability to detect the duration of a highly polluted signal duration that often difficult to be determined manually. For example, in Fig. 6 (b), the accuracy of duration detection is improved for nearly an entire duration band comparing to the manually picked duration.

After detecting the microseismic signal duration, the onset time can be picked, which is essential for microseismic source location. The error distribution of onset time picking of 3118 noisy signals in Fig. 7(a) shows that the average error is 1.39 ms, and 70.82 per cent of these noisy signals have an error smaller than 1.5 ms; and 93.07 per cent smaller than 2.5 ms. Fig. 7(b) reveals that the average SNR of the 3118 noisy signals increases by 7.54 dB, and the maximum improvement reaches 34.09 dB. The 3118 noisy signals are also divided into 20 types in Fig. 7(c) with SNRs ranging from 0 to 20 dB to study the relationship between SNR levels and errors of onset time picking. The maximum of the average error of the low SNR noisy signal is 2.04 ms, as shown in Fig. 7(c), which is not much different from the minimum error 1.2 ms of the high SNR signal. It means the DDTN has a good tolerance for noisy signals with low SNRs, and the accuracy of onset time picking can be better guaranteed. The higher the average SNR of the noisy signals, the lower the onset time picking error (Fig. 7d). Although the DDTN is trained on semisynthetic data, it could be well extended to real noisy signals and can be directly applied to actual engineering for denoising and onset time picking.

4 COMPARISON STUDIES

The influence of the denoising network on the waveform is first examined in this section, denoising performance in the time domain of the proposed method, a noisy signal with the same size of DDTN inputs is constructed using a clean microseismic signal and recorded cyclic noise (Fig. 8). Fig. 9 shows the comparison between the denoised signal by DDTN and the clean microseismic signal in Fig. 8(i). It can be found that the signal denoised by DDTN still keeps the waveform similar to the clean microseismic signal. When the noisy signals are at low SNRs before denoising, the shape and amplitude characters of the denoised signal are still mostly preserved. It indicates that the DDTN conduct the signal denoising in the time domain without distortion of waveform information.

To further demonstrate the superiority of the proposed method, noisy signals with different SNRs are formed by scaling the noise to vary its amplitude in Fig. 8. The VMD, DnCNN (Denoising convolutional neural network), and STA/LTA methods are used to compare the performance of the denoising and onset time picking of the method proposed in this paper. For the VMD method, the setting of parameters is tested based on the semisynthetic signals (Fig. 8) to achieve the good performance of denoising by monitoring the SNR improvement. The DnCNN was well trained based on the same training samples of the DDTN by evaluating the metrics on loss and accuracy of the model, and the SNR improvement was used to test the performance of denoising. Meanwhile, the threshold of the STA/LTA method was determined by minimizing the errors of onset time picking of the semisynthetic signals (Fig. 8).

Compared with the VMD and DnCNN, the improvements of SNR provided by the DDTN are the largest with a maximum SNR improvement of 64.85 dB (Fig. 10a). The highest correlation coefficient of denoised signals indicates that the DDTN causes less waveform distortion. In contrast, the VMD method introduces relatively high distortion of the waveform during denoising, so that the maximum correlation coefficient only reached 0.79 (Fig. 10b). Fig. 10(c) shows that the max amplitude changes of the denoised signals by the DnCNN are better than the DDTN when the SNR is less than 3 dB. However, the DDTN is superior to DnCNN in terms of the max amplitude changes of noisy signals with the increase of SNR levels, indicating that the amplitude characteristics of the denoised signal calculated by DDTN are closer to the clean microseismic signal in Fig. 8(i). The STA/LTA method is applied to pick the onset time of the noisy signal and that denoised by the VMD and DnCNN. It can also be found that the accuracy of the onset time picking is significantly improved after denoising by the three methods (VMD, DnCNN, and DDTN). The accuracy becomes higher as the SNR level increases and the DDTN has the best performance (Fig. 10d). The DDTN can also achieve sufficient accuracy regarding the onset time picking even when the SNR level is low, which also represents an improvement compared with the STA/LTA method.

The hit-rate and average deviation are also used to evaluate the performance of onset time picking. One hit is defined as one successful onset time picking if the absolute error of onset time picking between the DDTN and manual results is less than 2.5 ms, which is set based on the sampling frequency (20 kHz). A higher hit-rate represents a more successful onset time picking. The average deviation reflects the accuracy of onset time picking, i.e. the lower average deviation brings the smaller error of onset time picking. The hit-rate and average deviation are calculated by eqs (6) and (7):

$$\text{hit-rate} = \frac{N_{\text{hit}}}{N_{\text{pick}}} \times 100\,\text{per cent}, \tag{6}$$

$$\text{avgd} = \frac{\sum_{i=1}^{N_{hit}} d_i}{N_{hit}}, \tag{7}$$

where N_{hit} and N_{pick} represent the number of hit and onset time picking, and d_t represents the absolute error of onset time picking. The calculated hit-rates for noisy signals, noisy signals by VMD and DnCNN, and by DDTN are 36.2, 44.4, 52.2 and 100 per cent, respectively. The average deviations of that are 2.36, 1.59, 1.63 and 0.45 ms, respectively. High hit-rate and low average deviation indicate that the DDTN provides the best accuracy for the onset picking.

Indexes of Precision, Recall, and F1-score are used to evaluate the performances of the model (Mousavi *et al.* 2019):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{9}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{10}$$

where TP, FP and FN denote true positives, false positives and false negatives, respectively. The value of *Precision* is defined as the ratio of true positives and total of predicted positives; *Recall* is defined as the ratio of true positive and total real positives; F1-score combines the above two variables to eliminate the effects of unbalanced sample size. In our test, a pick is considered as a true positive when the onset time picking within $\pm$ 0.25 ms of the true arrival. All the tested signals have distinctively different SNR and contain various types of noise. Results in Table 2 show that the DDTN increases the both of the three indexes for the onset time picking, which means the denoised signal can be detected by the methods with 100 per cent correction. The combination of VMD or DnCNN, adaptive methods, with STA/LTA fails to achieve such a good accuracy.

The DDTN shows its robustness by maintaining an optimal accuracy under the training of complicated semisynthetic data, and generally reach much higher accuracy of duration detection and onset time picking compared to the manually picked duration. However, in some case, the accuracy of onset time picking cannot be guaranteed (see Fig. 5b), and about 6.93 per cent of the samples have the onset time error greater than 2.5 ms in the data set of a real project (Section 3.2), which did not meet a successful hit. For untrained new samples of noise or seismic signals, the method may not achieve performance as good as the presented results in the paper, which needs to be further studied with the enlargement of the real-field database in the future. The current estimation of the microseismic signal is based on the output of masks and non-mask prediction may be the direction of the next research (Zhu *et al.* 2019). The increase of denoising and training data can continuously improve the accuracy of onset time picking, which will be the goal of the next research. Furthermore, it is also possible to combine traditional methods and neural networks. For example, a neural network could be used to determine the possible range of onset time picking, and then the specific onset time can be picked up by using AIC, Kurtosis methods, etc. (Zhao *et al.* 2019).

5 CONCLUSION

This paper developed an advanced processing method based on the deep neural network for microseismic signals with optimized efficiency and accuracy. The method integrated two homo-structured encoder–decoder networks to solve the denoising and duration detection of the microseismic signal in the time domain. The signal and noise components in the data are properly recognized and separated, even if the signal is heavily polluted by noise. Denoising further improves the accuracy of duration detection of signals to meet the engineering requirements of magnitude estimation and onset time picking. Compared with the existing methods, this method significantly improves the SNRs and introduces less distortion in the waveform which allows better recovery of the real waveform. The method maintains a good ability for onset time picking even noisy signals at low SNR levels, which perform high hit-rate and low average deviation. While this study is motivated by the need for efficient and automated microseismic signal processing, it should be noted that the proposed method can be seamlessly extended to signal analysis for disaster estimation in geophysical and geotechnical fields; such as hydraulic fracturing, mining industry, shale-gas exploitation, and earthquake.

ACKNOWLEDGEMENTS

This work is financially supported by the National Natural Science Foundation of China (grant numbers 41807255, U19A20111 and 51808458), State Key Laboratory of Geohazard Prevention and Geoenvironment Protection Independent Research Project (grant number SKLGP2020Z010) and Science and Technology Project of Sichuan (grant number 2019YJ0465).

REFERENCES

Allen, R.V., 1978. Automatic earthquake recognition and timing from single traces, *Bull. seism. Soc. Am.,* **68**(5), 1521–1532.

Bekara, M. & Baan, M.V.D., 2009. Random and coherent noise attenuation by empirical mode decomposition, *Geophysics,* **74**(5), 89–98.

Bonar, D. & Sacchi, M., 2012. Denoising seismic data using the nonlocal means algorithm, *Geophysics,* **77,** A5–A8.

Boschetti, F., List, R.D. & Dentith, M.D., 1996. A fractal-based algorithm for detecting first arrivals on seismic traces, *Geophysics,* **61**(4), 1095–1102.

Chen, W., Xie, J.Y., Zu, S.H., Gan, S.W., & Chen, Y.K., 2017. Multiple-reflection noise attenuation using adaptive randomized-order empirical mode decomposition, *IEEE Geosci. Remote Sens. Lett.,* **14**(1), 18–22.

Chen, Y., Ma, J. & Fomel, S., 2016. Double-sparsity dictionary for seismic noise attenuation, *Geophysics,* **81**(2), V17–V30.

Diehl, T., Deichmann, N., Kissling, E. & Husen, S., 2009. Automatic S-wave picker for local earthquake tomography, *Bull. seism. Soc. Am.,* **99**(3), 1906–1920.

Dong, X.T, Li, Y., & Yang, B.J., 2019. Desert low-frequency noise suppression by using adaptive DnCNNs based on the determination of high-order statistic, *Geophys. J. Int.,* **219**(2), 1281–1299.

Dong, X.T, Zhong, T. & Li, Y., 2020. New suppression technology for low-frequency noise in desert region: the improved robust principal component analysis based on prediction of neural network, *IEEE Trans. Geosci. Remote Sens.,* **58**(7), 4680–4690.

Galiana-Merino, J.J., Rosa-Herranz, J., Giner, J., Molina, S. & Botella, F., 2003, De-noising of short-period seismograms by wavelet packet transform, *Bull. seism. Soc. Am.,* **93,** 2554–2562.

Gentili, S. & Michelini, A., 2006. Automatic picking of P and S phases using a neural tree, *J. Seismol.,* **10**(1), 39–63.

Goodfellow, I., Bengio, Y., & Courville, A., 2016. *Deep Learning,* Vol. 1, MIT Press.

Hashemi, H., Javaherian, A., & Babuska, R., 2008. A semi-supervised method to detect seismic random noise with fuzzy GK clustering, *J. Geophys. Eng.,* **5,** 457–468.

Huang, N.E. *et al.*, 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. A,* **454,** 903–995.

Ioffe, S., & Szegedy, C., 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift, in *Proceedings of the 32nd International Conference on Machine Learning,* Lille, France, pp.448–456.

Jia, R.S., Tan, Y.L., Sun, H.M., & Hong, Y.F., 2015. Method of automatic detection on micro-seismic P-arrival time under low signal-to-noise ratio, *J. China Coal Soc.,* **40**(8), 1845–1852. (in Chinese)

Li, H., Wang, R., Cao, S., Chen, Y. & Huang, W., 2016. A method for low frequency noise suppression based on mathematical morphology in microseismic monitoring, *Geophysics,* **81**(3), V159–V167.

Liu, S.C., & Xun, C., 2014. Seismic signals wavelet packet denoising method based on improved threshold function and adaptive threshold, *Comput. Modelleling New Technol.,* **18,** 1291–1296.

Ma, H.T., Yan, J., & Li, Y. 2020. Low-frequency noise suppression of desert seismic data based on variational mode decomposition and low-rank component extraction, *IEEE Geosci. Remote Sens.,* **17**(2), 337–341.

Mantovani, E., Albarello, D. & Mucciarelli, M., 1986. Seismic activity in North Aegean Region as middle-term precursor of Calabrian earthquakes, *Phys. Earth planet. Inter.,* **44,** 264–273.

Mao, X.J., Shen, C.H., & Yang, Y.B., 2016. Image restoration using very deep convolutional encoder–decoder networks with symmetric skip connections, in *Proceedings of Advances in Neural Information Processing Systems (NIPS)*, Barcelona, Spain, p.221.

Mccomack, M.D., Zaucha, D.E. & Dushek, D.W., 1993. First-break refraction event picking and seismic data trace editing using neural networks, *Geophysics,* **58**(1), 67–78.

Mousavi, S.M. & Langston, C.A., 2016a. Hybrid seismic denoising using higher-order statistics and improved wavelet block thresholding, *Bull. seism. Soc. Am.,* **106,** 1380–1393.

Mousavi, S.M. & Langston, C.A., 2016b. Adaptive noise estimation and suppression for improving microseismic event detection, *J. Appl. Geophys.,* **132,** 116–124.

Mousavi, S.M., & Langston, C.A., 2016c. Fast and novel microseismic detection using time–frequency analysis, in *86th Annual International Meeting,* SEG, Expanded Abstracts, pp. 2632–2636.

Mousavi, S.M., Zhu, W.Q., Sheng, Y.X. & Beroza, G.C., 2019. CRED: a deep residual network of convolutional and recurrent units for earthquake signal detection, *Sci. Rep.,* **9,** 10267.

Oropeza, V. & Sacchi, M., 2011. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis, *Geophysics,* **76**(3), V25–V32.

Perol, T., Gharbi, M. & Denolle, M., 2018. Convolutional neural network for earthquake detection and location, *Sci. Adv.,* **4**(2), e1700578.

Ronneberger, O., Fischer, P. & Brox, T., 2015. U-Net: convolutional networks for biomedical image segmentation, *Springer, Cham*.

Ross, Z.E., Meier, M.A., & Hauksson, E., 2018. P-wave arrival picking and first-motion polarity determination with deep learning, *J. geophys. Res.,* **123**(6), 5120–5129.

Saragiotis, C.D., Hadjileontiadis, L.J., RekanoS, I.T., & Panas, S.M., 2004. Automatic P phase picking using maximum kurtosis and statistics criteria, *IEEE Geosci. Remote Sens. Lett.,* **1**(3), 147–151.

Takanami, T. & Kitagawa, G., 1991. Estimation of the arrival times of seismic waves by multivariate time series model, *Ann. Inst. Stat. Math.,* **43**(3), 407–433.

Tselentis, G.A., Martakis, N., Paraskevopoulos, P., Lois, A. & Sokos, E., 2012. Strategy for automated analysis of passive microseismic data based on s-transform, Otsu's thresholding, and higher order statistics. *Geophysics,* **77**(6), KS43–KS54.

Yuan, S.Y., Liu, J.W., Wang, S.X., Wang, T.Y., & Shi, P.D., 2018. Seismic waveform classification and first-break picking using convolution neural networks, *IEEE Geosci. Remote Sensi. Lett.,* **15**(2), 1–5.

Yung, S.K. & Ikeller, L.T., 1997. An example of seismic time-picking by third-order bicoherence, *Geophysics,* **62**(6), 1947–1951.

Zhang, H., 2003. Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single component recordings, *Bull. seism. Soc. Am.,* **93**(5), 1904–1912.

Zhao, D.P., Liu, X.Q., Li, H. & Zhou, Y., 2012. Detection of regional seismic events by Kurtosis method and automatic identification of direct P-wave first motion by Kurtosis-AIC method, *J. Seismol. Res.,* **35**(2), 220–225. (in Chinese)

Zhao, M., Chen, S., Fang, L.H., & Yuen, D.A., 2019. Earthquake phase arrival auto-picking based on U-shaped convolutional neural network, *Chin. J. Geophys.,* **62**(8), 3034–3042. (in Chinese)

Zheng, J., Lu, J., Peng, S. & Jiang, T., 2018. An automatic microseismic or acoustic emission arrival identification scheme with deep recurrent neural networks, *Geophys. J. Int.,* **212**(2), 1389–1397.

Zhu, W.Q. & Beroza, G.C., 2018. Phasenet: a deep-neuralnetwork-based seismic arrival time picking method, *Geophys. J. Int.,* **216**(1), 261–273.

Zhu, W.Q., Mousavi, S.M. & Beroza, G.C., 2019. Seismic signal denoising and decomposition using deep neural networks, *IEEE Trans. Geosci. Remote Sens.,* **57**(11), 9476–9488.