Identification of Seismic Wave First Arrivals from Earthquake Records via Deep Learning

Yang Yu[1], Jianfeng Lin[1], Lei Zhang[2], Guiquan Liu[1](✉), Jing Hu[3], Yuyang Tan[3], and Haijiang Zhang[3]

[1] School of Computer Science and Technology, University of Science and Technology of China, Hefei, China
{koinu,linjf}@mail.ustc.edu.cn, gqliu@ustc.edu.cn
[2] School of Computer Science and Technology, Anhui University, Hefei, China
zl@ahu.edu.cn
[3] School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China
jhsa26@mail.ustc.edu.cn, {tanyy0,zhang11}@ustc.edu.cn

Abstract. For seismic location and tomography, it is important to pick P- and S-wave first arrivals. However, traditional methods mainly determine P- and S-wave first arrivals separately from a signal processing perspective, which requires the extraction of waveform attributes and tuning parameters manually. Also, traditional methods suffer from noise as they are operated on the whole earthquake record. In this paper, we propose a deep neural network framework to enhance picking P- and S-wave first arrivals from a sequential perspective. Specifically, we first transform the picking first arrival problem as a sequence labelling problem. Then, the rough ranges for P- and S-wave first arrivals are determined simultaneously through the proposed deep neural network model. Based on these rough ranges, the performance of existing picking methods can be greatly enhanced. Experimental results on two real-world datasets demonstrate the effectiveness of the proposed framework.

Keywords: Wave first arrivals · Sequence labelling · Deep learning

1 Introduction

An earthquake record consists of three components, recording waveform under different directions. When an earthquake occurs, there are two kinds of seismic wave traveling inside the earth, known as P-wave and S-wave. P-wave always arrives earlier than S-wave. Figure 1 gives a three-component earthquake record sliced from continuous waveform data. The problem we focused on in this paper is to effectively identify P- and S-wave first arrivals from earthquake records.

Many automatic picking methods were proposed [3,12]. These methods focus on detecting waveform samples where various waveform attributes change significantly. However, these methods suffer from three main challenges: First, these

© Springer Nature Switzerland AG 2018
W. Liu et al. (Eds.): KSEM 2018, LNAI 11062, pp. 274–282, 2018.
https://doi.org/10.1007/978-3-319-99247-1_24

methods are generally sensitive to noise. They tend to perform poorly when processing low signal-to-noise ratio (SNR) data. Second, manually designed waveform attributes may fail to fully utilize information contained in sequences. Third, these methods usually need a lot of effort on tuning parameters manually. With the development of neural networks, there have been some automatic picking methods utilizing Artificial Neural Networks (ANNs) [8]. However, these ANN methods still need designing input attributes. Besides, these methods do not make full use of sequential characteristic in the earthquake data.

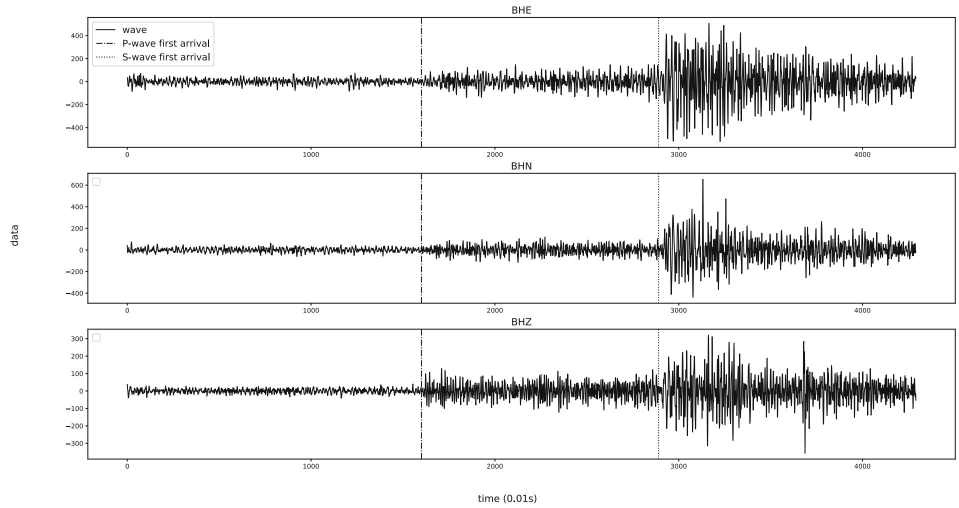

Fig. 1. A three-component earthquake record.

To solve the above challenges, in this paper, we propose a deep neural network framework to enhance picking P- and S-wave first arrivals from earthquake records. To be specific, the picking first arrival problem is transformed as a sequence labelling problem. Then, we use the proposed deep neural network model for determining the rough ranges of P- and S-wave first arrivals simultaneously. By using these rough ranges, the performance of existing picking methods can be greatly enhanced since the influence of noise can be reduced and effort of tuning parameters can be saved a lot as the ranges of parameters like window size have been narrowed. In summary, we have made several contributions in this paper:

- We transform the picking first arrival problem into a sequence labelling problem, with the aim to capture the temporal relationship of sequential data.
- We propose a general framework based on deep learning for the transformed problem. Specifically, our framework takes waveform data as input, produces a state sequence indicating the rough ranges for P- and S-wave first arrivals. Based on rough ranges, we can pick P- and S-wave first arrivals precisely.
- We have conducted extensive experiments on two real-world datasets to evaluate our framework. The results indicate that our framework outperforms

the state-of-art methods for identification of seismic wave first arrivals from earthquake records.

2 Related Works

2.1 Existing First Arrival Picking Methods

One classical picking method is the short term average to long term average ratio (STA/LTA) [3]. This method is based on comparison between a Short Term Average (STA) of a characteristic function of the signal and a Long Term Average (LTA) of this characteristic function. The point where the STA/LTA ratio exceeds the threshold value is regarded as wave first arrival.

Another classical method is autoregressive Akaike information criterion (AR-AIC) [12]. It is assumed that a seismogram can be divided into locally stationary segments, where each is modeled as an AR process and the intervals before and after the wave first arrival are two different stationary processes [12]. This method uses AIC function to find the point representing optimal separation of stationary time series which is interpreted as the wave first arrival [12].

In addition, there have been some attempts on picking first arrivals using ANNs [8]. Different attributes are manually designed as input to ANNs. It is worth noting that all these above works need designing attributes manually and ignore the sequential characteristic of earthquake record.

Therefore, in this paper, we apply deep neural networks to pick first arrivals from earthquake records. One work similar to the spirits of ours is [15]. In their work, they proposed a model of 7 Long Short-Term Memory (LSTM) [5] layers to pick first arrivals directly. However, the length of sequence can reach several thousand. Without performing dimensionality reduction, their method fails to pick first arrivals from earthquake records with long sequence.

2.2 Sequence Labelling

In order to utilize the sequential characteristic of earthquake records, we transform the picking first arrival problem as a sequence labelling problem.

Before widely use of neural networks, some graphical models have been proposed, such as hidden Markov Models (HMMs) [9], conditional random fields (CRFs) [6] and their variants.

With the development of neural networks, HHM-neural network hybrid approaches have been extensively studied [10,11]. The basic idea is to use HMMs to model the sequential structure of the data, and the neural networks to provide localized classifications [4]. Early hybrid methods used ANNs to achieve segment classification and HMMs to align the classification results into a temporal classification of the entire label sequence [10,11]. Since the main function of ANNs is to introduce contextual information, Recurrent Neural Networks (RNNs) [4] seem to be a better choice. Some HMM-RNN hybrid methods have been studied [11].

3 The Proposed Framework

Our proposed framework works as follows. First, we pass preprocessed three-component earthquake records to the proposed $PICKINGNET$ network to get rough P- and S-wave first arrival points. Then, we determine rough ranges based on rough first arrival points. Finally, we apply traditional methods on waveform sliced according to two rough ranges to pick first arrivals precisely. Fig. 2 presents the proposed framework.

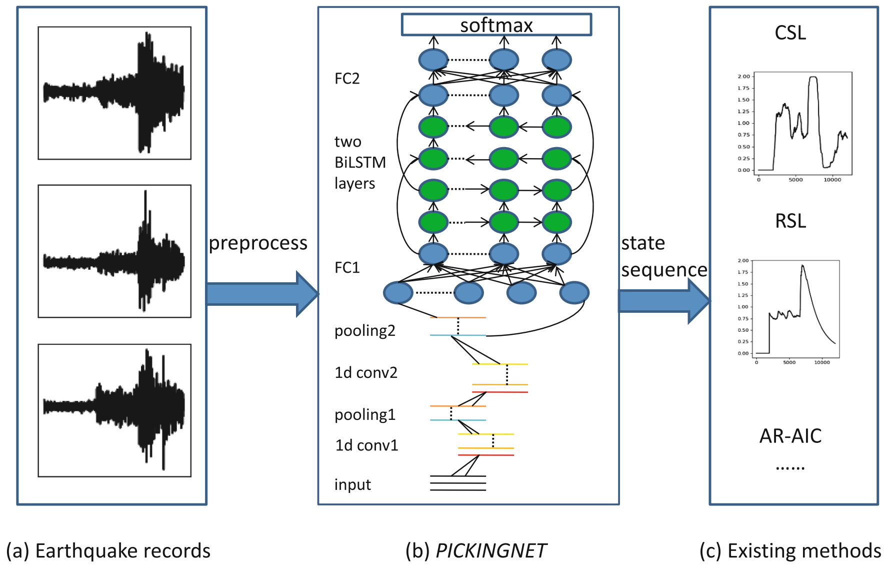

Fig. 2. Proposed framework.

3.1 Proposed Network Framework

The length of sequence can vary from hundreds to several thousand, it is hard to label the sequence directly to pick P- and S-wave first arrivals precisely. Therefore, we design the proposed $PICKINGNET$ according to the following thought. First, we perform dimensionality reduction on original sequence to get sequence s'. Every point in s' represents a segment of original sequence. Then, we model the temporal relationship of s' through RNNs. At last, we label every point of s' with three kinds of states, which are noise, P-wave and S-wave. The output of proposed network is a sequence of states of s', where every point indicates the corresponding segment of original sequence is noise, P-wave or S-wave. In this paper, we determine every point in s' represents a 1-second segment of original sequence. Our proposed $PICKINGNET$ is shown in Fig. 2(b).

We denote original sequence length as l. As input is three-component record, we can regard the dimensionality of input data as $l \times 3$.

Our proposed $PICKINGNET$ consists of two parts: (1) Dimensionality reduction; (2) Temporal modelling and labelling.

Dimensionality Reduction. We know CNNs [7] are good at extracting features. However, here we use CNNs to reduce the length of sequence as well. We use two 1d convolutional layers, each layer followed by a pooling layer. Specifically, *conv1* has 8 feature maps with the length of 4 and *conv2* has 16 feature maps with the length of 4. Both pooling window size and stride of pooling layers are 2 so that we can reduce the length of sequence.

The output dimensionality is $(l/4) \times 16$ where $l/4$ and 16 denote the length of sequence and the feature dimensionality of every point, respectively. With sampling rate of 100 Hz, to achieve every point in dimensionality reduced sequence represents a 1-second segment of original sequence, the sequence length should be reduced to $l/100$. Therefore, by concatenating features of every 25 points, the shape of output is $(l/100) \times 400$. Further, followed by the fully connected layer *FC1*, the feature dimensionality of every point is reduced to 100.

Temporal Modelling and Labelling. Then we pass the output of fully connected layer *FC1* to RNN layers. In this problem, we can access the past and future contexts of every point. Also, we know a general earthquake sequential record includes noise, P-wave, S-wave and noise in order. Therefore, we use two layers of Bidirectional Long Short-Term Memory (BiLSTM) [13] to model temporal relationship, where every LSTM cell has 100 units.

After temporal modeling, we pass the output of the BiLSTMs to the fully connected layer *FC2* followed by a softmax layer to label the sequence obtained from previous step. Every point of the output indicates the state of the corresponding 1-second segment of the original sequence. There are three kinds of states: noise, P-wave and S-wave.

3.2 Picking First Arrivals Precisely

With the state sequence produced by network, we pick up the first point labelled with p as the P-wave rough first arrival, the first point labelled with s as the S-wave rough first arrival. With rough first arrivals, we get the rough ranges for both P- and S-wave first arrivals by slicing waveform before and after the P- and S-wave rough first arrivals. We apply existing methods on waveform segments within two rough ranges to pick first arrivals precisely.

4 Experiment

4.1 Experimental Setup

Datasets. We evaluate the proposed approach on two real-world datasets. The first dataset contains continuous waveform data in August, 2008 collected by 16 seismic stations located on Sichuan Province, China and its neighboring province [16]. In this dataset, there are 14431 earthquake records. The second dataset contains continuous waveform data in August, 2014, after Napa earthquake in Northern California, USA [1]. In Napa dataset, there are 178 earthquake records. P- and S-wave first arrivals of both datasets are manually picked.

For the first dataset, we divide dataset into training set, validation set and test set with proportion of 60%, 10% and 30%. The labelled data of second dataset is not large enough for network training. To verify the transferability of our model, we use the network trained from Sichuan dataset to test performance on the Napa dataset.

Comparison Schemes. We use the rough ranges for P- and S-wave first arrivals to enhance three traditional first arrival picking methods, which are classical STA/LTA (CSL) [14], recursive STA/LTA (RSL) [14] and AR-AIC [12]. The common way is to employ these traditional methods on the whole earthquake record.

With the rough ranges, we enhance three traditional methods by changing slicing scheme. In this paper, we slice waveform of 3 s before and after P- and S-rough first arrivals as rough ranges. As a record has three components, we run existing methods and our proposed method on three components. We choose the earliest P- and S-wave first arrivals of three components as final picking results. All the input sequences have been preprocessed with detrending and bandpass filter [2].

Evaluation Criterion. Our metrics are *hit-rates* and *average deviations* of P- and S-wave picking performance, which are represented as as *hit-p*, *hit-s*, *avgd-p* and *avgd-s* respectively. We regard manually picking results as standard first arrivals. In this paper, we define a first arrival picking deviating from standard first arrival less than 1 s as a hit picking.

The hit-rate is the ratio of hit picking records to the all records, reflecting the success rate of picking performance. The average deviation is average picking deviation on hit picking records, reflecting the precision of picking performance.

Model Parameters. We add L2 regularization to *conv*1 and *conv*2 layers and use dropout and clip gradients tricks to prevent overfitting, the dropout rate is 0.3. The loss function we use is cross-entropy loss. The learning rate is set to be 0.0005. For convolutional layers and fully connected layers, we take ReLU as activation function. For BiLSTM layers, we take Tanh as activation function.

4.2 Experimental Results

Effectiveness of Proposed Method. Table 1 shows the comparison results of three baselines and our method on Sichuan test set in terms of *hit-rates* and *average deviations* of P- and S-wave picking performance. From Table 1, we can see that the performance of existing methods has been improved a lot with rough ranges. Without rough ranges, the performance of existing methods degenerates a lot with SNR declining, which confirms that our method is robust to noise.

Table 1. The comparison results of three baselines and our method on Sichuan test set in terms of *hit-rates* and *average deviations*, where RG denotes rough range.

SNR	Method	hit-p(+RG)	hit-s(+RG)	avgd-p(+RG)	avgd-s(+RG)
>20	CSL	96.22% (**99.11%**)	71.41% (**88.45%**)	0.3586 s (**0.3086 s**)	0.4561 s (**0.4179 s**)
	RSL	96.22% (**99.04%**)	51.27% (**89.20%**)	0.3359 s (**0.2762 s**)	0.4562 s (**0.4000 s**)
	AR-AIC	92.78% (**99.38%**)	71.62% (**91.07%**)	0.2455 s (**0.1771 s**)	0.3462 s (**0.2639 s**)
2–20	CSL	82.26% (**94.94%**)	53.78% (**79.39%**)	0.4315 s (**0.3862 s**)	0.5038 s (**0.4667 s**)
	RSL	80.89% (**93.38%**)	48.59% (**80.39%**)	0.4230 s (**0.3713 s**)	0.4886 s (**0.4511 s**)
	AR-AIC	86.26% (**96.88%**)	60.59% (**88.88%**)	0.3068 s (**0.2195 s**)	0.3760 s (**0.2854 s**)
<2	CSL	74.84% (**89.33%**)	40.81% (**72.12%**)	0.4511 s (**0.3994 s**)	0.5472 s (**0.5025 s**)
	RSL	72.51% (**89.64%**)	40.97% (**73.21%**)	0.4492 s (**0.3958 s**)	0.5414 s (**0.4837 s**)
	AR-AIC	79.91% (**93.61%**)	52.65% (**85.75%**)	0.3339 s (**0.2471 s**)	0.4189 s (**0.3119 s**)

Transferability of Proposed Method. Table 2 shows the result on Napa test set. Note that due to the lack of labelled data in Napa data set, here we use network obtained from Sichuan data set. We can see, with rough ranges, the performance of existing methods has been improved a lot except the *avgd-s* of AR-AIC. We believe it is because the *hit-s* of AR-AIC is low which means AR-AIC only picks S-wave first arrivals successfully on easily picked records, which results in a small *avgd-s* on hit picking records. The result confirms the transferability of our method, we believe it is because our network has captured the pattern of seismic waves.

Table 2. The comparison results of three baselines and our method on Napa test set in terms of *hit-rates* and *average deviations*, where RG denotes rough range.

Method	hit-p(+RG)	hit-s(+RG)	avgd-p(+RG)	avgd-s(+RG)
CSL	91.01% (**96.63%**)	75.28% (**87.64%**)	0.2134 s (**0.1583 s**)	0.3492 s (**0.2737 s**)
RSL	88.76% (**94.94%**)	53.93% (**93.26%**)	0.1814 s (**0.1530 s**)	0.3199 s (**0.2932 s**)
AR-AIC	89.89% (**96.63%**)	69.10% (**93.82%**)	0.0653 s (**0.0471 s**)	**0.1555 s** (0.2096 s)

5 Conclusion

In this paper, we studied the problem of picking P- and S-wave first arrivals from earthquake records. Unlike most existing neural network methods, we proposed a deep learning based framework for effectively picking P- and S-wave first arrivals. The proposed framework takes waveform data as input, transforms picking first arrival problem into a sequence labelling problem. Then we used the proposed deep neural network model $PICKINGNET$ for determining the rough ranges for P- and S-wave first arrivals simultaneously. By using these rough ranges, the performance of existing picking methods can be greatly enhanced. Our framework makes full use of the sequential characteristic of earthquake records and saves effort designing features manually. Finally, we demonstrated the effectiveness of our method on two real-world datasets. In future work, we plan to extend our approach on continuous seismic waveform data, with the aim to detect earthquakes and pick first arrivals simultaneously.

Acknowledgement. This work was supported in part by the Natural Science Foundation of China (Grant No. 61502001) and by the Academic and Technology Leader Imported Project of Anhui University (No. J01006057). The authors would like to thank Data Management Centre of China National Seismic Network at Institute of Geophysics, China Earthquake Administration and Northern California Earthquake Data Center (NCEDC) for providing waveform data for this study.

References

1. NCEDC: Northern California Earthquake Data Center, UC Berkeley Seismological Laboratory. Dataset (2014). https://doi.org/10.7932/ncedc
2. Akram, J.: Downhole microseismic monitoring: processing, algorithms and error analysis. Ph.D. thesis. University of Calgary (2014)
3. Allen, R.V.: Automatic earthquake recognition and timing from single traces. Bull. Seismol. Soc. Am. **68**(5), 1521–1532 (1978)
4. Graves, A., et al.: Supervised Sequence Labelling with Recurrent Neural Networks. Springer, Berlin (2012). https://doi.org/10.1007/978-3-642-24797-2
5. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
6. Lafferty, J., McCallum, A., Pereira, F.C.: Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of the 18th International Conference on Machine Learning, vol. 951, pp. 282–289 (2001)
7. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
8. Maity, D., Aminzadeh, F., Karrenbach, M.: Novel hybrid artificial neural network based autopicking workflow for passive seismic data. Geophys. Prospect. **62**(4), 834–847 (2014)
9. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE **77**(2), 257–286 (1989)
10. Renals, S., Morgan, N., Bourlard, H., Cohen, M., Franco, H.: Connectionist probability estimators in hmm speech recognition. IEEE Trans. Speech Audio Process. **2**(1), 161–174 (1994)
11. Robinson, A.J.: An application of recurrent nets to phone probability estimation. IEEE Trans. Neural Netw. **5**(2), 298–305 (1994)
12. Sleeman, R., van Eck, T.: Robust automatic p-phase picking: an on-line implementation in the analysis of broadband seismogram recordings. Phys. Earth Planet. Inter. **113**(1), 265–275 (1999)
13. Thireou, T., Reczko, M.: Bidirectional long short-term memory networks for predicting the subcellular localization of eukaryotic proteins. IEEE/ACM Trans. Comput. Biol. Bioinform. **4**(3), 441–446 (2007)
14. Withers, M., et al.: A comparison of select trigger algorithms for automated global seismic phase and event detection. Bull. Seismol. Soc. Am. **88**(1), 95–106 (1998)

15. Zheng, J., Lu, J., Peng, S., Jiang, T.: An automatic microseismic or acoustic emission arrival identification scheme with deep recurrent neural networks. Geophys. J. Int. **212**(2), 1389–1397 (2017)
16. Zheng, X.F., Yao, Z.X., Liang, J.H., Zheng, J.: The role played and opportunities provided by igp dmc of china national seismic network in wenchuan earthquake disaster relief and researches. Bull. Seismol. Soc. Am. **100**(5B), 2866–2872 (2010)