Artificial Intelligence in Geosciences 3 (2022) 115–122

Contents lists available at ScienceDirect

Artificial Intelligence in Geosciences

journal homepage: www.keaipublishing.com/en/journals/artificial-intelligence-in-geosciences

Original research articles

A study on small magnitude seismic phase identification using 1D deep residual neural network

Wei Li [a], Megha Chakraborty [a,b], Yu Sha [c,a,d], Kai Zhou [a,d,e], Johannes Faber [a,e], Georg Rümpker [a,b], Horst Stöcker [a,d,e,f], Nishtha Srivastava [a,b,*]

[a] Frankfurt Institute for Advanced Studies, Frankfurt am Main, 60438, Germany
[b] Institute of Geosciences, Goethe-University Frankfurt, Frankfurt am Main, 60438, Germany
[c] Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Academy of Advanced Interdisciplinary Research, Xidian University, Xian, 710071, China
[d] Xidian-FIAS international Joint Research Center, Giersch Science Center, Frankfurt am Main, 60438, Germany
[e] Institut für Theoretische Physik, Goethe Universität Frankfurt, Frankfurt am Main, 60438, Germany
[f] GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt, 64291, Germany

ARTICLE INFO

Dataset link: https://github.com/srivastavareserarchgroup/Seismic-phase-Classification

Keywords:
Deep learning
Residual neural network
Earthquake detection
Seismic phase identification

ABSTRACT

Reliable seismic phase identification is often challenging especially in the circumstances of low-magnitude events or poor signal-to-noise ratio. With improved seismometers and better global coverage, a sharp increase in the volume of recorded seismic data has been achieved. This makes handling seismic data rather daunting by using traditional approaches and therefore fuels the need for more robust and reliable methods. In this study, we develop 1D deep Residual Neural Network (ResNet), for tackling the problem of seismic signal detection and phase identification. This method is trained and tested on the dataset recorded by the Southern California Seismic Network. Results demonstrate that the proposed method can achieve robust performance for the detection of seismic signals and the identification of seismic phases. Compared to previously proposed deep learning methods, the introduced framework achieves around 4% improvement in earthquake detection and a slightly better performance in seismic phase identification on the dataset recorded by Southern California Earthquake Data Center. The model generalizability is also tested further on the STanford EArthquake Dataset. In addition, the experimental result on the same subset of the STanford EArthquake Dataset, when masked by different noise levels, demonstrates the model's robustness in identifying the seismic phases of small magnitude.

1. Introduction

The detection of earthquakes is crucial for seismologists to monitor tectonic activities in a region. In order to achieve reliable earthquake monitoring, many automated methods for seismic phase identification have been developed. The most state-of-the-art conventional algorithms for earthquake detection or seismic phase picking include template matching (Peng and Zhao, 2009; Ross et al., 2017) and short-time average/long-time average (STA/LTA) (Zhou et al., 2022). Template matching involves measuring the similarity between earthquake waveforms and cataloged waveforms of seismic events. However, it relies heavily on the pre-defined events, which makes the detection challenging in case of unseen data (Ross et al., 2018). STA/LTA refers to measuring the ratio between the amplitude of the signal on a short time window and that on a long time window; an event is detected when the ratio exceeds some pre-defined threshold. STA/LTA might not perform well for the signals with low signal-to-ratio (SNR) and low magnitudes.

Over the past decades, due to the development of seismic equipment and seismic monitoring networks, remarkable improvements have been achieved in seismic event detection systems which brings about a huge and rapidly-increasing seismic database. This calls for robust and sensitive methods to process the ever-growing volume of seismic data. Therefore, seismic event detection and phase picking algorithms are becoming increasingly important to automatically deal with large amounts of seismic data. Deep learning, with its recent development — especially in computer vision, is capable of processing big data with a large number of different features such as lines, edges, and image segments. The task of earthquake signal detection or seismic phase identification can be recognized as similar to the identification of objects in computer vision. Therefore, recent advances in the field of computer vision have great potential for seismological applications. Recently, deep learning has been widely used to detect seismic events,

* Corresponding author at: Frankfurt Institute for Advanced Studies, Frankfurt am Main, 60438, Germany.
E-mail address: srivastava@fias.uni-frankfurt.de (N. Srivastava).

https://doi.org/10.1016/j.aiig.2022.10.002
Received 22 July 2022; Received in revised form 25 October 2022; Accepted 25 October 2022
Available online 4 November 2022

identify seismic phases, and estimate seismic phase arrival-time picking (Ross et al., 2018; Zhu and Beroza, 2019; Mousavi et al., 2020; Saad and Chen, 2020; Li et al., 2021; Chakraborty et al., 2021; Saad and Chen, 2021; Chakraborty et al., 2022). For example, in Ross et al. (2018), a convolutional neural network associated with fully-connected layers was trained to detect short windows of P-waves, S-waves, and noise. Capsule Neural Network-based earthquake signal detection and seismic phase identification was presented in Saad and Chen (2020, 2021). The authors in Li et al. (2021) combined the attention mechanism and U-shaped neural network for seismic phase arrival-time estimation. Both convolution layer and Bidirectional Long-Short Term Memory units were used in Chakraborty et al. (2021) and Chakraborty et al. (2022) for earthquake identification and magnitude estimation.

In convolutional neural networks (CNN)-based phase picking methods, the pooling layer is mainly used to reduce the network dimensionality. However, its use may result in the loss of some key features. As a result, the use of a pooling layer at the deeper network could increase the risk of more information loss, which may lead to a decrease in the network performance (Zeiler and Fergus, 2014; Sabour et al., 2017). Besides, CNN is not sensitive to the shift or orientation of the object (Sabour et al., 2017). Furthermore, deep CNNs are challenging to train because of the vanishing gradient problem, and the model performance becomes saturated or begins to degrade rapidly with the increase in network depth (He et al., 2016). To tackle the problem of the pooling layer in CNN, the capsule neural network (Sabour et al., 2017) was introduced by Sabour et al. (2017), using the dynamic routing algorithm. Dynamic routing enables a capsule to attend to some active capsules at a low level while ignoring others. Therefore, in this manner, the coupling coefficients among capsules could be iteratively updated. However, Peer et al. (2021) found that capsule networks cannot be viewed as universal approximators due to the limited representation capabilities of routing-by-agreement and EM-routing to symmetric functions (Peer et al., 2021). Furthermore, Capsule networks also suffer from performance issues when dealing with larger datasets including complex data (Xi et al., 2017). The residual neural networks (ResNet) were developed by He et al. (2016) to address the degradation problem and enhance the model performance in the deeper network using residual blocks (see Section 2.1 for more details). In this study, 1D-ResNet adapted from the original ResNet architecture is implemented for earthquake signal detection and seismic phase identification given short-window seismic waveforms similar with Ross et al. (2018), Saad and Chen (2020, 2021), while continuous seismic waveforms are adopted in Zhu and Beroza (2019), Mousavi et al. (2020), Li et al. (2021) and Chakraborty et al. (2021, 2022). Within this work, the seismic data fed into the introduced model is not labeled with specific phase arrival time, however, the proposed model could further be used for the arrival-time estimation of continuous seismic waveforms. Therefore, it is capable of reducing the efforts and time of arrival-time labeling.

In this work, both earthquake detection and seismic phase identification are formulated as a supervised classification problem. Considering that Residual Neural Network (ResNet) (He et al., 2016) achieves superiority in image classification by adopting skip connection and 1D ResNet works well for time-series data (e.g., valve acoustic signals) (Sha et al., 2022), we develop 1D ResNet34 based on the residual module (He et al., 2016) for the defined tasks, here the Reset34 denotes a 34 layers convolution neural network combined with multiple residual blocks. The seismic data of the Southern California Seismic Network (Center, 2013) labeled as ‘P-phase’, ‘S-phase’, and ‘Noise’ is employed to train the model and test its performance. In addition to this, the model’s generalization ability is also verified on the data selected from the STanford EArthquake Dataset (STEAD) (Mousavi et al., 2019). The results well indicate that the proposed method is capable of reliably identifying small magnitude earthquakes and respective P-phase and S-phase.

This work is organized as follows: Section 2 comprehensively delineates the proposed methodology. Section 3 briefly describes the used dataset, and shows the experimental setting and metrics for the performance evaluation. Section 4 details the results of the performed experiments and finally, Section 5 describes the conclusions of this work.

2. Methodology

The proposed method described in this section takes a window of three-channel seismogram data, i.e., three-channel normalized waveform within the duration of 4 s, as input, and outputs the probabilities over three classes: P-wave window, S-wave window, and noise window.

2.1. Residual neural network

In principle, the neural network could be considered as the function approximator $\mathcal{H}(x)$ which maps the input x with the output. In He et al. (2016), the authors showed that when adding more layers to the neural network to enrich the features of the model, training the neural network becomes more challenging due to difficulty in optimizing the model parameters caused by vanishing/exploding gradient. Furthermore, the accuracy of the model either gets saturated at a particular value or slowly degrades. Consequently, the model performance deteriorates both in the training phase and testing phase. To tackle this, the popular neural network architecture known as ResNet was proposed in He et al. (2016), where the layers are explicitly reformulated as learning residual functions with respect to the layer input. The idea is that it is easier to optimize the residual function $\mathcal{F}(x)$ than the original function $\mathcal{H}(x) = \mathcal{F}(x) + x$. Thus, the stacked nonlinear layers aim at fitting the mapping $\mathcal{F}(x) := \mathcal{H}(x) - x$, not the original mapping $\mathcal{H}(x)$. They also provided extensive empirical evidence to indicate that it is not only easier to optimize these residual networks, but also the accuracy could be enhanced from the considerably increased depth. The residual block adopted from He et al. (2016) is displayed at the bottom of Fig. 1. On one hand, the skip connections in ResNet succeeded in dealing with the issue of vanishing gradient in deep neural networks by allowing the gradient to flow directly through the alternative shortcut path backward from the latter layers to the former layers. On the other hand, these connections allow the model to learn the identity functions, which guarantees that the higher layer could perform at least as well as the lower layer.

2.2. Model architecture

In this study, the existent ResNet component, as shown in Fig. 1, is utilized as the fundamental block and adapted for earthquake detection and seismic phase classification using 1D seismic data. The only difference between these two tasks is the output size. In our work, the seismic records with three components are treated as images by the devised 1D ResNet to identify seismic phases to P-wave, S-wave, or noise window, correspondingly.

It should be noted that in this work the two tasks (earthquake detection and phase identification) are separately implemented when using 1D ResNet. Therefore, in the task of earthquake detection, the output is labeled as an earthquake (including but not distinguishing P-wave and S-wave) and noise signals (Saad and Chen, 2020), while in the case of seismic phase identification, the model is trained to classify noise window, P-wave window, and S-wave window, respectively, similar to Ross et al. (2018) and Saad and Chen (2021). For those two tasks, in the model training process, the labels are defined as follows: (i) for earthquake detection — ‘zero’ for earthquake signal and ‘one’ for noise. (ii) for phase classification — ‘zero’ for P-wave windows, ‘one’ for S-wave windows, and ‘two’ for noise windows.

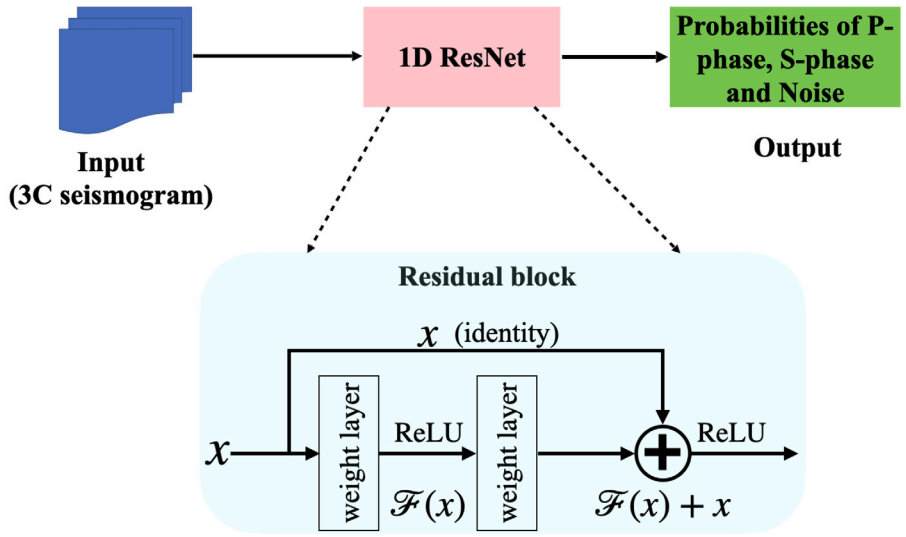

Fig. 1. Model framework. The residual block is reproduced from the work (He et al., 2016), where x is the input vector of the layer, and $\mathcal{F}(x)$ denotes the residual function to be learned.

3. Experiments

3.1. Seismic dataset

In this study, the open-source dataset (SCEDC Center, 2013) recorded by the Southern California Seismic Network is used to train and test the proposed model. This dataset was first used in Ross et al. (2018) and the following preprocessing steps were used: the data was first detrended and high-pass filtered with a corner frequency of 2 Hz to remove microseismic noise, and then resampled at 100 Hz; strong-motion records were integrated to velocity. The 3-component records included in the dataset consist of 1.5 million P-wave seismograms, 1.5 million S-wave seismograms, and 1.5 million noise windows, with each window being exactly 4 s (400 samples) in duration. The magnitude range of the data is $-0.81 < M < 5.7$, while only records with epicentral distances less than 100 km are used. Both P-wave and S-wave windows are centered on the arrival pick, while each noise window starts 5 s before each P-wave arrival pick. Fig. 2 visualizes the waveforms including one P-phase window, one S-phase window, and one noise window.

The STEAD dataset (Mousavi et al., 2019) is also used for model evaluation. In order to help better understand the model output, Fig. 3 visualizes the pipeline of the testing process and the testing result (the predicted probability) using the pie chart of three different input data extracted from one raw waveform of STEAD dataset when fed to the trained model. Similar to the SCEDC dataset (Center, 2013), the P-phase and S-phase windows are centered on the respective arrival time, and the noise windows are extracted starting 5 s before the P arrival time. Each pie chart shows the predicted probability for each class (P-phase, S-phase, Noise), in which the sum over all the probabilities is one.

3.2. Parameters setting

In this study, the ADAM algorithm is used for optimization with a learning rate of 0.001, and the model is trained for 50 epochs when the training loss is stable. The proposed models are implemented in Pytorch (Paszke et al., 2019) and trained on an NVIDIA A100 Graphics Processing Unit. The loss function used is cross-entropy and training is performed in the mini-batches of 480 records. A dropout rate of 0.2 is adopted for all dropout layers. Note that here data augmentation is not used on the training data and at the same time, we do not use any ensemble methods in the model training phase.

Aiming to make the comparison with CapsNet (Saad and Chen, 2020) for the earthquake detection task, 50% of the whole dataset is used for model training and 25% of the data is utilized for testing. However, in CapsNet (Saad and Chen, 2020), the training dataset is balanced, i.e., the number of the data labeled by zero ('earthquake') is same as the number of the data with labels of one ('noise'). Please note that, from the beginning, the whole dataset is labeled with one of the three classes: P-phase, S-phase, or noise. On the other hand, in the case of earthquake and non-earthquake detection, the wave windows including P-phase and S-phase are re-labeled as earthquake signals. The above-mentioned processes make the training dataset biased, where the number of the data labeled by zero (earthquake) is not equal to that labeled by one (noise). For seismic phase identification, 90% of the seismograms are used as a training set, and the remaining 5% of the seismograms are used to build the testing set same as CapsPhase (Saad and Chen, 2021).

3.3. Evaluation metrics

The following metrics are used to evaluate the model performance. First, the accuracy is defined as the ratio of correctly identified instances over all testing samples, which is usually regarded as the basic measurement of a classifier's performance.

$$Accuracy = \frac{N_C}{N_T} \tag{1}$$

where N_C denotes the number of correctly labeled samples and N_T represents the total number of testing samples.

Then, in order to further estimate the model's effectiveness, the confusion matrix (Stehman, 1997) is employed to reflect the classification result. Furthermore, given a confusion matrix, the precision, recall, and F1-score can be defined as follows:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1_{score} = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

where TN, FN, FP, and TP are the true negative, false negative, false positive, and true positive, respectively.

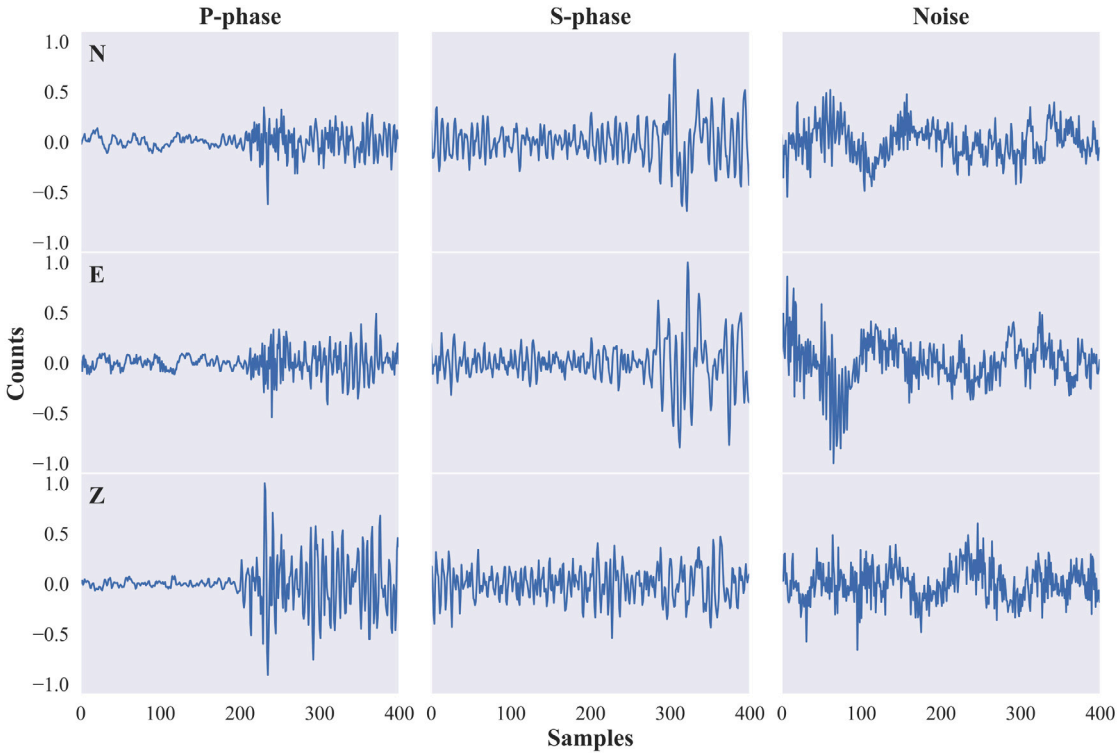

Fig. 2. SCEDC dataset (Center, 2013) waveforms. These waveforms are resampled at 100 Hz and normalized using the absolute maximum amplitude over three components.

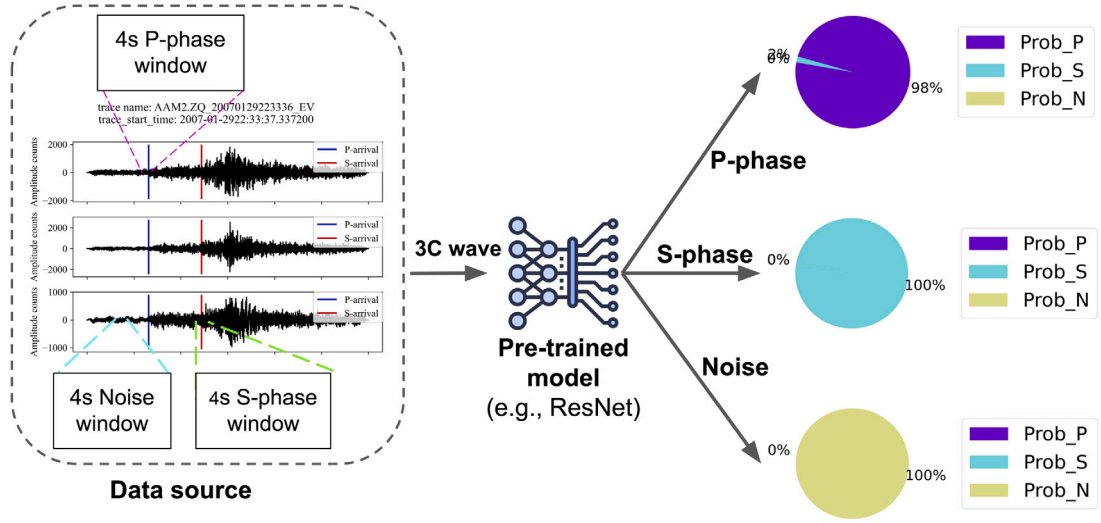

Fig. 3. Output visualization of one testing example from STEAD dataset (Mousavi et al., 2019), where the pie chart displays the predicted probability for the classes corresponding to different input data.

4. Results and discussions

4.1. Phase identification performance

In order to find an optimum architecture for seismic phase identification, we first tested several ResNet blocks such as 1D-ResNet18, 1D-ResNet34, and 1D-ResNet50. The results show that 1D-ResNet34 (98.70%) performs better than both 1D-ResNet18 (98.63%), and 1D-ResNet50 (98.62%) even though 1D-ResNet34 has fewer parameters compared with 1D-ResNet50.

The overall testing accuracy for earthquake detection and phase classification of different methods is compared and summarized in Table 1. We can find that the testing accuracy of 1D ResNet is 98.83%. The result demonstrates that our proposed model achieves better performance for earthquake detection than CapsNet (Saad and Chen, 2020) (here, the result of CapsNet is stated from Saad and Chen (2020)). For seismic phase classification, 1D ResNet demonstrates its superiority compared with CapsPhase (Saad and Chen, 2021) (here, the result of CapsPhase is adopted from Saad and Chen (2021)). The potential reason behind achieving higher performance could be that the residual blocks in ResNet (He et al., 2016) contribute to improving the classification accuracy since it is capable to learn some meaningful features from the input.

The results of different metrics including precision, recall and F1-score for earthquake identification are shown in Table 2. It is found that 1D ResNet can achieve compatible results with CapsNet.

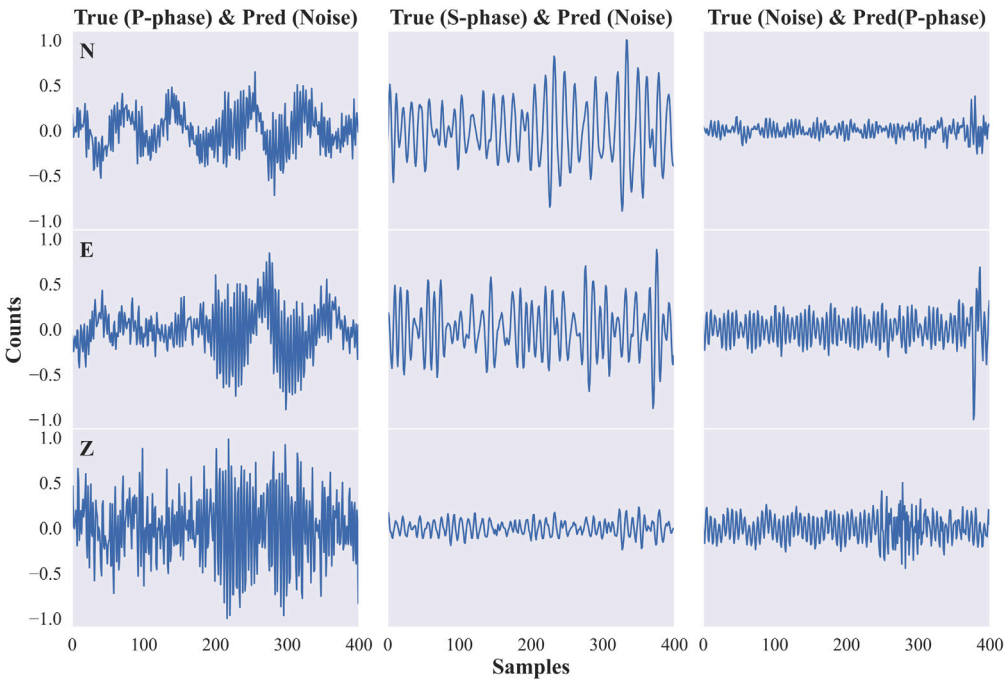

Fig. 4. Visualization of misclassified SCEDC data. These waveforms are resampled at 100 Hz and normalized using the absolute maximum amplitude over three components.

Table 1
Testing accuracy for two tasks on the SCEDC dataset.

Model	Earthquake detection	Phase classification
CapsNet	98.40%	–
CapsPhase	–	98.67%
1D ResNet	**98.83%**	**98.70%**

Table 2
Testing results for earthquake detection on the SCEDC dataset.

Category	Model	Precision	Recall	F1-score
Earthquake	CapsNet	98.64%	98.98%	98.80%
	1D-ResNet	**99.18%**	98.06%	**99.12%**
Noise	CapsNet	97.96%	97.30%	98.70%
	1D-ResNet	**98.13%**	**98.37%**	98.25%

Table 3
Testing results for phase classification on the SCEDC dataset.

Category	Model	Precision	Recall	F1-score
P-phase	CapsPhase	98.68%	98.99%	98.76%
	1D-ResNet	**98.88%**	98.64%	**98.76%**
S-phase	CapsPhase	98.40%	98.88%	98.70%
	1D-ResNet	**98.72%**	**98.94%**	**98.83%**
Noise	CapsPhase	98.90%	98.17%	98.53%
	1D-ResNet	98.52%	**98.54%**	**98.53%**

Finally, Table 3 summarizes the classification result of different metrics for seismic phase identification. It can be observed that compared with CapsPhase (Saad and Chen, 2021), the proposed model achieves better performance, especially for S wave identification.

As reported in Ross et al. (2018), P- and S-waves windows are obtained from the earthquake events, whereas, the noise windows are selected starting 5 s before each P-wave arrival. Hence, there are similarities between these windows. Fig. 4 visualizes an example for each class including the P-phase window, S-phase window, and noise window from the SCEDC dataset that is misclassified by the proposed 1D-ResNet in this work. As one can see, the resemblance between data belonging to different classes could lead to misclassification. In the case where the noise is either misclassified as P- or S-phase, we notice an increase in either amplitude or sudden change in frequency or both which might have contributed to the wrong detection. Whereas, most of the P- or S-phase which are misclassified as noise were noisy signals.

4.2. Model generalizability

To further investigate the model generalization ability, the proposed model is evaluated on the STEAD dataset (Mousavi et al., 2019) and compared with previous deep learning-based phase identification methods.

For this purpose, local earthquake signals of the STEAD dataset after normalization are used. As in Ross et al. (2018), only seismic data satisfying the following requirements are used. First, P- and S-wave windows are centered on the ground truth phase pick corresponding to the dataset metadata, and each noise window is clipped starting 5 s before each P-phase arrival time. Second, the arrival-time difference between P-wave and S-wave picks of each signal is larger than 6 s with the goal to reduce the effect of the waveform overlapping between P-phase and S-phase arrivals. Third, only the seismograms with the 'source_distance' labeled in the metadata of the STEAD dataset less than 100 km are chosen same as (Ross et al., 2018). Finally, the created dataset consists of 8111 P-wave seismograms, 8111 S-wave seismograms, and 8111 noise windows. Here the duration of each window is 400 samples with a sampling rate of 100 Hz. The earthquake magnitude of the used data ranges from 1 to 3. The distributions of the earthquake magnitude, source distance, and the time difference between P-phase and S-phase arrival time are displayed in Fig. 5.

The proposed method is first compared with CapsPhase (Saad and Chen, 2021). Here, these two models are pre-trained on the SCEDC dataset, and the best-saved model of CapsPhase published by the authors is directly used without model retraining. In contrast to the original work, Saad and Chen (2021), in this comparison, no threshold is used for the model output. The overall testing accuracies of the used two methods are 94.39% (CapsPhase), and 95.68% (1D-ResNet), respectively. Thus, we see find that 1D-ResNet used in this study achieves better performance, especially in the S-phase identification.

The confusion matrices on the selected data for the three methods are depicted in Fig. 6 and the precision, recall, and F1-score are

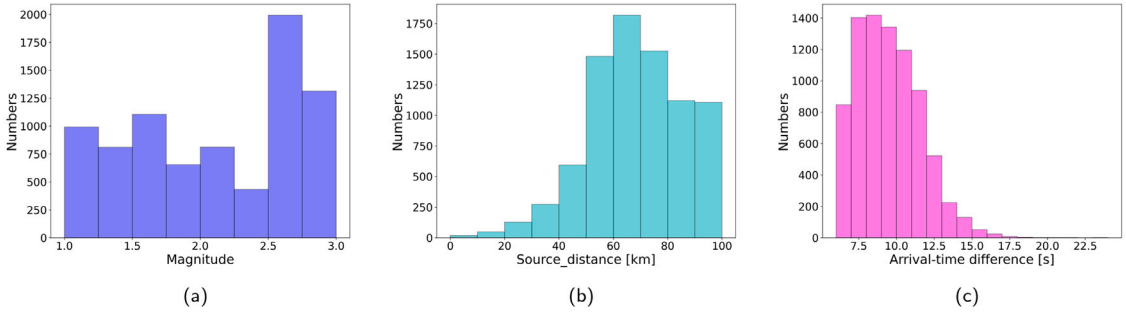

Fig. 5. Distributions of the (a) magnitude, (b) source distance, and (c) arrival-time difference between P- and S-phase arrival time for the used STEAD data.

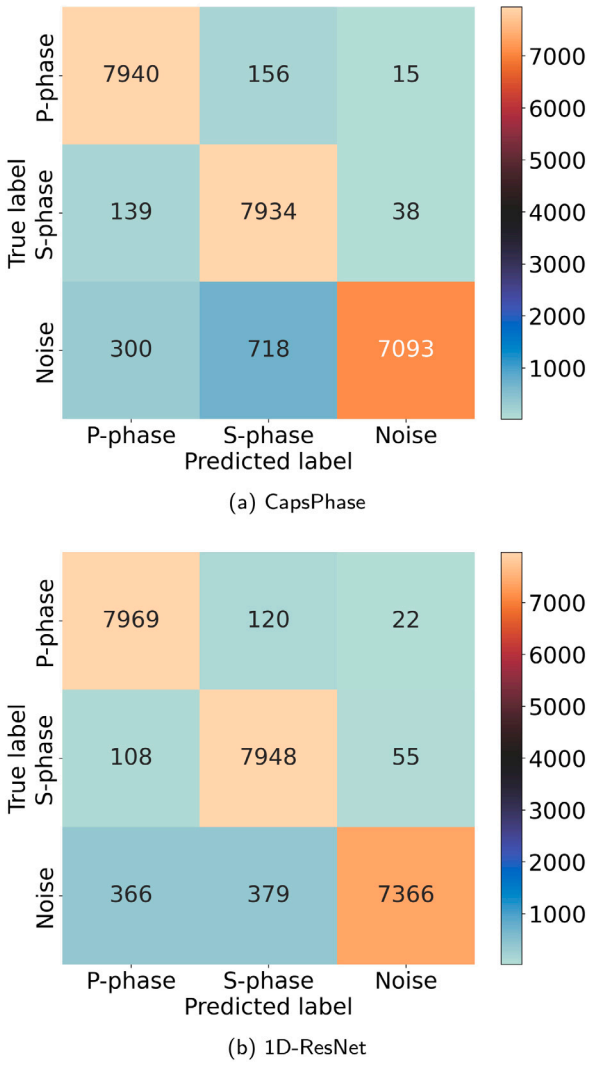

Fig. 6. Confusion matrices for (a) CapsPhase and (b) 1D-ResNet on the STEAD dataset.

summarized in Table 4. Table 4 shows that the proposed 1D-ResNet performs better in S-phase identification, and achieves a comparable performance in P-phase and noise window identification compared with CapsPhase.

Finally, the proposed method is also compared to the generalized phase detection method (GPD) (Ross et al., 2018). Hence, 75% of the SCEDC dataset is used to retrain the proposed 1D-ResNet, which is the same as GPD, and the built STEAD data is further utilized to test the model performance. Here, the best-saved model of the GPD method is

Table 4
Testing results of CapsPhase and 1D-ResNet for phase identification on the STEAD dataset.

Category	Method	Precision	Recall	F1-score
P-phase	CapsPhase	94.76%	97.89%	96.30%
	1D-ResNet	94.39%	98.25%	96.28%
S-phase	CapsPhase	90.08%	97.82%	93.79%
	1D-ResNet	94.09%	97.99%	96.00%
Noise	CapsPhase	99.26%	87.45%	92.98%
	1D-ResNet	98.96%	90.81%	94.72%

Table 5
Testing results of GPD and 1D-ResNet for phase identification on the STEAD dataset.

Category	Method	Precision	Recall	F1-score
P-phase	GPD	93.55%	97.67%	95.57%
	1D-ResNet	93.15%	97.96%	95.50%
S-phase	GPD	94.13%	98.26%	96.15%
	1D-ResNet	95.12%	97.58%.	96.34%
Noise	GPD	99.24%	90.52%	94.69%
	1D-ResNet	98.72%	91.06%	94.73%

used without retraining. The testing accuracy of GPD reaches 95.48%, and the accuracy of 1D-ResNet reaches 95.54%. Compared with GPD, 1D-ResNet achieves a slight improvement. The confusion matrices and the results of different metrics are shown in Fig. 7 and Table 5, where we can find that 1D-ResNet achieves competitive performance with the GPD method (Ross et al., 2018).

4.3. Discussions

To investigate the model performance, when facing more noisy data, the same subset of the STEAD dataset used for model generalization evaluation in the study is masked by Gaussian noise (similar to the approach used in Mousavi et al. (2020)). Then, the noisy data is fed to the pre-trained model to test the performance. The signal-to-noise ratio (SNR) of the data before adding noise ranges from 0 to 70 dB. Here, the SNR is the mean value of SNR over three-components for each signal. To study the impact of different noise levels on model performance, the subset is masked by the Gaussian noise (similar to the method used in Mousavi et al. (2020)) with mean $\mu = 0$ and standard deviation $\delta = 0.01$, 0.05, 0.1, and 0.5, respectively. The testing accuracies are summarized in Table 6 below.

The results in Table 6 show that large noise reduces the model performance, and compared with CapsPhase and GPD, 1D-ResNet34 achieves a better performance.

The SCEDC dataset used in this study consists only of windows centered around the first arrivals of P- and S-waves. Therefore, during the training phase only windows with P- and S-arrivals in the center are encountered by the model; this could affect the model performance when tested on windows not centered around P- and S-arrivals. To study the case where the P- and S-waves are not at the center of the

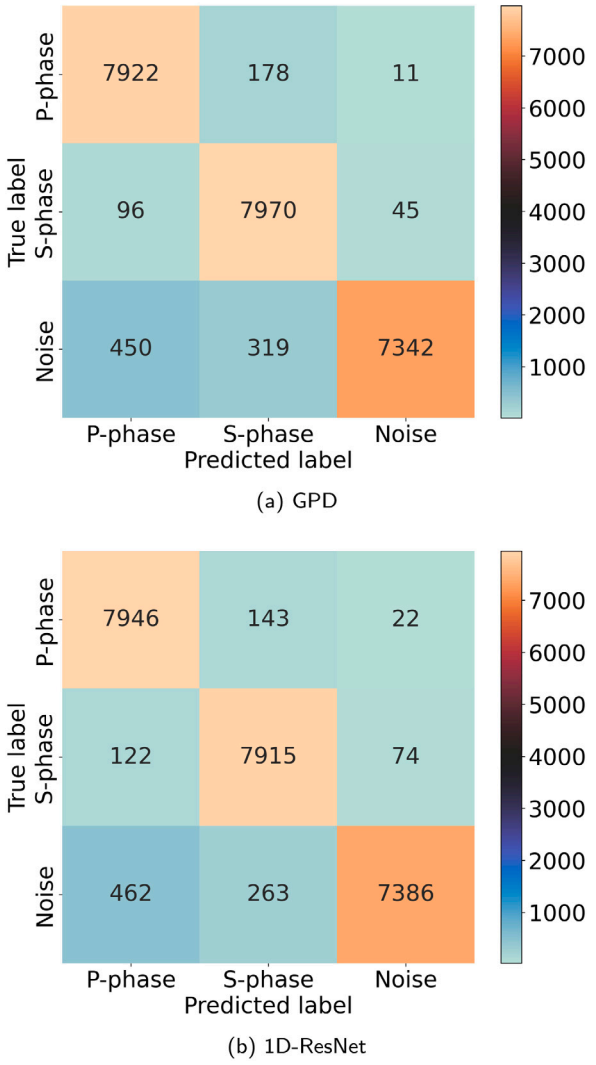

Fig. 7. Confusion matrices for (a) GPD and (b) 1D-ResNet on the STEAD dataset.

Table 6
Testing results of different noise levels for phase identification on the STEAD dataset.

Noise level	0.01	0.05	0.1	0.15
1D-ResNet34	96.30%	96.46%	93.22%	89.16%
Capsphase	95.28%	95.43%	92.80%	88.90%
GPD	95.85%	95.89%	92.83%	88.82%

windows, we created a subset of the STEAD where both P- and S-waves are randomly located from 1 s to 3 s after the window start time. The testing accuracy, as expected in this case, was reduced to 80.66%, and the confusion matrix is shown in Fig. 8. The reason behind the decrease in accuracy is most likely the difference between the position of P- and S-arrivals in the training data and the testing data which could have resulted in incorrect decisions.

5. Conclusions

In this study, we investigate the 1D residual neural network (ResNet) for earthquake detection and seismic phase classification. This model is trained and tested on the Southern California Seismic Dataset and further evaluated on the STEAD dataset. Extensive experimental results verify that the 1D-ResNet achieves a better performance than previous deep learning-based approaches, and performs better over the baseline

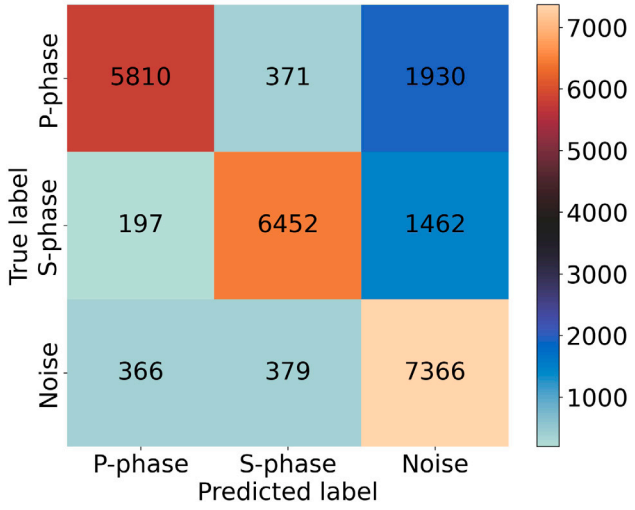

Fig. 8. Confusion matrix for 1D-ResNet on the phase-shifted data.

methods when facing more noisy data. The proposed model can be utilized by seismologists to identify earthquake signals and phases, especially in the case of noisy low-magnitude earthquake waveforms. Future work will focus on a hierarchical classifier to hierarchically complete the tasks of earthquake detection and seismic phase identification. Note that this study only focuses on seismic phase identification, in follow-up work, we will investigate how to use the pre-trained phase classifier to improve the phase arrival time picking on continuous seismic waveforms.

CRediT authorship contribution statement

Wei Li: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Megha Chakraborty:** Writing – review & editing. **Yu Sha:** Methodology. **Kai Zhou:** Methodology, Writing – review & editing. **Johannes Faber:** Methodology, Writing – review & editing. **Georg Rümpker:** Conceptualization, Writing – review & editing. **Horst Stöcker:** Writing – review & editing. **Nishtha Srivastava:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data and resources

The Southern California seismic data that support this study can be accessed in Center (2013) and the STEAD dataset is available at Mousavi et al. (2019). The code is available on GitHub at https://github.com/srivastavaresearchgroup/Seismic-phase-Classification.

Acknowledgments

This work is supported by the "KI-Nachwu chswissenschaftlerinnen" - grant SAI 01IS20059 by the Bundesministerium für Bildung und Forschung - BMBF. Calculations were performed at the Frankfurt Institute for Advanced Studies' GPU cluster, funded by BMBF for the project Seismologie und Artifizielle Intelligenz (SAI). Horst Stöcker gratefully acknowledges the Judah M. Eisenberg Laureatus - Professur at Fachbereich Physik, Goethe Universität Frankfurt, funded by the

Walter Greiner Gesellschaft zur Förderung der physikalischen Grundlagenforschung e.V. Yu Sha would like to thank the support from Xidian-FIAS International Joint Research Center. Dr. Kai Zhou would like to thank the AI grant at FIAS through SAMSON AG, and the BMBF funding through the ErUM-Data project. We would like to take this opportunity to acknowledge the time and effort devoted by reviewers to improving the quality of the manuscript.

References

Center, S.C.E.D., 2013. Southern California Earthquake Data Center. Dataset, California Institute of Technology, http://dx.doi.org/10.7909/C3WD3xH1.

Chakraborty, M., Fenner, D., Li, W., Faber, J., Zhou, K., Rüempker, G., Stoecker, H., Srivastava, N., 2022. Creime: A convolutional recurrent model for earthquake identification and magnitude estimation. J. Geophys. Res: Solid Earth 127 (7), e2022JB024595.

Chakraborty, M., Li, W., Faber, J., Rüempker, G., Stoecker, H., Srivastava, N., 2021. A study on the effect of input data length on deep learning based magnitude classifier. arXiv preprint arXiv:2112.07551.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.

Li, W., Chakraborty, M., Fenner, D., Faber, J., Zhou, K., Rüempker, G., Stoecker, H., Srivastava, N., 2021. Epick: Multi-class attention-based u-shaped neural network for earthquake detection and seismic phase picking. arXiv preprint arXiv:2109.02567.

Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y., Beroza, G.C., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature Commun. 11 (1), 1–12.

Mousavi, S.M., Sheng, Y., Zhu, W., Beroza, G.C., 2019. Stanford earthquake dataset (stead): A global data set of seismic signals for ai. IEEE Access 7, 179464–179476.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al., 2019. Pytorch: An imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. 32.

Peer, D., Stabinger, S., Rodriguez-Sanchez, A., 2021. Limitation of capsule networks. Pattern Recognit. Lett. 144, 68–74.

Peng, Z., Zhao, P., 2009. Migration of early aftershocks following the 2004 parkfield earthquake. Nat. Geosci. 2 (12), 877–881.

Ross, Z.E., Meier, M.-A., Hauksson, E., Heaton, T.H., 2018. Generalized seismic phase detection with deep learning. Bull. Seismol. Soc. Am. 108 (5A), 2894–2901.

Ross, Z.E., Rollins, C., Cochran, E.S., Hauksson, E., Avouac, J.-P., Ben-Zion, Y., 2017. Aftershocks driven by afterslip and fluid pressure sweeping through a fault-fracture mesh. Geophys. Res. Lett. 44 (16), 8260–8267.

Saad, O.M., Chen, Y., 2020. Earthquake detection and p-wave arrival time picking using capsule neural network. IEEE Trans. Geosci. Remote Sens..

Saad, O.M., Chen, Y., 2021. Capsphase: Capsule neural network for seismic phase classification and picking. IEEE Trans. Geosci. Remote Sens..

Sabour, S., Frosst, N., Hinton, G.E., 2017. Dynamic routing between capsules. Adv. Neural Inf. Process. Syst. 30.

Sha, Y., Faber, J., Gou, S., Liu, B., Li, W., Schramm, S., Stoecker, H., Steckenreiter, T., Vnucec, D., Wetzstein, N., et al., 2022. A multi-task learning for cavitation detection and cavitation intensity recognition of valve acoustic signals. Eng. Appl. Artif. Intell. 113, 104904.

Stehman, S.V., 1997. Selecting and interpreting measures of thematic classification accuracy. Remote Sens. Environ. 62 (1), 77–89.

Xi, E., Bing, S., Jin, Y., 2017. Capsule network performance on complex data. arXiv preprint arXiv:1712.03480.

Zeiler, M.D., Fergus, R., 2014. Visualizing and understanding convolutional networks. In: European Conference on Computer Vision. Springer, pp. 818–833.

Zhou, Y., Yue, H., Fang, L., Zhou, S., Zhao, L., Ghosh, A., 2022. An earthquake detection and location architecture for continuous seismograms: Phase picking, association, location, and matched filter (palm). Seismological Society of America 93 (1), 413–425.

Zhu, W., Beroza, G.C., 2019. Phasenet: a deep-neural-network-based seismic arrival-time picking method. Geophys. J. Int. 216 (1), 261–273.