OBSU: a deep-learning seismic phase picker for OBS data using transfer learning and Unet

Yi Xie [1, a], Yongming Huang [1, b]

[1] School of Automation, Southeast University Nanjing 210096, P. R. China

[a] xiey@seu.edu.cn, [b] huang_ym@seu.edu.cn

Abstract. Seismic phase identification and ariivals picking are two essential steps in the processing of seismic monitoring data. We propose OBSU, a transfer learning and Unet-based seismic phase picker for ocean earthquakes with multimodal inputs, using the land seismic dataset INSTANCE for pre-training to enhance performance. We test it using the ocean bottom seismometer (OBS) dataset and achieve mean absolute deviations of 0.17 s and 0.23 s for P-wave and S-wave, respectively. Meanwhile, the model results after using transfer learning are significantly better than before using it.

Keywords: Deep learning, Transfer learning, Ocean bottom seismometer, Seismic phase picking

1. Introduction

The ocean encompasses 70.8% of the Earth's surface, and approximately 60% of earthquakes with a magnitude of 5 or higher occur in oceanic and nearshore regions worldwide. Earthquakes in marine and nearshore areas, along with the secondary hazards they trigger (such as tsunamis), pose substantial risks to human existence and progress that warrant serious consideration. Moreover, earthquakes serve as crucial tools and windows for studying the internal structure, evolution, and geophysics of the Earth[1].

Recent advancements in computing power, coupled with the emergence of large seismic datasets, have paved the way for the development of several deep learning (DL) models for earthquake detection and phase picking, which exhibit outstanding accuracy and efficiency in land data[2-4].

In contrast to land seismic data, the utilization of ocean bottom seismometer (OBS) data in the advancement of DL phase pickers has been relatively overlooked, mainly owing to the scarcity of well-annotated extensive datasets. Additionally, OBS data commonly presents inferior signal-to-noise ratios and lower-quality phase arrivals compared to data recorded at land stations, attributed to various factors such as underwater currents, instrument tilt, and instrument coupling[5].

Some researchers have explored the effectiveness of directly applying land-based earthquake detection algorithms to OBS data. With the rise of deep learning, some seismologists have begun to directly apply land DL phase pickers to OBS data for phase picking[6-9]. However, experimental results indicated that onshore seismic picking models may not necessarily generalize well to OBS data. Transfer learning is a machine learning technique that aims to enhance the learning efficiency and performance of new tasks by leveraging existing knowledge and experience[10]. Benefiting from abundant land seismic data and leveraging the similarity between ocean bottom seismic signals and land seismic signals, transfer learning is utilized to enhance the performance of phase picking in OBS data. Niksejel et al.[5] employed the widely trained land phase picker EqT as the base model and developed an efficient OBS phase picker called obstrtransformer (OBST) through transfer learning. PickBlue[11] trained on three-component seismic records and hydrophone channel data, using pre-trained PhaseNet and EqT to process OBS data. Meanwhile, Bornstein et al. [11] pointed out that OBS data often include an additional hydrophone channel not available at onshore seismometers. Cheng et al. [12] utilized a transfer learning approach based on the existing U-GPD model, to develop an automatic phase detection model (OBS phase detection, OBSPD) for OBS data, showing that transfer learning can achieve lower model loss and less overfitting compared to training the model from scratch.

In this paper, we have developed a high-performance marine seismic phase picker named OBSU using transfer learning, which is applied to OBS data. The results indicate that our model exhibits outstanding performance of seismic phases picking for OBS data.

2. Data

In this paper, we analyzed an extensive database of waveforms from local earthquakes in various submarine tectonic environments, with the majority being four-component waveforms, complied by Bornstein et al. [11]. The comprehensive OBS dataset consists of manually selected phases from 15 deployments and a total of 355 stations. The dataset comprises 13190 events, 109210 traces and 153338 picks (about 90000 P and 63000 S picks). The OBS data was divided into training, development and holdout sets in a 66.8%-12.8%-20.4% ratio, consistent with the partitioning in pickblue.

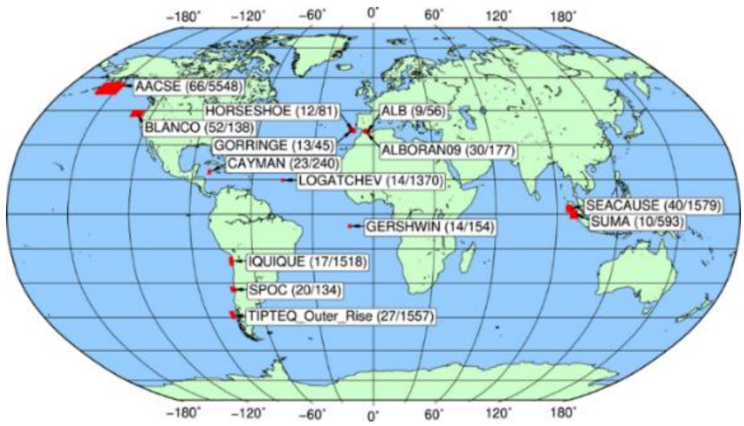

Fig. 1 Global map showing the distribution of the OBS networks [11]

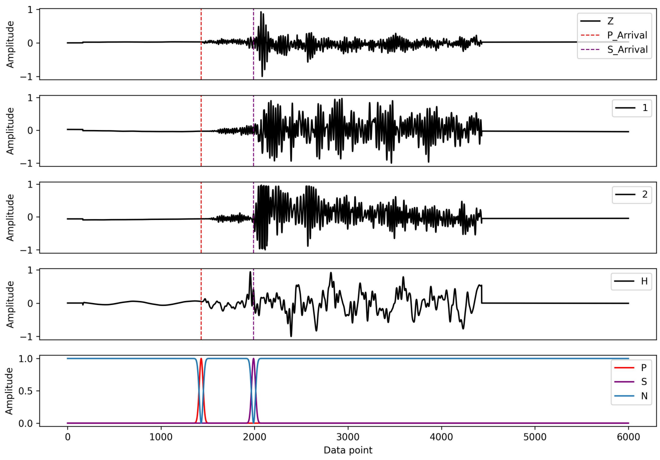

Fig. 2 A sample from the data set. From top to bottom: trace with four components after preprocessing (resampling, cutting of waveform, and normalization); the converted probability masks for P and S picks. The red and purple vertical dashed lines are manually selected P and S arrival times. Z: vertical component; 1 and 2: two horizontal components; and H: Hydrophone channel.

We performed the following steps for OBS waveform pre-processing:1) resampling data at 100 Hz; 2) cutting 60 s earthquake waveform randomly; 3) normalizing each channel of the data; 4) performing demean and detrending on the data. Fig. 2 shows the waveform input after the preprocessing step used for training.We trained the network to replicate a feature function where Gaussian peak values with amplitudes 1 and a half-width 0.1s are centered around manually picked arrivals of P and S phases. The characteristic function of noise includes all data points that are not the first arrival of P or S waves. By applying a Gaussian distribution mask transformation, we can extract the accurate arrival time from the peak of the probability distribution predicted by the model in this paper.

In addition to waveform input, we also utilized the Spectrogram Transformation technique based on Short-Time Fourier Transform (STFT) for the vertical component, reflecting the signal's characteristics in the frequency domain[13]. The spectrograms are represented along the time and frequency axes to analyze the seismic signal features as well as P- and S-waves more clearly . They are fed into the CNN model to represent the image features more clearly. The STFT of signal $\mathrm{x(t)}$ is defined as:

$$STFT(t,f)=\int_{-\infty}^{+\infty}x(\tau)h(\tau-t)e^{-j2\pi f\tau}d\tau$$

where $\mathrm{x(t)}$ represent the input signal, $h(\mathrm{t})$ is the window function. Earthquake signals are discontinuous functions, therefore, the vertical-component spectrogram matrix $\mathrm{S^Z}$ is computed from $\mathrm{x^Z(t)}$ using the squared absolute value of discrete STFT with a Hann window function w in the length of $\mathrm{N}=20$ (i.e., 0.2 s) and a hop size $\mathrm{H}=8$ (i.e., 0.08 s) that determines how many samples to shift across $\mathrm{x^Z(t)}$. After pre-processing, each waveform follows a 60s time window and contains $\mathrm{L}\ =\ 6000$ data points. Before computing the STFT, we pad zeros to both ends of $\mathrm{x^Z(t)}$ to ensure that the maximum time frame is equal to $\mathrm{L/H}\ +\ 1\ =\ 751$. The number of frequency bins is $\mathrm{M/2}\ +\ 1\ =\ 51$. As a result, the shape $\mathrm{S^Z}$ of the frequency domain encoder has a shape of $(2,51,751)$, where the first axis represents the real and imaginary values. The real and imaginary parts are fed to the neural network as two separate channels so that the network is able to learn from both the time and phase information.

3. Methodology

In this paper, we proposed a OBSU that incorporates both multimodal fusion and transfer learning techniques. When data is scarce, transfer learning is an effective strategy that can reduce overfitting and improve model training performance.

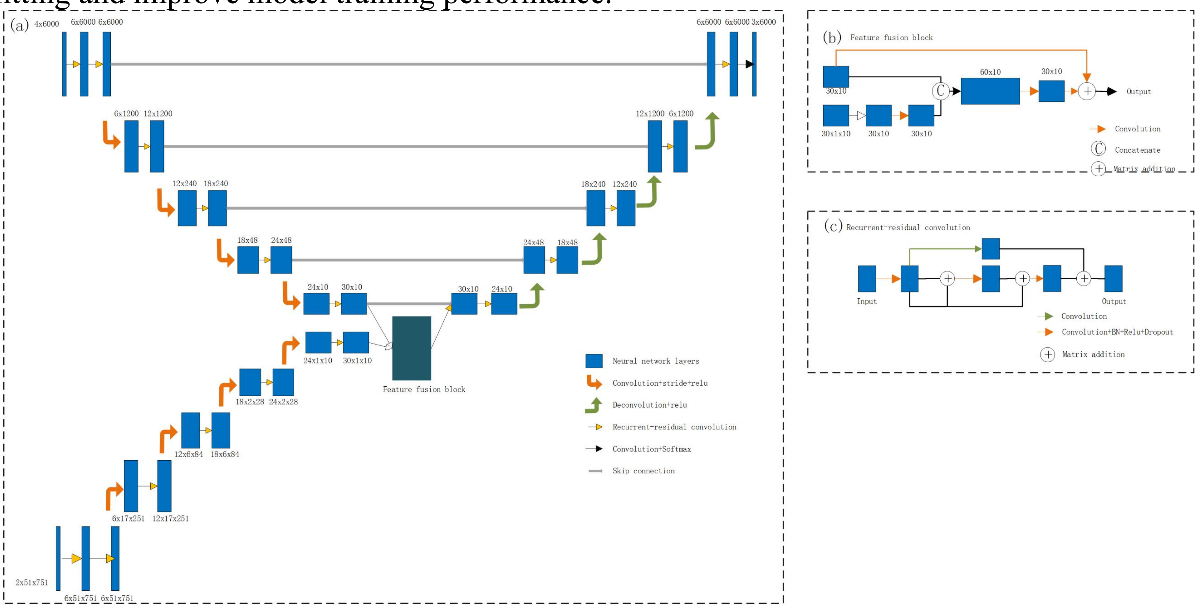

Fig. 3 The network architecture.

Our model uses two encoders in order to realize multimodal inputs (Fig. 3), then performs information fusion at the feature layer, and finally outputs the P-phase and S-phase probabilities with a decoder. The mapping to our problem is to localize the properties of our time-series into three classes: P pick, S pick and noise.

The seismic data undergoes four downsampling stages and four upsampling stages. In each waveform data feature extraction stage, we applied one-dimensional convolution and Rectified Linear Unit (ReLU) activation, with the one-dimensional convolution size set to seven data points and the downsampling stride set to 5 data points, resulting in channel lengths being compressed to one-fifth of their original size after each stride. In each frequency domain information feature extraction stage, we apply two-dimensional convolution with a convolution size of (3, 3), the stride step for down-sampling is set to (3, 3). The downsampling process aims to extract useful information from the seismic data and shrink it to a few neurons, so that each neuron in the final layer constitutes a wide receptive window. The upsampling process expands and converts this information into probability distributions of P wave, S wave, and noise for each time point. Skip connections for each depth directly link the output from the left to the right layers without passing through deeper layers, which helps improve convergence during training[14]. Deconvolution operations[15] are used for upsampling, expanding the compressed layers five times to restore their previous lengths. Padding is added before and after each layer in the convolution process to ensure that the input and output sequences have the same length.

In the feature fusion block, the information encoded by two encoders is input into the block, and the features are fused as the input for the subsequent decoder through a series of convolution and matrix addition operations. Additionally, we utilized recurrent residual networks[16], employing feature summation at different time steps to obtain more expressive features, which also helps extract lower-level features. This approach, different from traditional Convolution+Relu layers, effectively increases the depth of the network.

4. Experiment

Transfer learning strategies rely on the observation that the different parts of CNN learn different things. The initial convolutional layers extract fundamental components like edges and basic textures, while the deeper layers discern more intricate patterns relevant to the specific task and dataset. Two widely used transfer learning strategies in deep transfer learning frameworks in seismology are full fine-tuning and feature extractor. The 'full fine-tuning' has been shown to be an effective method in several seismological studies, outperforming the model training from scratch[17-19].For cross-domain applications, the training or pre-training datasets must have a distance range similar to the target application domain[20]. Furthermore, source models trained on large datasets typically exhibit optimal performance after fine-tuning. Due to the presence of local earthquakes in the OBS data, we opted to initialize weights using the INSTANCE dataset, then employed full fine-tuning for transfer learning. Based on the designed deep learning network model, the model is trained on the Pytorch framework.The Adam algorithm was adopted in the training process, and the cross-entropy loss function was used to optimize the network parameters. The learning rate was set at 0.001, and the number of training samples in each batch was set at 256.To conduct the performance comparison, we follow the convention of Mousavi et al. [2] and define the picking residuals Δt as the difference between ground truth and model picks. Meanwhile, we have chosen the following evaluation metrics to quantify and compare the pickers' performance: precision, recall, F1 score, the modified mean absolute error (MAE). Peak probabilities above 0.3 were counted as positive picks. For P-waves, arrival time residuals less than 0.5 seconds ($\Delta t < 0.5$ seconds) were counted as true positives. Since the S-wave is more noisy, time-of-arrival residuals less than 1 second ($\Delta t < 1$ second) are counted as true positives.

4.1 Comparison with other method

We compare our results with those obtained by BluePhasenet[11]. Also when calculating the MAE, to prevent the effect of some extreme errors, residuals with an absolute value exceeding 1s were set to 1s.

Table 1 Comparison between the proposed and existing model.

Evaluation indicator	Phase	OBSU	BluePhasenet
Precision	P	0.943	-
	S	0.730	-
Recall	P	0.888	-
	S	0.845	-
F1 score	P	0.915	-
	S	0.767	-
Accuracy	P	0.895	0.875
	S	0.830	0.806
MAEs	P	0.17s	0.23s
	S	0.23s	0.32s

When the trained model is applied to test on the test set, the results are shown in Table 1, which shows that the F1 score exceeds 0.9 for P and 0.8 for S.The MAEs for OBSU were 0.17s and 0.23s, respectively.For P onsets and S onsets, OBSU outperforms BluePhasenet with MAEs of 0.23 versus 0.32 s, whereas both models show comparable results for S onsets. Compared to BluePhasenet, OBSU picks up seismic phase with higher accuracy and lower MAE, demonstrating improved capabilities for detecting P-wave and S-wave.Overall, the picking performance of P is better than that of S. This may be related to the composition of seismic waves, as P are the fastest propagating seismic waves, which arrive relatively earlier and usually have a more distinct initial waveform. S, on the other hand, tend to be submerged between various seismic fluctuations due to their slower propagation speed, while the noise level is relatively high on the horizontal component of the OBS data, thus S arrivals usually have a lower signal-to-noise ratio, leading to difficulties in picking.

4.2 Effectiveness of transfer learning

Our model was trained on INSTANCE then fine-tuned using OBS data, with random weights for the hydrophone channels. Hence, we will evaluate the effect of this transfer learning.

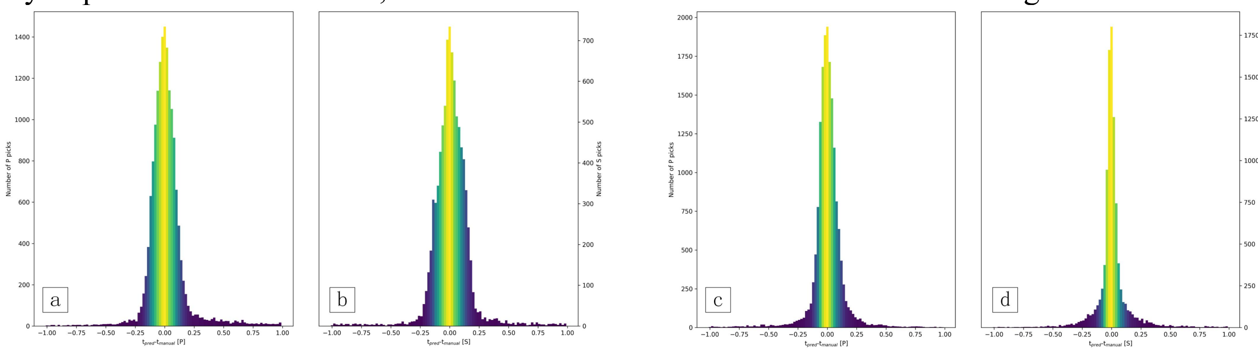

Fig. 4 Histogram illustrating the residuals between manual picks and those made by OBSU for P phases (panels (a) and (c)) as well as S phases (panels (b) and (d)). (a, b) Models pre-trained on INSTANCE, then trained on ocean bottom seismometer (OBS) data set. (c, d) Models trained on OBS data set without pre-training. each bin represents a 0.02-second interval.

Fig. 4 shows the performance of the model without transfer learning and with transfer learning from INSTANCE in terms of P- and S-wave residuals. As can be seen in Fig. 4 , after the transfer learning, the residual distributions of the Ppicks and Spicks are much narrower, which is consistent with the fact that the effect of transfer leanring. Migration learning enables lower model loss and less overfitting, and in the case of seismic event detection, it leads to a significant improvement in seismic phase picking.

For the P arrivals picking, the residual distribution looks more symmetric, while for the residual distribution of picking S arrivals, the central part is very slightly biased towards later arrivals, resulting in too late picks on average. The residuals follow an approximate Laplace distribution with or without the use of transfer learning, which is a side-effect of the fact that our labelling is largely plausible, but due to the quality of the data, our phase picks are not as good as the land seismic as a whole.

5. Summary

In this paper, we developed a multimodal and Unet-based seismic phase picker. The model is pre-trained using massive land seismic data INSTANCE and the OBS dataset is trained by transfer learning, for the land dataset three component waveforms are used as inputs, whereas for the OBS dataset, four component waveforms are used as data inputs, and random weights are used for the additional hydrophone channels. By comparing with BluePhasenet, our OBSU shows some superiority in arrival picking.

Acknowledgments

This paper is supported by Key Researchand Development Program of Jiangsu Province(BE2023022).

References

[1] Wan Y. G., Sheng S. Z., Li X., et al. Stress influence of the 2015 Nepal earthquake sequence on Chinese mainland. Chinese Journal of Geophysics-Chinese Edition, 2015, 58 (11): 4277-4286.

[2] Mousavi S. Mostafa, Ellsworth William L., Zhu Weiqiang, et al. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature Communications, 2020, 11 (1): 3952.

[3] Zhu Weiqiang, Beroza Gregory C. PhaseNet: a deep-neural-network-based seismic arrival-time picking method. Geophysical Journal International, 2018, 216 (1): 261-273.

[4] Ross Z. E., Meier M. A., Hauksson E., et al. Generalized Seismic Phase Detection with Deep Learning. Bulletin of the Seismological Society of America, 2018, 108 (5A): 2894-2901.

[5] Niksejel A., Zhang Miao. OBSTransformer: a deep-learning seismic phase picker for OBS data using automated labelling and transfer learning. Geophysical Journal International, 2024, 237 (1): 485-505.

[6] Ruppert N. A., Barcheck G., Abers G. A. Enhanced Regional Earthquake Catalog with Alaska Amphibious Community Seismic Experiment Data. Seismological Research Letters, 2023, 94 (1): 522-530.

[7] Gong J. H., Fan W. Y., Parnell-Turner R. Microseismicity Indicates Atypical Small-Scale Plate Rotation at the Quebrada Transform Fault System, East Pacific Rise. Geophysical Research Letters, 2022, 49 (3).

[8] Chen H., Yang H. F., Zhu G. H., et al. Deep Outer-Rise Faults in the Southern Mariana Subduction Zone Indicated by a Machine-Learning-Based High-Resolution Earthquake Catalog. Geophysical Research Letters, 2022, 49 (12).

[9] Pilot Matthias, Schlindwein Vera. A Practical Approach to Automatic Earthquake Catalog Compilation in Local OBS Networks Using Deep-Learning and Network-Based Algorithms. Seismological Research Letters, 2024.

[10] Zhuang Fuzhen, Luo Ping, He Qing, et al. Survey on Transfer Learning Research. Journal of Software, 2015, 26 (1): 26-39.

[11] Bornstein Thomas, Lange Dietrich, Münchmeyer Jannes, et al. PickBlue: Seismic phase picking for ocean bottom seismometers with deep learning. Earth and Space Science, 2024, 11: e2023EA003332.

[12] Cheng Zhangbao, Tan Yen Joe, Zhang Fan, et al. An Efficient Transfer Learning - Based OBS Seismic Phase Picker (OBSPD) Trained on Cascadia Subduction Zone Earthquake Dataset. Seismological Research Letters, 2023.

[13] Thomazella R., Lopes W. N., Aguiar P. R., et al. Digital signal processing for self-vibration monitoring in grinding: A new approach based on the time-frequency analysis of vibration signals. Measurement, 2019, 145: 71-83.

[14] Li Hao, Xu Zheng, Taylor Gavin, et al. Visualizing the loss landscape of neural nets.Proceedings of the 32nd International Conference on Neural Information Processing Systems,2018: 6391–6401.
[15] Noh Hyeonwoo, Hong Seunghoon, Han Bohyung. Learning Deconvolution Network for Semantic Segmentation,2015: 1520-1528.
[16] Ming Liang, Xiaolin Hu. Recurrent convolutional neural network for object recognition,2015: 3367-3375.
[17] Chai Chengping, Maceira Monica, Santos-Villalobos Hector J., et al. Using a Deep Neural Network and Transfer Learning to Bridge Scales for Seismic Phase Picking. Geophysical Research Letters, 2020, 47 (16): e2020GL088651.
[18] Lapins Sacha, Goitom Berhe, Kendall J. Michael, et al. A Little Data Goes a Long Way: Automating Seismic Phase Arrival Picking at Nabro Volcano With Transfer Learning. Journal of Geophysical Research: Solid Earth, 2021, 126 (7): e2021JB021910.
[19] El Zini J., Rizk Y., Awad M. A Deep Transfer Learning Framework for Seismic Data Analysis: A Case Study on Bright Spot Detection. Ieee Transactions on Geoscience and Remote Sensing, 2020, 58 (5): 3202-3212.
[20] Münchmeyer Jannes, Woollam Jack, Rietbrock Andreas, et al. Which picker fits my data? A quantitative evaluation of deep learning based seismic pickers. Journal of Geophysical Research: Solid Earth, 2022, 127 (1): e2021JB023499.