REAVER: Real-time Earthquake Prediction with Attention-based Sliding-Window Spectrograms

Lotfy Abdel Khaliq[1] , **Sabine Janzen**[1] and **Wolfgang Maaß**[1]

[1]Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)

{lotfy.abdel_khaliq, sabine.janzen, wolfgang.maass}@dfki.de

Abstract

Predicting earthquakes with precision remains an ongoing challenge in earthquake early warning systems (EEWS), that struggle with accuracy and fail to provide timely warnings for impending earthquakes. Recent efforts employing deep learning techniques have shown promise in overcoming these limitations. However, current methods lack the ability to capture subtle frequency changes indicative of seismic activity in real-time, limiting their effectiveness in EEWS. To address this gap, we propose REAVER, a novel approach for real-time prediction of P- and S-waves of earthquakes using attention-based sliding-window spectrograms. REAVER leverages Mel-Spectrogram signal representations to capture temporal frequency changes in seismic signals effectively. By employing an encoder-decoder architecture with attention mechanisms, REAVER accurately predicts the onset of P- and S-waves moments when an earthquake occurs. We benchmark the effectiveness of REAVER, showing its performance in terms of both accuracy and real-time prediction capabilities compared to existing methods. Additionally, we provide a web-based implementation of REAVER, allowing users to monitor seismic activity in real-time and analyze historical earthquake waveforms.

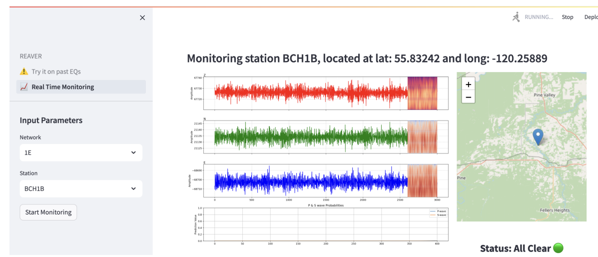

Figure 1: REAVER's interface, showing real-time monitoring of seismic activity at station BCH1B in Canada. It illustrates Mel-Spectrogram visualizations for three channels (Z, N, E) and the probabilities of P and S waves, with a map indicating the station's location.

1 Introduction

Despite the pressing need to alert populations and safeguard essential infrastructure, up til now it is not possible to predict specific earthquakes with certainty, e.g., the Turkey–Syria earthquakes in 2023 [Kwiatek *et al.*, 2023].

Existing earthquake early warning systems (EEWS) like the ShakeAlert system used in US West Coast [Kohler *et al.*, 2020] use on-site $\tau_c - P_d$ algorithm to estimate the arrival of a P-phase, which is the primary seismic wave produced by an earthquake. This algorithm is based on τ_c parameter that captures the wave's dominant period and P_d for its initial amplitude, in combination with classic short-term average and long-term average (STA/LTA) detection approaches [Gaol *et al.*, 2021]. Those approaches are limited in their accuracy as well as the time period between warning and occurrence of the earthquake; furthermore, they cannot generalize to earthquake signals from other regions [Lara *et al.*, 2023].

Recently, deep learning methods were applied to overcome those limitations in EEWS. Several methods use fully-connected network, CNNs and RNNs applied on the raw earthquake waveforms to classify between earthquakes and impulsive noises [Ku *et al.*, 2020; Fauvel *et al.*, 2020; Huang *et al.*, 2020; Meier *et al.*, 2019; Mallouhy *et al.*, 2019] or to estimate the magnitude of earthquakes and its source characterization [Munchmeyer *et al.*, 2021; Ochoa *et al.*, 2018]. Furthermore, CNN-based architectures are applied to phase-picking for predicting arrival times of P waves as well as S waves, i.e., more destructive secondary or shear waves of earthquakes, e.g., PhaseNet [Zhu and Beroza, 2019]. This approach is extended by the EQTransformer by employing a combination of CNNs, RNNs and Transformer attention applied on longer 60-seconds waveforms [Mousavi *et al.*, 2020]. While such methods show high detection accuracy of P and S waves, they are trained with raw long waveforms containing centered P or S waves from the earthquake signal. Other methods combine CNN-RNN architectures with Mel-Spectrograms for binary earthquake classification [Shakeel *et al.*, 2021; Mukherjee *et al.*, 2021]. These methods do not capture the instant tiniest changes in frequency when an earthquacke occurs and limits their performance in EEWS to enable real-time detection of an earthquake. But, predicting an earthquake even just a second before it happens can have significant positive impacts on reducing damage and en-

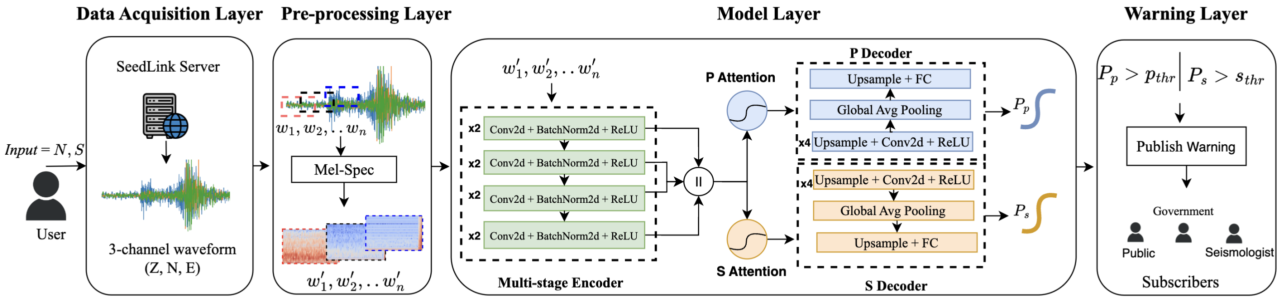

Figure 2: REAVER's Architecture with its four layers: Data Acquisition, Pre-Processing, Model Prediction, and Warning Generation.

hancing public safety [McBride *et al.*, 2022], e.g., automated systems activation like shutting off gas lines, brief warnings for allowing people to take cover under a desk or move away from windows and switch of life-support systems in hospitals to emergency power to ensure continuity.

In this paper, we propose REAVER - an approach for predicting P- and S-waves of earthquakes in real-time based on attention-based sliding-window spectograms. Instead of relying on raw 3-channel waveforms, we adopt the widely used Mel-Spectogram signal representation in sound to represent the frequency amplitudes in each channel in the signal across time [Gong *et al.*, 2022; Meng *et al.*, 2019]. Combined with a sliding window approach, we are able to instantly capture the temporal frequency changes in the signal using an encoder equipped with attention based separate decoders for the continuous probabilities of P and S waves. REAVER not only significantly enhances detection speed, needing only 0.08 seconds of the waveform to predict P-waves with up to 70% faster than existing approaches, but also achieves a higher classification accuracy of 98.81%. The proposed approach was implemented into a web-based EEWS that allows users to monitor seismic stations for real-time earthquake warnings or to analyze past earthquake waveforms (cf. Fig. 1). The source code of REAVER is available at the following GitHub repository[1] along with a demonstration video showing how it works[2].

2 System Overview

REAVER is composed of 4 layers: 1) Data acquisition layer that gets real-time and historical waveforms, 2) Pre-processing layer which divides the signal into overlapping windows and computes the Mel-Spectogram for each window, 3) Model layer which predicts the continuous probabilities of P and S waves, and 4) Warning layer which publishes warnings to subscribers when the detection of P or S waves exceeds a threshold to take an instant response (cf. Fig. 2).

2.1 Data Acquisition Layer

The data acquisition layer is responsible for collecting the historical waveform data for model training and real-time data from seismic stations for inference. For real-time data acquisition, the user chooses a network and a station from an inventory provided by Obspy[3]- a python package for seismology. The chosen variables are encapsulated within a request to acquire real-time data from IRIS real-time server[4] for later processing. For model training, the layer collects historical waveforms data from STEAD [Mousavi *et al.*, 2019], which contains 1.2 million earthquake and noise waveforms with a sampling rate of 100 Hz. Each waveform is represented as 3-channel time-series $X_c(t)$ where c corresponds to the vertical (Z), north-south (N), and east-west (E) components of seismographs.

2.2 Pre-Processing Layer

The pre-processing layer receives the raw waveform and slices it into overlapping windows across each channel. Unlike [Mousavi *et al.*, 2020; Zhu and Beroza, 2019] which use the full raw waveform, the slicing approach allows a more detailed and continuous analysis of the signal over time, which is crucial for real-time EEWS. Formally, we segment each channel into overlapping windows as follows:

$$S = W - \left(\frac{O}{100} \times W\right), \quad N = 1 + \left\lfloor \frac{L - W}{S} \right\rfloor \tag{1}$$

where S is the step size, W is the window length, $\frac{O}{100}$ is the overlap percentage, L is the length of each channel in the signal, and N is the total number of windows per channel. The output is a list of windows $w_i^c = X_c[s_i : e_i]$ for each channel c and window i, where s_i and e_i are the start and end indices of the i^{th} window, respectively. Instead of using the raw waveform of signal amplitudes, we convert the signal in each window into its Mel-Spectrogram representation by computing the Short-Time Fourier Transform (STFT) as follows:

$$\text{STFT}\{x(t)\}(m, \omega) = \sum_{n=-\infty}^{\infty} x(n) \cdot w(n-m) \cdot e^{-j\omega n} \tag{2}$$

where $w(n)$ is the window function, m is the time index, and ω is the frequency index. Then, a set of k triangular filters $\{H_k(m)\}$ is applied to the STFT, where each filter corresponds to a Mel frequency. The log Mel-Spectrogram for each window is then obtained as:

$$w_i'^c = \log\left(\sum_m |\text{STFT}\{w_i^c(t)\}(m, \omega)| \cdot H_k(m)\right) \tag{3}$$

[1]https://github.com/InformationServiceSystems/pairs-project/tree/main/Modules/CrisisImaginator/REAVER

[2]https://www.youtube.com/watch?v=PIRmRzfhN-k

[3]https://github.com/obspy/obspy/wiki/

[4]https://www.iris.edu/hq/sage

where c is the channel index.

2.3 Model Layer

The model layer takes input the Mel-Spectogram windows for each channel w' and passes it to the REAVER model. The model is composed of a multi-stage encoder, an attention layer for each stage, and two separate decoders for P and S waves. The encoder consists of 4 stages, where each stage S_i is defined as a sequence of two blocks. Each block applies a Conv2D layer followed by BatchNorm2d and ReLU activation function. Given the set of Mel-Spectogram windows W', the operation of a single block can be represented as:

$$\text{Block}(W') = \text{ReLU}\left(\text{BatchNorm2d}\left(\text{Conv2d}\left(W'\right)\right)\right)$$

Therefore, each stage in the encoder can be expressed as the composition of two such blocks:

$$S_i = \text{Block}(\text{Block}(W_i'))$$

where W_i' is the output from the previous stage or the input to the encoder for $i = 1$. Subsequently, P Decoder and S Decoder each apply spatial self-attention to the outputs from each encoder stage, which can be written as:

$$\text{Attn}(S_i) = \text{Conv2d}\left(S_i \odot \sigma\left(\text{Conv2d}(S_i)\right)\right)$$

where σ denotes the sigmoid function and $\odot$ represents element-wise multiplication. The attention mask generated by the sigmoid function modulates the feature map for each branch to focus on relevant features for P and S waves separately. Each decoder consists of sequences of such $Attn()$ layers followed by upsampling and convolution operations. Formally, for each stage S_i of the encoder, the corresponding decoder block output can be represented as:

$$D_i = \text{Upsample}\left(\text{Conv}\left(\text{Attn}(S_i)\right)\right)$$

where D_i is the output of the i-th decoder block. Finally, the continuous probabilities for P and S wave, denoted as P_p for the P Decoder and P_s for the S Decoder, can be expressed as:

$$P_p = \text{FC}\left(\text{Upsample}\left(\text{GlobalAvgPool}(D_P)\right)\right)$$
$$P_s = \text{FC}\left(\text{Upsample}\left(\text{GlobalAvgPool}(D_S)\right)\right)$$

where D_P and D_S are the aggregated outputs of the P Decoder and S Decoder blocks, respectively. We utilize BinaryCrossEntropy loss for supervising both P_p and P_s.

2.4 Warning Layer

The warning layer evaluates the probabilities P_p and P_s from the model layer and issues warnings to subscribers, including individuals or government entities, when $P_p > P_{\text{thr}}$ or $P_s > S_{\text{thr}}$, aiming for timely Earthquake response. P_{thr} and S_{thr} are hyperparameters which are set to 0.7 upon extensive experiments to optimize the precision-recall trade-off.

3 Evaluation

To evaluate our method we use a dataset consisting of 51,510 earthquakes and 11,773 noise waveforms from STEAD. We simulate real-time detection by applying a 4-second sliding window with 99% overlap before the P- and S-wave arrival times. We compare our method against classical STA/LTA method [Choubik *et al.*, 2020], and deep learning methods Phasenet [Zhu and Beroza, 2019] and EQTransformer [Mousavi *et al.*, 2020] which were originally trained using 30- and 60-second windows respectively. We apply the same sliding window configuration across all methods for a fair comparison. In our evaluation, we prioritize two critical dimensions of earthquake detection performance: the speed of P-wave detection and the accuracy of distinguishing between earthquake and noise waveforms. Speed of detection is assessed through Δt, representing the time difference between the algorithm's detection of the P-wave and its actual occurrence. Additionally, we evaluate classification accuracy, measuring each method's ability to correctly identify earthquake waveforms against impulsive noise.

Table 1: Performance comparison of P-wave detection time statistics, showing the mean time difference Δt_{mean}, standard deviation $\sigma_{\Delta t}$, median Δt_{median}, 25th percentile (Q1) $\Delta t_{25\%}$, and 75th percentile (Q3) $\Delta t_{75\%}$.

Method	Δt_{mean} (s)	$\sigma_{\Delta t}$ (s)	Median (s)	Q1 (s)	Q3 (s)
Phasenet	1.67	2.01	0.83	0.59	1.93
Eqtransformer	3.73	4.48	1.62	1.20	4.00
STA/LTA	0.27	0.31	0.15	0.04	0.40
REAVER	**0.08**	**0.16**	**0.04**	**0.03**	**0.12**

Table 2: Classification performance on the testing dataset.

Method	Precision	Recall	F1 Score	Accuracy (%)
Phasenet	0.98	0.96	0.97	95.46
Eqtransformer	0.99	0.91	0.95	92.50
STA/LTA	0.96	0.85	0.89	85.56
REAVER	**0.99**	**0.98**	**0.99**	**98.81**

Table 1 shows the distribution statistics of Δt over the testing dataset for all the methods. We can observe that REAVER significantly enhances detection timeliness, improving by 70% over STA/LTA and around 95% over Phasenet and Eqtransformer in mean detection time difference (Δt_{mean}). In Table 2, we show the classification performance of differentiating between earthquake and noise waveforms, recording a correct detection when either a P or S wave was detected in the waveform. We observe that REAVER outperforms other methods in accurately detecting earthquakes with overall accuracy 98.81%.

4 Conclusion

In this paper, we proposed REAVER, a novel method for real-time earthquake prediction utilizing attention-based sliding-window spectrograms. Our evaluation demonstrates that our methos not only achieves high accuracy in differentiating earthquakes from noise, but also offers faster detection times compared to existing methods. Furthermore, REAVER's web-based implementation allows for real-time earthquake monitoring and historical waveform analysis, making it a valuable tool for both individuals and professionals in seismology.

References

[Choubik *et al.*, 2020] Youness Choubik, Abdelhak Mahmoudi, Mohammed Majid Himmi, and Lahcen El Moudnib. Sta/lta trigger algorithm implementation on a seismological dataset using hadoop mapreduce. *IAES International Journal of Artificial Intelligence*, 9(2):269, 2020.

[Fauvel *et al.*, 2020] Kevin Fauvel, Daniel Balouek-Thomert, Diego Melgar, Pedro Silva, Anthony Simonet, Gabriel Antoniu, Alexandru Costan, Véronique Masson, Manish Parashar, Ivan Rodero, et al. A distributed multi-sensor machine learning approach to earthquake early warning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 403–411, 2020.

[Gaol *et al.*, 2021] YH Lumban Gaol, RK Lobo, SS Angkasa, A Abdullah, I Madrinovella, S Widyanti, A Priyono, SK Suhardja, AD Nugraha, Z Zulfakriza, et al. Preliminary results of automatic p-wave regional earthquake arrival time picking using machine learning with sta/lta as the input parameters. In *IOP conference series: earth and environmental science*, volume 873, page 012060. IOP Publishing, 2021.

[Gong *et al.*, 2022] Yuan Gong, Cheng-I Lai, Yu-An Chung, and James Glass. Ssast: Self-supervised audio spectrogram transformer. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 10699–10709, 2022.

[Huang *et al.*, 2020] Xin Huang, Jangsoo Lee, Young-Woo Kwon, and Chul-Ho Lee. Crowdquake: A networked system of low-cost sensors for earthquake detection via deep learning. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 3261–3271, 2020.

[Kohler *et al.*, 2020] Monica D Kohler, Deborah E Smith, Jennifer Andrews, Angela I Chung, Renate Hartog, Ivan Henson, Douglas D Given, Robert de Groot, and Stephen Guiwits. Earthquake early warning shakealert 2.0: Public rollout. *Seismological Research Letters*, 91(3):1763–1775, 2020.

[Ku *et al.*, 2020] Bonhwa Ku, Gwantae Kim, JaeKwang Ahn, Jimin Lee, and Hanseok Ko. Attention-based convolutional neural network for earthquake event classification. *IEEE Geoscience and Remote Sensing Letters*, 18(12):2057–2061, 2020.

[Kwiatek *et al.*, 2023] G Kwiatek, P Martínez-Garzón, Dirk Becker, Georg Dresen, Fabrice Cotton, Gregory C Beroza, D Acarel, S Ergintav, and Marco Bohnhoff. Months-long seismicity transients preceding the 2023 mw 7.8 kahramanmaraş earthquake, türkiye. *Nature Communications*, 14(1):7534, 2023.

[Lara *et al.*, 2023] Pablo Lara, Quentin Bletery, Jean-Paul Ampuero, Adolfo Inza, and Hernando Tavera. Earthquake early warning starting from 3 s of records on a single station with machine learning. *Journal of Geophysical Research: Solid Earth*, 128:e2023JB026575, 2023.

[Mallouhy *et al.*, 2019] Roxane Mallouhy, Chady Abou Jaoude, Christophe Guyeux, and Abdallah Makhoul. Major earthquake event prediction using various machine learning algorithms. In *2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*, pages 1–7. IEEE, 2019.

[McBride *et al.*, 2022] Sara K McBride, Hollie Smith, Meredith Morgoch, Danielle Sumy, Mariah Jenkins, Lori Peek, Ann Bostrom, Dare Baldwin, Elizabeth Reddy, Robert de Groot, et al. Evidence-based guidelines for protective actions and earthquake early warning systems. *Geophysics*, 87(1):WA77–WA102, 2022.

[Meier *et al.*, 2019] Men-Andrin Meier, Zachary E Ross, Anshul Ramachandran, Ashwin Balakrishna, Suraj Nair, Peter Kundzicz, Zefeng Li, Jennifer Andrews, Egill Hauksson, and Yisong Yue. Reliable real-time seismic signal/noise discrimination with machine learning. *Journal of Geophysical Research: Solid Earth*, 124(1):788–800, 2019.

[Meng *et al.*, 2019] Hao Meng, Tianhao Yan, Fei Yuan, and Hongwei Wei. Speech emotion recognition from 3d log-mel spectrograms with deep learning network. *IEEE access*, 7:125868–125881, 2019.

[Mousavi *et al.*, 2019] S Mostafa Mousavi, Yixiao Sheng, Weiqiang Zhu, and Gregory C Beroza. Stanford earthquake dataset (stead): A global data set of seismic signals for ai. *IEEE Access*, 7:179464–179476, 2019.

[Mousavi *et al.*, 2020] S Mostafa Mousavi, William L Ellsworth, Weiqiang Zhu, Lindsay Y Chuang, and Gregory C Beroza. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications*, 11(1):3952, 2020.

[Mukherjee *et al.*, 2021] Tonumoy Mukherjee, Chandrani Singh, and Prabir Kumar Biswas. A novel approach for earthquake early warning system design using deep learning techniques, 2021.

[Munchmeyer *et al.*, 2021] Jannes Munchmeyer, Dino Bindi, Ulf Leser, and Frederik Tilmann. Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network. *Geophysical Journal International*, 226(2):1086–1104, 2021.

[Ochoa *et al.*, 2018] Luis H. Ochoa, Luis F. Niño, and Carlos A. Vargas. Fast magnitude determination using a single seismological station record implementing machine learning techniques. *Geodesy and Geodynamics*, 9(1):34–41, 2018. Seismological advances in Latin America.

[Shakeel *et al.*, 2021] Muhammad Shakeel, Katsutoshi Itoyama, Kenji Nishida, and Kazuhiro Nakadai. Detecting earthquakes: a novel deep learning-based approach for effective disaster response. *Applied Intelligence*, 51(11):8305–8315, 2021.

[Zhu and Beroza, 2019] Weiqiang Zhu and Gregory C Beroza. Phasenet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1):261–273, 2019.