National College of Ireland

PhaseNet and EfficientNet-B0 for Phase Detection and Arrival Time Prediction

MSc Research Project
Data Analytics

Nagalakshmi Ramakrishna
Student ID: x23183829
x23183829@student.ncirl.ie

School of Computing
National College of Ireland

Supervisor: Christian Horn

National College of Ireland Project Submission Sheet School of Computing

Student Name:	Nagalakshmi Ramakrishna
Student ID: x23183829@student.ncirl.ie	x23183829
Programme:	Data Analytics
Year:	2024
Module:	MSc Research Project
Supervisor:	Christian Horn
Submission Due Date:	12/12/2024
Project Title:	PhaseNet and EfficientNet-B0 for Phase Detection and Arrival Time Prediction
Word Count:	8,284
Page Count:	22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Nagalakshmi Ramakrishna
Date:	11th December 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

PhaseNet and EfficientNet-B0 for Phase Detection and Arrival Time Prediction

Nagalakshmi Ramakrishna
x23183829
x23183829@student.ncirl.ie

Abstract

Seismic phase detection and arrival time prediction are crucial for earthquake monitoring and early warning systems. This study evaluates the performance of two advanced deep learning models, PhaseNet and EfficientNet-B0, on the INSTANCE dataset for phase detection and the picking of P and S waves using spectrogram and waveform data. PhaseNet achieved a testing accuracy of 94.7%, demonstrating its effectiveness in classifying seismic events. Conversely, EfficientNet-B0 excelled in arrival time prediction, achieving a P-wave MAE of 279 ms and an S-wave MAE of 255 ms, surpassing PhaseNet in regression tasks. While PhaseNet exhibited faster training convergence, EfficientNet-B0 delivered superior accuracy and gen- eralization. This paper highlights the complementary strengths of PhaseNet and EfficientNet-B0 in seismic phase detection and arrival time picking, contributing to advancements in seismic monitoring methodologies.

Keywords: Seismic phase detection, P-wave, S-wave, PhaseNet, EfficientNet- B0, Deep learning, Earthquake monitoring, INSTANCE dataset

1 Introduction

1.1 Background

Earthquake is a natural disaster which hard to predict and occurs suddenly that results in a large scale of properties and in some cases hundreds of losses of lives. They attack when least expected – wiping off structures, ripping through concrete roads, and ways of living. Some of the fallout remains a state of shock and grief in the affected communities many of which take years in rebuilding

The automatic detection and phase picking of earthquakes are two important aspects of seismology research ScienceDirect (2024). The identification of earthquake signals is referred to as phase detection, while predicting the arrival times of P and S waves is known as phase picking. Phase detection plays a crucial role in the early prediction of earthquakes, while phase picking aids in determining the magnitude, pinpointing the hy- pocenter location, and conducting spectral analysis.

Early earthquake detection methods are classified into three main types. The first involves amplitude- and energy-ratio-based techniques, such as Short-Time Average (STA) /Long-Time Average (LTA), effective for seismic event detection (Allen (1978); Withers

et al. (1998)). The second includes waveform similarity-based approaches like template matching, which offers high catalog completeness but requires pre-prepared templates, has low efficiency, and cannot directly determine seismic phase arrival times (Peng and Zhao (2009); Shelly et al. (2007)). The third category, early machine learning-based methods, was limited by hardware performance, struggled with complex calculations, and saw limited adoption (Wang and Teng (1997); Gentili and Michelini (2006)).

Deep learning has become a widely utilized and impactful approach across various applications. In earthquake monitoring, there is a growing demand for efficient and reliable tools to handle the ever-growing volumes of data. With its straightforward conceptual framework and the abundance of labeled datasets, earthquake detection and phase pick- ing present compelling opportunities for the application of advanced machine learning techniques in seismology. **?**.

1.2 Research Question

"How does the application of the EfficientNet architecture to spectrogram and waveform plots of seismic signals compare to PhaseNet in terms of accur- acy and performance for phase detection and picking of both P and S waves on the INSTANCE dataset?"

1.3 Research Objectives

The objectives of this research are as follows.

- To evaluate the performance of the EfficientNet model for phase detection and picking of P and S waves using the INSTANCE dataset.
- Investigate how raw waveform data and spectrogram representations contribute to phase detection and picking accuracy and efficiency in seismic event analysis.
- To evaluate the performance of the PhaseNet model for phase detection and picking of P and S waves using the INSTANCE dataset.
- Comparison of the performance of EfficientNet and PhaseNet in terms of accuracy and efficiency for phase detection and picking of P and S waves on the INSTANCE dataset.

1.4 Scope of Study

In this research,By leveraging advanced deep learning algorithms, specifically PhaseNet (Zhu and Beroza (2019)) and EfficientNet (Tan and Le (2019)), to enhance earthquake detection and seismic wave arrival time picking. PhaseNet, while highly accurate in de- tecting P-wave arrivals, has demonstrated challenges in reliably detecting S-waves. To overcome this limitation, we explore the potential of EfficientNet, a state-of-the-art model known for its scalable architecture across depth, width, and resolution, offering a more efficient solution for detecting both P and S waves in seismic data.

Phase detection and picking techniques often rely on raw ground motion signals, which

are processed through neural networks. However, transforming these signals into spectrograms or waveform images can significantly improve the performance of convolutional neural networks (CNNs). Spectrograms, which represent the frequency domain of seismic signals over time, provide valuable insights into the frequency characteristics of seismic events, allowing models to better differentiate between earthquakes and noise. Waveform images, on the other hand, retain the temporal dynamics of seismic events, making them critical for detecting the precise arrival times of P and S waves.

While CNNs are commonly designed with a fixed resource budget, their performance can be enhanced by scaling them to improve accuracy based on available resources. Efficient- Net offers a principled approach to model scaling, achieving remarkable accuracy while maintaining computational efficiency. This research aims to evaluate the effectiveness of EfficientNet applied to both spectrogram and waveform plots of seismic signals, compar- ing its performance to PhaseNet in terms of accuracy, efficiency, and overall robustness. Additionally, we investigate the generalizability of both models when trained on the IN-STANCE dataset and assess their performance across various seismic signal types and environmental conditions.

2 Related Work

For decades, the field of seismology has aimed to automate the detection and picking of seismic phase arrivals, a critical step in earthquake monitoring and cataloging. These efforts have transitioned from conventional energy-based methods to sophisticated deep- learning models, offering notable improvements in accuracy, robustness, and efficiency.

2.1 Conventional Algorithms for Phase Picking and Detection

The Short-Time Average to Long-Time Average (STA/LTA) method, introduced by Allen (1978), remains one of the most widely adopted techniques in seismic phase picking. This method calculates the ratio of amplitudes within short- and long-time windows, triggering a detection when the ratio exceeds a predefined threshold. STA/LTA's simplicity made it a cornerstone for local earthquake detection, but its limitations, particularly in handling noisy environments and weak signals, became evident over time.

Subsequent enhancements by Withers et al. (1998) introduced adaptive window lengths and spectral adjustments, significantly improving the method's ability to adapt to vari- able noise conditions. These refinements extended STA/LTA's applicability to teleseismic and regional events, addressing some of the challenges posed by high-noise environments. Similarly, Baer and Kradolfer (1987) proposed an automatic phase picker based on nonlin- ear amplifiers and dynamic thresholds. This innovation enabled more precise detection of weak signals, providing a solution for applications where traditional STA/LTA struggled. Template matching, another influential technique, was introduced by Shelly et al. (2007). By employing cross-correlation with known earthquake waveforms, this method effectively identified small or previously undetected events. However, its reliance on pre- existing templates restricted its ability to detect novel seismic phenomena, a significant drawback for real-time applications Duputel et al. (2019). Moreover, template matching's computational demands limited its scalability, particularly for large datasets. Similarly, auto-correlation methods have shown promise in identifying recurring events but remain memory-intensive despite optimizations like fingerprinting Yoon et al. (2015).

While these conventional methods laid the foundation for automated phase picking, they often struggled with adaptability and computational efficiency, paving the way for modern deep-learning approaches.

2.2 Deep Learning Approaches

Deep-learning techniques have transformed seismic analysis, addressing limitations in traditional methods and enabling the processing of large-scale datasets. These models excel in generalizing across diverse conditions, reducing noise, and accurately picking seismic phases Mousavi and Beroza (2022).

PhaseNet and U-Net-Based Models: PhaseNet, introduced by Zhu and Beroza (2019), was among the first deep-learning models designed explicitly for seismic phase picking. Built on the U-Net architecture, PhaseNet processes three-component seismo- grams and outputs probability distributions for P-waves, S-waves, and noise. Trained on over 700,000 labeled waveforms from the Northern California Earthquake Data Center, it achieved F1 scores of 0.896 for P arrivals and 0.801 for S arrivals, surpassing traditional STA/LTA-based methods. However, PhaseNet's reliance on region-specific training data limits its generalizability to new geographic regions or areas with insufficient labeled datasets.

ARRU Phase Picker Liao et al. (2021), a U-Net extension, addresses this limitation by integrating soft attention gates (AGs) and recurrent-residual convolution units (RRCUs). These architectural enhancements enable ARRU to filter irrelevant waveform responses and capture temporal relationships. The model achieved F1 scores of 98.62% for P-phases and 95.16% for S-phases, demonstrating robust performance across diverse datasets, in- cluding unseen Southern California data. This generalizability makes ARRU particularly well-suited for large-scale seismic monitoring.

Despite their success, both PhaseNet and ARRU struggle in regions with unique geolo- gical characteristics or noise profiles. To address this, Chai et al. (2020) demonstrated the efficacy of transfer learning (TL) in adapting pre-trained models to new datasets. TL has been shown to improve precision and recall by up to 10%, significantly reducing manual effort in seismic phase picking. When combined with double-difference tomography, TL enhances seismic event location accuracy in complex environments like hydraulic fractur- ing. These integrations not only improve accuracy but also achieve over 99% reductions in manual intervention, underscoring their scalability and efficiency.

SEA-net and Sequence Attention Models: SEA-net, proposed by Hou et al. (2023), is a Sequence Attention Network that combines temporal convolutional layers with attention mechanisms. This architecture focuses on critical waveform features, enabling robust detection of seismic events amidst noise. SEA-net achieved a 93.8% accuracy and 95.2% precision for P- and S-wave detections, significantly outperforming earlier methods like PhaseNet. Additionally, SEA-net reduced computational costs by 35%, making it suitable for real-time applications like earthquake early warning systems. Compared to traditional methods, SEA-net improved P-wave detection rates by 22% and S-wave accuracy by 18%, setting a benchmark for precision and computational efficiency.

Spectrogram-Based Approaches: Spectrograms have emerged as a powerful tool in seismic analysis, offering frequency-domain features often overlooked in raw waveform data. Choi et al. (2024) demonstrated the effectiveness of transforming seismic signals into spectrograms using Short-Time Fourier Transform (STFT). These spectrograms, processed through a U-Net architecture, achieved strong performance metrics, including

an MSE of 0.0031 and an MAE of 0.0177 for precise P-wave arrival detection. Simil- arly, CRED Mousavi et al. (2019), a hybrid model combining convolutional layers and bidirectional LSTMs, utilized spectrograms to detect microearthquakes as small as 1.3 ML, even under noisy conditions. These findings highlight the value of incorporating frequency-domain data into seismic phase picking workflows.

Comparative Studies and Hybrid Models: Comparative analyses between PhaseNet and EQTransformer, conducted by Jiang et al. (2021), reveal the strengths and weak- nesses of these models. EQTransformer, which integrates CNN and RNN layers with hierarchical attention mechanisms, achieved superior precision and recall scores of 99.0% for P-waves and 96.0% for S-waves on the Stanford Earthquake Dataset (STEAD). While EQTransformer outperformed PhaseNet in most scenarios, its increased complexity can hinder efficiency in simpler applications. Both models exhibit reduced performance in new regions, emphasizing the importance of retraining or transfer learning to address data variability.

Hybrid models like U-Net++ Guo (2021) incorporate Gaussian noise during training, significantly enhancing their robustness to noisy environments. These models outperform traditional STA/LTA techniques, particularly in scenarios with high background noise, further demonstrating the adaptability of advanced deep-learning architectures.

Phase Association and Multi-Station Data: Phase association, or the linking of seismic phases across multiple stations, is essential for constructing comprehensive earthquake catalogs. Traditional approaches often rely on travel-time data, which can struggle in high-noise environments or during periods of heightened seismicity. FastLink Yu and Wang (2022), which leverages CNN-based methods and GPU acceleration, has demonstrated remarkable performance in phase association. During the Yangbi Ms 6.4 aftershock sequence, FastLink processed over 2 million picks, associating 17,000 events with an accuracy of 93.5%, achieving a 12-fold reduction in processing time compared to conventional methods.

Multi-station data integration, exemplified by ArrayConvNet Shen and Shen (2021), further improves model generalization and robustness. By leveraging spatially distributed networks, ArrayConvNet achieved 99.4% accuracy in earthquake localization in Hawai'i, highlighting the potential of combining spatial and temporal data for enhanced seismic monitoring.

2D Image-Based Models: Several studies have explored treating seismic data as images, applying 2D convolutional neural networks (CNNs) to detect and classify seismic phases. MT-Net ZHANG et al. (2021), an extension of U-Net, leverages 2D convolutional operations on spectrograms to enhance the detection of first-arrival phases (FAPs) amidst noise. Similarly, Yang et al. (2021) demonstrated the robustness of image-based models in handling diverse datasets, showcasing their applicability for long-term monitoring where station configurations vary significantly.

The evolution of seismic phase picking, from foundational STA/LTA methods Al- len (1978) to advanced deep-learning models like EQTransformer Mousavi et al. (2020), reflects a remarkable trajectory of innovation. By integrating traditional insights with modern techniques, seismic monitoring systems continue to push the boundaries of ac- curacy, scalability, and adaptability, paving the way for future advancements.

3 Methodology

This research analyzes the performance of the PhaseNet and EfficientNet architectures in classifying earthquake and noise signals. The EfficientNet regression model is also employed to predict the arrival times of P and S waves. The study adopts a modified CRISP-DM methodology, tailored for scientific research, and outlines the steps involved in classifying seismic signals using both spectrograms and waveform plots derived from raw seismic data.

3.1 Problem Understanding

Seismic phase detection and picking are critical for understanding earthquake dynamics and enhancing early warning systems. This research addresses the challenges of ac- curately classifying earthquake and noise signals and predicting P and S wave arrival times. The objective is to evaluate the effectiveness of deep learning models, specifically PhaseNet and EfficientNet, in handling seismic data, leveraging both waveform plots and spectrogram representations.

3.2 Data Understanding

The INSTANCE dataset is a comprehensive resource of seismic waveform data and metadata, specifically designed for machine learning-based analysis. It includes data from 54,008 earthquakes, amounting to 1,159,249 three-channel waveforms, along with 132,330 three-channel noise waveforms. Each waveform is accompanied by detailed metadata con- taining 115 attributes, providing information about the station, trace, source, path, and quality. The dataset spans 19 networks and 620 seismic stations, offering diverse geo- graphic and seismic event coverage. For our research, we have used this dataset to train and evaluate the PhaseNet and EfficientNet models for classifying seismic signals and predicting the arrival times of P and S waves. *(Michelini et al. (2021)).*

For this research, we used version 3 of the INSTANCE dataset, which includes wave- form data in HDF5 format and metadata in CSV format. Metadata files are linked to waveform data through trace names for efficient analysis.

Waveform Data	Instance.events counts 10k.hdf5	10,000 events
	Instance.events gm 10k.hdf5	10,000 events
	Instance.noise 1k.hdf5	1,000 instances
Metadata	metadata.Instance noise 1k.csv	Shape(1000,43)
	metadata.Instance events 10k.csv	Shape (10000 ,115)

Table 1: Dataset Description

Table 1provides a brief description of the details of the dataset used in this research. Since our study focuses on distinguishing between noise and earthquake phases and ac- curately picking P and S wave arrivals, we used the event-related HDF5 file. This dataset provides labeled seismic events critical for phase detection and picking tasks, whereas the ground motion data introduces additional complexity not required for this scope.

Since the number of events and noise samples are unequal, we utilized additional noise data from the main INSTANCE dataset, which originally has a shape of (132,288,

43). To prevent model overfitting, we balanced the dataset, ensuring an equal number of earthquake and noise samples for phase detection. For phase picking, however, the noise data is excluded, as P and S arrival times are only present in the event data

3.3 Data Pre-processing and Analysis

Before training the neural network, the dataset was pre-processed to ensure it was well-structured and balanced.
The event data consists of four key types of metadata: *source*, *station*, *trace*, and *path (event)*, resulting in a shape of 10*,* 000 × 115, with each record providing detailed inform- ation about seismic events. In contrast, the noise data includes two types of metadata: *station* and *trace (noise)*. The original noise dataset contains 132,288 events, but to maintain a balanced dataset and prevent model bias during phase detection, the noise data was reduced to 10,000 samples, matching the number of earthquake events.

Spectrogram and Waveform Generation for Phase Detection and Picking

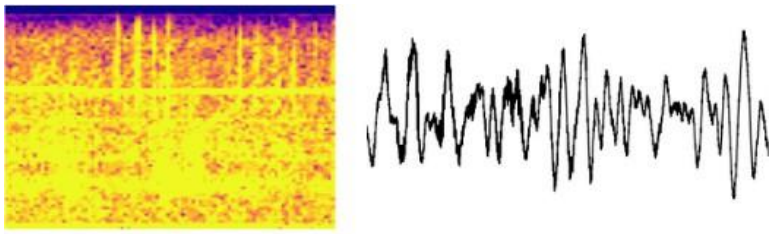


Figure 1: A combined visualization of seismic data: a) Spectrogram and b) Waveform.

This study utilizes both spectrograms and waveforms to analyze seismic data for phase detection and phase picking. Spectrograms, created using short-time Fourier transform (STFT), highlight frequency intensity over time, aiding in distinguishing seismic events from noise and identifying patterns for P and S wave arrivals. Waveforms, as raw time- series representations, are critical for phase picking by allowing precise identification of P and S wave arrival times. Waveforms and spectrograms will be generated using the *trace name* (source ID) from the CSV metadata, which matches with waveform data in the HDF5 file. Each _image will be saved in its respective directory based on the event type (e.g., earthquake or noise). Both representations are systematically generated and stored, leveraging their strengths to create a robust framework for seismic analysis.

Distribution of Station Channels and Network Codes

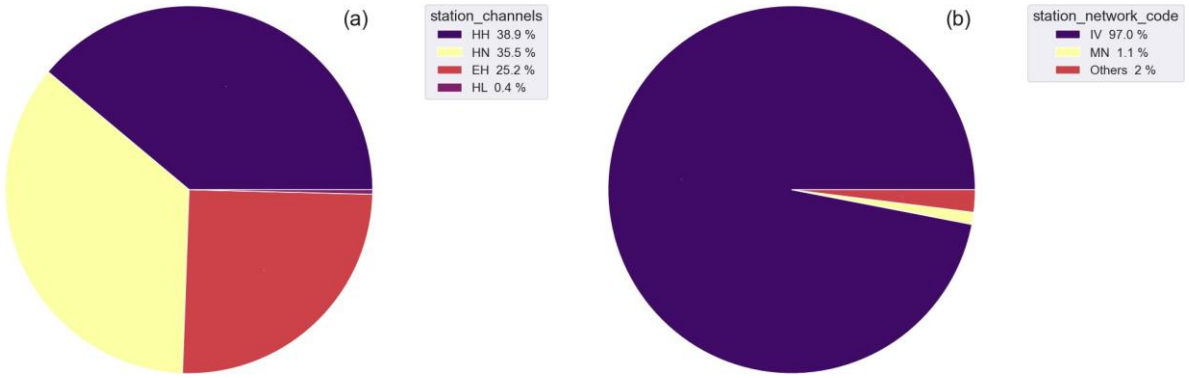


Figure 2: Distribution of seismic data: (a) station channels and (b) station network codes.

Figure 2 illustrates the distribution of seismic data based on station channels and network codes. Subfigure (a) represents the proportion of station channels in the dataset, categorized into four types: HH (38.9%), HN (35.5%), EH (25.2%), and HL (0.4%). These categories indicate different seismic instrument configurations, with HH and HN being the most dominant, reflecting high-frequency and strong motion data, respectively.

Subfigure (b) shows the distribution of station network codes, with IV accounting for 97.0% of the data, MN representing 1.1%, and other networks contributing 2%. The dominance of the IV network highlights its extensive coverage and contribution to the dataset.

These distributions provide valuable insights into the dataset's structure, ensuring a clear understanding of the diversity in station channels and networks. Such information is crucial for validating the generalizability of models trained on this dataset.

Relationship Between Source Magnitude and Hypocentral Distance

Figure 3 shows the relationship between source magnitude and hypocentral distance. Most events have magnitudes between 1 and 3 and occur within 200 km of the source. The color intensity represents the density of events, with darker shades indicating higher concentrations.

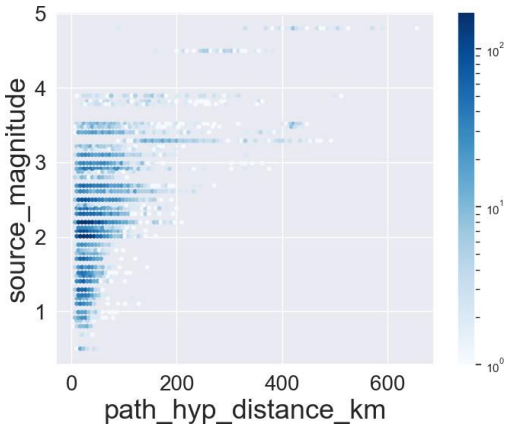


Figure 3: Scatter plot showing the relationship between source magnitude and hypocent- ral distance.

Histograms of P and S Arrival Samples

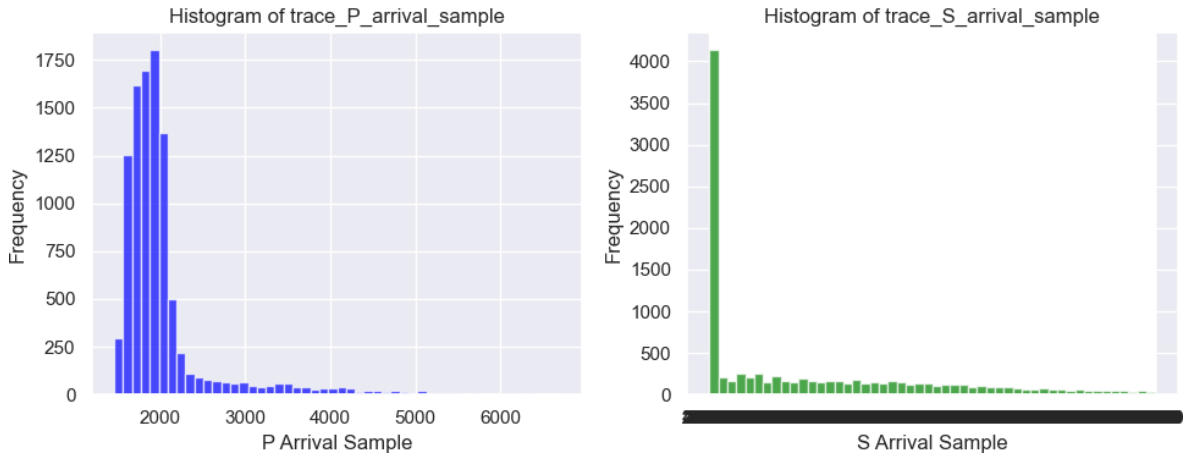


Figure 4: Histograms of arrival samples: (a) P wave arrival samples and (b) S wave arrival samples.

Figure 4 shows the distribution of P and S wave arrival samples in the dataset. Subfig- ure (a) illustrates the histogram for P wave arrivals, with a peak around 1,750 to 2,000 samples, indicating that most P waves arrive in a consistent time range. Subfigure (b) shows the histogram for S wave arrivals, with a sharp peak at the start and a more spread-out distribution, reflecting greater variability in S wave arrival times.
These patterns are critical for phase picking, as accurately identifying P and S wave ar- rivals depends on understanding their distributions. The concentrated P wave arrivals aid in precise picking, while the variability in S arrivals requires models to adapt to a broader range of times.

The target variables for this research are derived from the metadata columns. For phase detection, the variable *source type* is used to classify signals as noise or earthquake. For phase picking, the variables *trace P arrival time* and *trace S arrival time* are used to determine the arrival times of P and S waves, respectively. By following this structured pre-processing approach, the dataset is balanced and optimized for robust neural network training and evaluation.

3.4 Modeling

Once the data was preprocessed, it was prepared for modeling. For phase detection, the dataset was divided into training, testing, and validation sets with a 60:20:20 split. Spectrogram images were organized into separate folders using the splitfolders library. The flow from directory function was used to load images from the respective train, test, and validation directories into data generators. These images were processed in batches of 32 and resized to the required input dimensions for the neural network architectures: $224 \times 224 \times 3$ for EfficientNet-B0 and $128 \times 128 \times 3$ for PhaseNet.

Similarly, waveform data was reshaped and split into training, testing, and validation sets. Using the flow from directory function, the waveform data was efficiently loaded and prepared for model training and evaluation, ensuring compatibility with the input requirements of the respective architectures.

For phase picking, the waveform image paths, along with the P and S arrival times, were organized into a pandas DataFrame. The data was then split into training, testing, and validation sets. Each waveform image path corresponded to an individual waveform image, enabling the models to associate the images with their respective P and S arrival times. This structured approach ensured the data was well-prepared for training models to accurately pick P and S wave arrivals.

3.5 Evaluation

The research evaluates the performance of the EfficientNet and PhaseNet models on the INSTANCE dataset. The metrics used for classification include Precision,Recall, and the F1-score. Precision measures the percentage of correct predictions when identifying earthquakes, while Recall (or sensitivity) evaluates the percentage of detected earthquakes among all actual earthquakes. Since minimizing false negatives is crucial in earthquake detection, Recall serves as a more critical metric than Precision. The F1-score combines Precision and Recall into a single measure, offering a balanced evaluation of the models' performance.

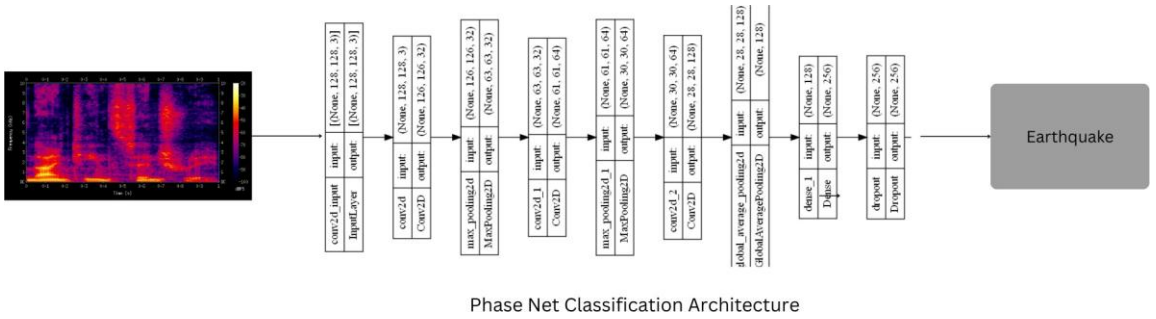
For the regression tasks, metrics such as Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE) were used to assess the models' ability to predict P and S wave arrival times. These metrics indicate how close the predicted values are to the actual values. The models' performance was further validated using baseline errors for comparison, ensuring robustness. Additionally, the F1-score, Precision, and Recall were calculated by considering a prediction as true positive if it fell within a 0.5-second threshold. This comprehensive evaluation framework ensured a reliable assessment of the models' effectiveness in classifying seismic signals and picking phase arrival times.

4 Design Specification

4.1 PhaseNet Architecture

The PhaseNet model, inspired by Zhu and Beroza (2019), has been adapted to a U- Net-like architecture designed for spectrogram inputs of size $128 \times 128 \times 3$. The model structure is modified from U-Net Ronneberger et al. (2015), a deep neural network origin- ally developed for biomedical image processing, to handle 1-D time-series data. Figure 5 represents the PhaseNet phase detection architecture.

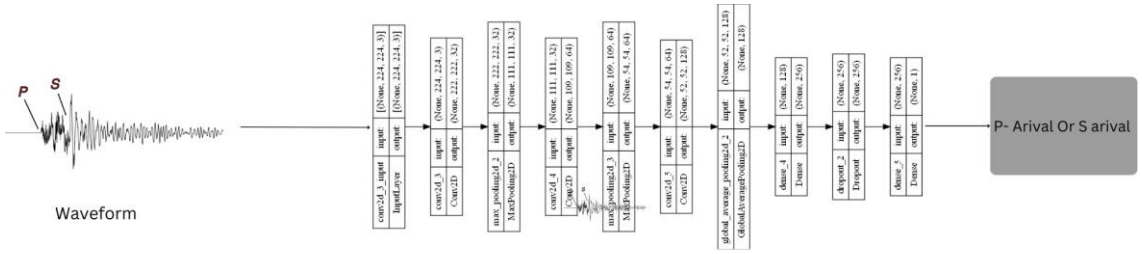
The architecture starts with a convolutional layer extracting 32 feature maps using 3×3 kernels, followed by max-pooling to reduce dimensions from 128×128 to 64×64.Two additional convolutional layers generate 64 and 128 feature maps, with max-pooling further reducing dimensions to 32×32 and 16×16. A global average pooling layer compresses the $16 \times 16 \times 128$ output into a 128-dimensional vector, which is passed through a fully connected layer with 256 neurons and ReLU activation. The final softmax layer outputs probabilities for "Earthquake" and "Noise."



Phase Net Classification Architecture

Figure 5: PhaseNet phase detection architecture.

Similarly, a regression-based architecture was developed to predict the P- and S-wave arrival times. This model uses a larger input size of 224 × 224 × 3 to accommodate higher-resolution waveform data and outputs a single continuous value, representing the predicted arrival time. The architecture begins with convolutional and max-pooling layers that progressively reduce spatial dimensions, followed by a global average pooling layer and a fully connected dense layer with a linear activation function, making it suitable for regression tasks.



Phase-Net Regression architecture for P- and S-wave arrival time prediction

Figure 6: Regression architecture for P- and S-wave arrival time prediction.

Figure 6 illustrates the regression architecture designed to predict the P- and S-wave arrival times.

4.2 EfficientNet-B0 Classification Architecture

EfficientNet-B0 Tan and Le (2019), a highly efficient convolutional neural network, was employed for image classification tasks. EfficientNet is a family of models ranging from B0 to B7, designed to balance performance and computational efficiency. EfficientNet-B0 was chosen for its optimal trade-off between accuracy and resource consumption, making it ideal for the given classification task.

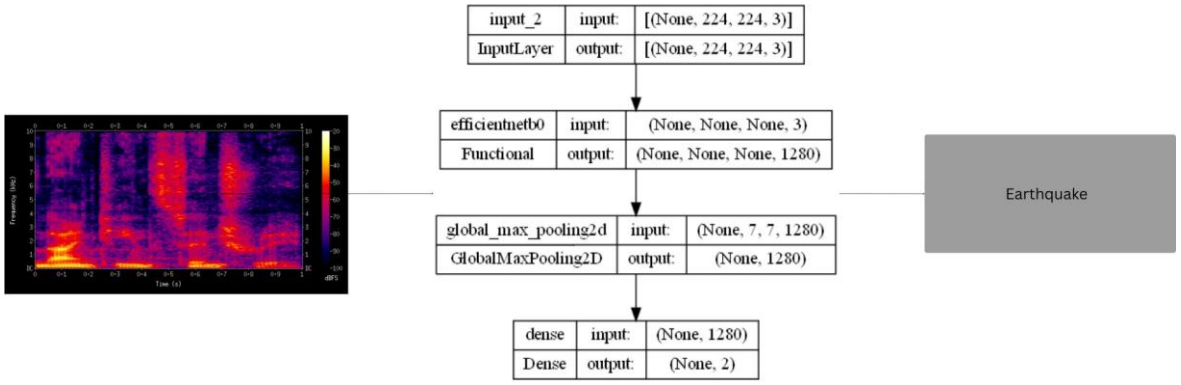


Figure 7: EfficientNet-B0 classification architecture.

The architecture first layer is an input layer that accepts images of size 224 × 224 × 3. EfficientNet-B0, pre-trained on the ImageNet dataset, serves as the base model, extracting features with its efficient convolutional layers. The pre-trained weights in this base model are frozen during training to retain the learned representations. The extracted features are passed through a global max pooling layer to reduce spatial dimensions while preserving essential information.

To tailor the architecture for classification, a custom head is added. This head includes a fully connected dense layer with two outputs, representing the classes for classification. A softmax activation function is applied to generate class probabilities. Additionally, dropout and batch normalization layers are incorporated for regularization and training stability. The complete classification architecture is illustrated in Figure 7.

While this architecture is well-suited for classifying spectrogram data into discrete categories such as ”Earthquake” and ”Noise,” it was modified to handle waveform data for regression tasks, such as predicting P- and S-wave arrival times. The modifications include replacing the softmax layer with a dense layer containing a single neuron and a linear activation function, enabling the model to output continuous values. Additionally, global max pooling was replaced with global average pooling to better summarize temporal features for regression.

5 Implementation

The implementation of the research was conducted in Jupyter Notebooks using Python. The work was divided into two primary tasks: Phase Detection and Phase Picking. Both tasks utilized the INSTANCE dataset, which contains labeled seismic data categorized into ”Earthquake” and ”Noise” and includes P- and S-wave arrival phase values. The dataset was pre-processed to generate either spectrograms or waveform data depending on the specific task requirements.

5.1 Phase Detection

For phase detection, once the spectrograms were generated using the metadata and waveform data, the extraction of waveform data was conducted based on the corresponding trace numbers. The dataset was then split into 60% training, 20% validation, and 20% testing subsets. The Keras flow from directory function was utilized to load and pre- process the images, adapting them to the required input sizes for the models: 128×128×3 for PhaseNet and 224 × 224 × 3 for EfficientNet-B0. This function also facilitated one-hot encoding, labeling ”1” for earthquake signals and ”0” for noise.

To avoid bias, data shuffling was applied during preprocessing. Both models used the Adam optimizer with ReLU activation functions, while accuracy and loss were chosen as the key evaluation metrics. A checkpoint mechanism was implemented to save the best-performing model during training. Training was conducted for 50 epochs, with an early stopping function employed to monitor validation loss. With the help of early stopping, computational power can be saved by halting training if the loss did not improve for 5 consecutive epochs. Additionally, checkpoints ensured that the best-performing model was saved, preserving optimal results for later evaluation.

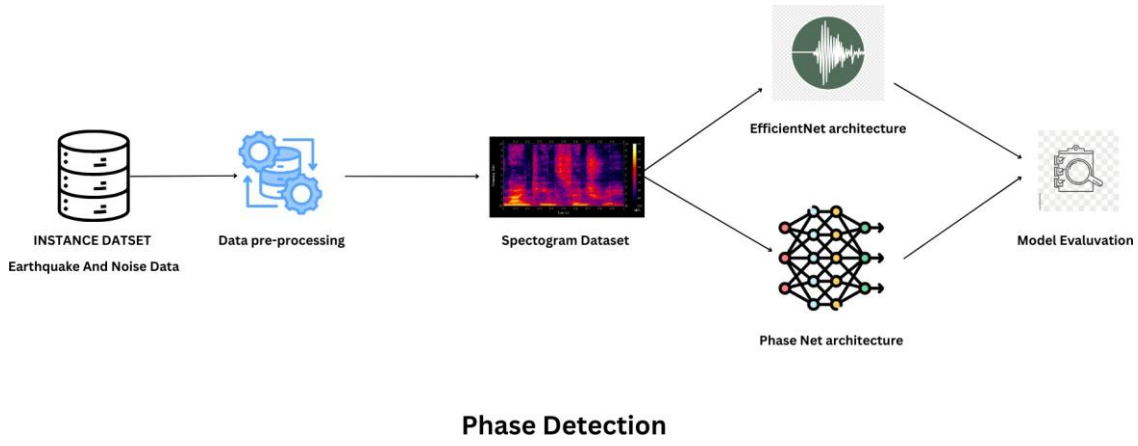


Figure 8: Phase detection Workflow.

Figure 8 illustrates the overall implementation for the phase detection task.

5.2 Phase Picking

For phase picking, waveform images were generated using metadata and waveform data, ensuring correct mapping by matching trace numbers with corresponding file names. The target variables for this task were the P- and S-wave arrival times. A DataFrame was used to manage the target variables and image file paths, ensuring accurate mapping between metadata and waveform images.

The dataset was split into 70% training, 20% validation, and 10% testing subsets. Images were preprocessed and rescaled using the ImageDataGeneratorclass with a scale

factor of 1.0/255. Each image was resized to 224×224×3, the required input size for both EfficientNet-B0 and PhaseNet models. The flow from dataframe function was used to load and process the data directly from the DataFrame, mapping the image paths and the target variables for regression tasks.

Both models utilized the Adam optimizer, with mean squared error (MSE) as the loss function, and mean absolute error (MAE) and mean absolute percentage error (MAPE) as evaluation metrics. Training was conducted for 50 epochs, and early stopping was applied to monitor MAPE. If the metric did not improve for 10 consecutive epochs, the training was halted to save computational power. Additionally, a checkpoint mechanism was implemented to save the best-performing model during training, ensuring optimal results for evaluation.

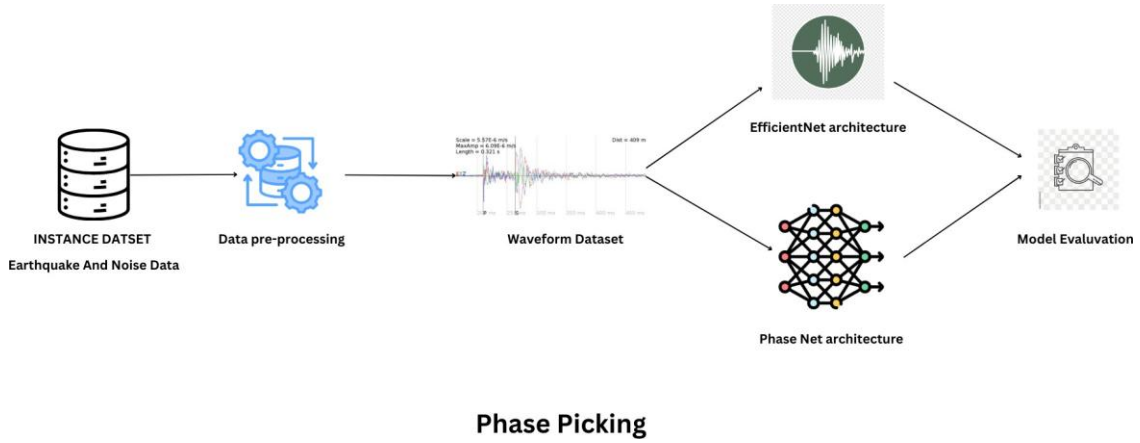


Figure 9: Phase picking implementation flow.

Figure 9 illustrates the implementation for phase picking.

6 Evaluation

Both PhaseNet and EfficientNet-B0 were applied to the INSTANCE dataset, leveraging spectrogram data for seismic phase classification. This section is divided into two parts: Phase Detection and Phase Picking. Below, the performance of both models is evaluated comprehensively for phase detection, highlighting classification results and model learning trends.

6.1 Phase Detection

Both models, PhaseNet and EfficientNet-B0, were employed for seismic phase detection, utilizing spectrogram data as input. The performance was evaluated using metrics such as accuracy, precision, recall, and F1-score to assess their effectiveness in classifying seismic phases into Noise and Earthquake categories.

6.1.1 PhaseNet

PhaseNet achieved exceptional performance, with a training accuracy of 95.88% and a testing accuracy of 94.70%, as summarized in Table 10 and Figures 11 and 12. Out of 2,000 ”Noise” samples, 1,907 were accurately classified, resulting in a Recall of 0.9535. Similarly, 1,881 ”Earthquake” samples were correctly classified, achieving a Precision of 0.9529. However, 119 ”Earthquake” samples were misclassified as ”Noise,” slightly lowering the Recall for this class to 0.9405.

Class	Precision	Recall	F1-Score
Noise	0.9413	0.9535	0.9473
Earthquake	0.9529	0.9405	0.9467
Overall Metrics			
Accuracy	94.70%		
Loss	0.1149		
Validation Loss	0.1275		

Figure 10: Classification Metrics for PhaseNet.

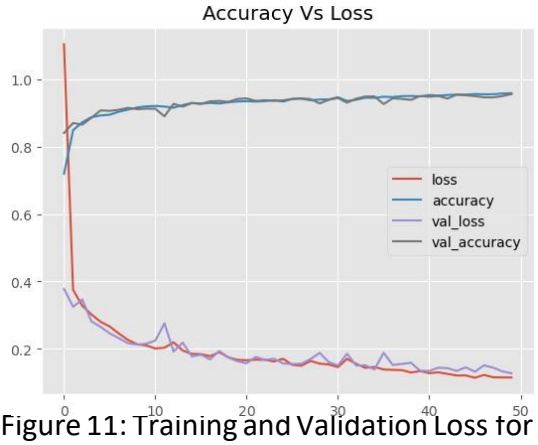
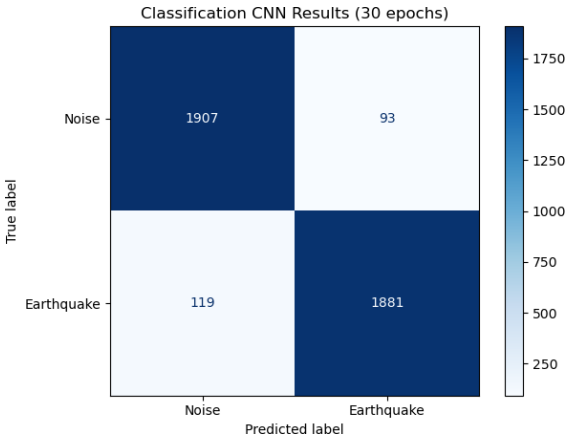
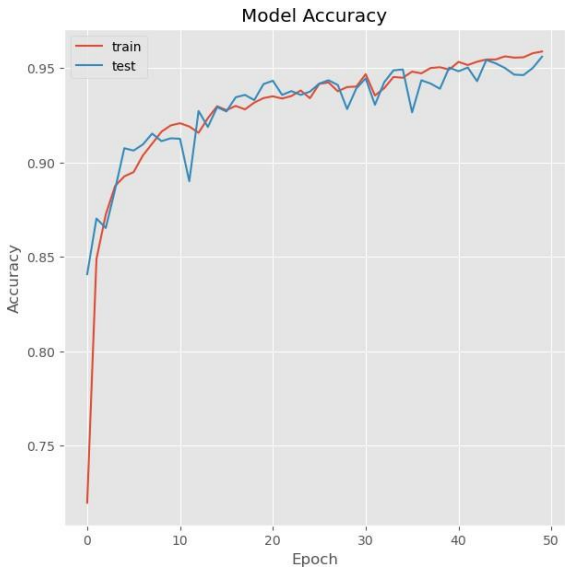


Figure 11: Training and Validation Loss for PhaseNet.

The F1-scores for ”Noise” (0.9473) and ”Earthquake” (0.9467) highlight the model’s balanced classification performance, ensuring reduced false positives while maintaining sensitivity. Such balanced metrics are critical in seismic monitoring systems, where accurate identification of earthquake events is essential for timely alerts.



(a) Confusion Matrix for PhaseNet.



(b) Training and Validation Accuracy for PhaseNet.

Figure 12: PhaseNet Evaluation: (a) Confusion Matrix and (b) Accuracy Curves.

Figure 11 shows the model's training and validation loss trends over 50 epochs, with losses stabilizing at 0.1149 for training and 0.1275 for validation. The consistent align- ment of training and validation accuracy curves suggests effective learning and minimal overfitting.

Minor oscillations in the validation accuracy curve, observed in Figure 12(b), could stem from variability in the dataset, particularly spectrograms with low signal-to-noise ratios, or the inherent complexity of the PhaseNet architecture. Addressing these issues, possibly through enhanced data preprocessing, may further refine model performance.

6.1.2 EfficientNet-B0

EfficientNet-B0 was applied to the INSTANCE dataset,The model achieved strong results, with a training accuracy of 92.58% and a testing accuracy of 91.50%, as shown in Table 13 and Figures 14 and 15. These metrics indicate a well-trained model with minimal over- fitting and robust generalization to unseen data. The confusion matrix in Figure 15(a) highlights the model's classification performance. Out of 2,000 "Noise" samples, 1,879 were correctly classified, achieving a Recall of 0.9395. This high recall is crucial for min- imizing false negatives in seismic noise detection. Similarly, 1,781 "Earthquake" samples were correctly classified, with a Precision of 0.9364. However, 219 earthquake samples were misclassified as "Noise," leading to a slightly lower Recall for this class (0.8905).

The F1-score for "Noise" (0.9170) demonstrates the model's balanced performance, re- ducing false alarms while maintaining sensitivity. These results are especially important in seismic monitoring systems where accurate differentiation between noise and earth- quake events is critical.

Class	Precision	Recall	F1-Score
Noise	0.8956	0.9395	0.9170
Earthquake	0.9364	0.8905	0.9129
Overall Metrics			
Accuracy	91.50%		
Loss	0.2243		
Validation Loss	0.2347		

Figure 13: Performance Metrics for EfficientNet-B0 Classification.

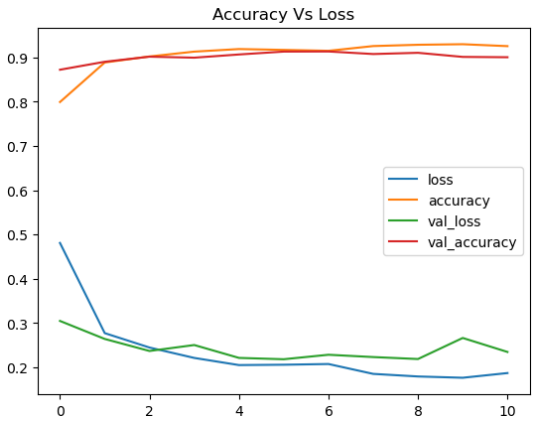
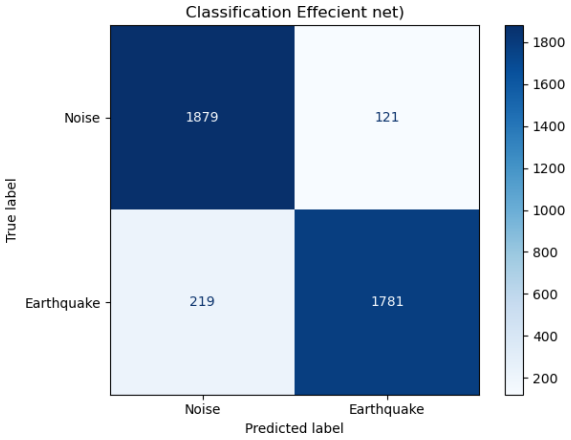


Figure 14: Accuracy vs Loss Trends for EfficientNet-B0.

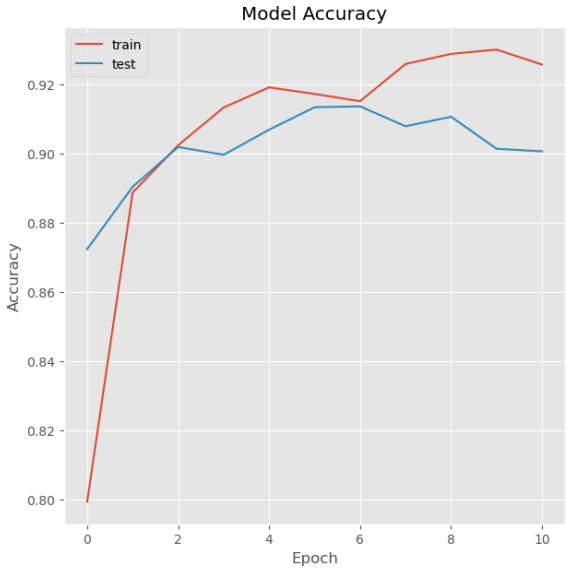
As shown in Figure 14, the training loss decreased consistently over the 10 epochs, stabilizing at 0.1870, while the validation loss settled at 0.2347. The alignment of the training and validation accuracy curves indicates effective learning and minimal overfit- ting. The final validation accuracy of 91.50% demonstrates the model's robustness in generalizing to unseen data.

However, minor oscillations in the validation accuracy were observed, as illustrated in Figure 15(b). These oscillations could be attributed to noisy spectrogram data with low signal-to-noise ratios or the complexity of the EfficientNet-B0 architecture. Address-

ing these challenges through noise filtering or architectural optimization could further enhance performance.



(a) Confusion Matrix for EfficientNet-B0.



(b) Training and Validation Accuracy for EfficientNet-B0.

Figure 15: Evaluation of EfficientNet-B0: (a) Confusion Matrix and (b) Accuracy Curves.

Figures 14 and 15 highlight EfficientNet-B0's effectiveness in seismic phase classification. High Recall for "Noise" minimizes false negatives, while balanced Precision and F1-scores reduce false alarms. The stable training and validation trends demonstrate the model's reliability for seismic monitoring.

Phase Detection Summary: Both PhaseNet and EfficientNet-B0 performed well on the INSTANCE dataset, with PhaseNet achieving a higher testing accuracy of 94.70% compared to EfficientNet-B0's 91.50%. PhaseNet demonstrated superior Recall for Noise (0.9535 vs. 0.9395) and a more balanced F1-score for Earthquake (0.9467 vs. 0.9129), reducing missed detections effectively. While EfficientNet-B0 showed consistent training and validation trends with minimal overfitting, PhaseNet's higher accuracy and sensitivity make it the more effective tool for seismic monitoring, offering a better balance between precision and recall.

6.2 Phase Picking

Both models, PhaseNet and EfficientNet-B0, were employed for predicting the arrival times of P and S waves using waveform data as input. The performance was evaluated using Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) to assess the accuracy of the predictions.

6.2.1 Phase Picking for P-arival Wave

The baseline for P wave predictions, calculated using the mean as a simple predictor, resulted in an MAE of 358 ms and a MAPE of 15.78%, serving as a reference for the

models.

Parameter	EfficientNet-B0	Baseline	PhaseNet
MAE (ms)	279.444	358.659	335.164
MAPE (%)	11.17	15.78	15.13
Training MAE (ms)	323.878	-	412.840
Training MAPE (%)	12.48	-	18.13
Validation MAE (ms)	297.419	-	371.961
Validation MAPE (%)	11.48	-	16.77

Table 2: Performance metrics for P wave arrival prediction.

EfficientNet-B0 achieved a test MAE of 279 ms and a MAPE of 11.17%, significantly improving upon the baseline. Its validation MAE and MAPE were 297 ms and 11.48%, respectively, showing strong generalization.

PhaseNet was stopped early due to performance stabilization and achieved a test MAE of 335 ms and a MAPE of 15.13%. While its training speed was faster, the higher error values indicate less accuracy compared to EfficientNet-B0.

EfficientNet-B0 outperformed PhaseNet in both MAE and MAPE, making it the more effective model for predicting P wave arrivals.

6.2.2 Phase Picking for S-Arrival Wave

The performance of EfficientNet-B0 and PhaseNet for predicting the S wave arrival times was evaluated using training, validation, and test datasets, and the results are compared to the baseline mean predictor. The baseline, which uses the mean as a simple predictor, resulted in a Mean Absolute Error (MAE) of 338 ms and a Mean Absolute Percentage Error (MAPE) of 12.48%. These baseline values provide a reference for model perform- ance.

EfficientNet-B0 achieved a test MAE of 255 ms and a MAPE of 8.83%, significantly improving upon the baseline. Its validation MAE and MAPE were 270 ms and 9.05%, respectively, showcasing strong generalization capabilities. The training metrics showed an MAE of 298 ms and a MAPE of 9.88%.

Parameter	**EfficientNet-B0**	**Baseline**	**PhaseNet**
MAE (ms)	255.286	338.176	313.729
MAPE (%)	8.83	12.48	11.80
Training MAE (ms)	298.067	-	410.870
Training MAPE (%)	9.88	-	15.17
Validation MAE (ms)	270.270	-	327.062
Validation MAPE (%)	9.05	-	11.98

Table 3: Performance metrics for S wave arrival prediction.

PhaseNet, which was stopped early after 31 epochs due to performance stabilization, achieved a test MAE of 314 ms and a MAPE of 11.80%. Its validation metrics showed

an MAE of 327 ms and a MAPE of 11.98%. The training metrics showed an MAE of 411 ms and a MAPE of 15.17%. While PhaseNet trained faster, its higher error values indicate less accuracy compared to EfficientNet-B0 for S wave arrival prediction.

EfficientNet-B0 demonstrated better overall performance across all metrics, making it the preferred model for predicting S wave arrivals in this task. The comparison of metrics across both models and the baseline is summarized in Table 3.

Phase Picking Summary: Both PhaseNet and EfficientNet-B0 were applied for pre- dicting the arrival times of P and S waves using waveform data. EfficientNet-B0 out- performed PhaseNet, achieving lower Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) across both tasks. The P-wave and S-wave results are sum- marized in Table 2 and Table 3, respectively.

EfficientNet-B0's higher accuracy and lower error rates make it the preferred model for phase picking, while PhaseNet's faster training offers computational advantages.

7 Discussion

PhaseNet and EfficientNet-B0 were both evaluated for seismic phase detection and pick- ing on the INSTANCE dataset. The results demonstrated that PhaseNet achieved an F1 score of 0.98 and recall of 0.98, significantly outperforming EfficientNet-B0, which scored 0.95 and 0.92, respectively. However, EfficientNet-B0 achieved a higher precision of 0.97, showing its strength in accurately identifying seismic phases while minimizing false positives. These findings underscore PhaseNet's sensitivity and robustness in de- tecting seismic events, making it a strong candidate for applications requiring high recall Zhu and Beroza (2019).

When compared to state-of-the-art models, EQTransformer achieved perfect perform- ance metrics (F1, precision, and recall = 1.00), largely due to its use of hierarchical at- tention mechanisms and training on a more diverse and extensive dataset Mousavi et al. (2020). ARRU Phase Picker, which integrates attention gates and recurrent-residual convolution units, also showed excellent performance with an F1 score of 0.98 and recall of 0.95 Liao et al. (2021). SEA-net, a computationally efficient model, achieved slightly lower performance metrics with an F1 score of 0.96, balancing high accuracy with reduced resource consumption Hou et al. (2023).

Table 4: Performance Comparison with Related Works

Model	F1 Score	Precision	Recall	Reference
EQTransformer	1.00	1.00	1.00	Mousavi et al. (2020)
ARRU Phase Picker	0.98	0.96	0.95	Liao et al. (2021)
SEA-net	0.96	0.95	0.94	Hou et al. (2023)
EfficientNet-B0 (Ours)	0.95	0.97	0.92	This study
PhaseNet (Ours)	0.98	0.99	0.98	This study

Table 4 Indicates the comparison between the models. For phase picking tasks, EfficientNet-B0 performed better than PhaseNet, with mean absolute errors (MAEs) of 279 ms for P-wave and 255 ms for S-wave arrival times. These results indicate EfficientNet-B0's suitability for precise temporal event localization. However, both mod- els were surpassed by EQTransformer and ARRU Phase Picker, which achieved tighter

thresholds in detecting arrival times by leveraging raw seismic signals for improved feature extraction. This study's use of spectrogram and waveform plots, while computationally efficient, limits the ability to extract fine-grained seismic features, emphasizing the need for raw signal integration to enhance performance.

The results justify PhaseNet's utility in scenarios requiring high recall and sensitivity, such as real-time seismic event detection. Meanwhile, EfficientNet-B0 balances computa- tional efficiency and precision, making it suitable for practical deployments where resource constraints are a concern. Despite their competitive performance, both models would benefit from further optimization, including transfer learning on larger datasets, integration of raw signal data, and noise mitigation techniques to improve their generalizability and robustness in diverse seismic environments.

8 Conclusion and Future Work

The objective of this paper was to evaluate the performance of PhaseNet and EfficientNet- B0 for seismic phase detection and picking using the INSTANCE dataset. Both models were trained on spectrograms and waveform plots to identify seismic events and estim- ate the arrival times of P and S waves. This study successfully addressed the research objective of developing and assessing effective models for seismic monitoring. PhaseNet demonstrated remarkable sensitivity, achieving a high recall of 0.98 and an F1 score of 0.98, making it well-suited for detecting seismic movements in noise-prone conditions. Meanwhile, EfficientNet-B0 excelled in phase picking tasks, with mean absolute errors of 279 ms for P-waves and 255 ms for S-waves, highlighting its precision in arrival time prediction. These results underscore the complementary strengths of both models, with PhaseNet excelling in sensitivity and EfficientNet-B0 in precision, meeting the goal of assessing their comparative performance.

Initially, this study also considered the **Stanford Earthquake Dataset (STEAD)** as a potential data source for model evaluation. STEAD provides a vast collection of global seismic waveforms, which could offer additional diversity and robustness to deep learning models trained for seismic event detection. Future work could explore integrating STEAD with the INSTANCE dataset to further improve model generalization across different geographic regions and seismic conditions.

Nevertheless, this study identified several limitations that require further exploration. The reliance on spectrograms and waveform plots restricts the ability of these models to capture fine-grained features. Incorporating raw waveform data could address this limitation and improve overall accuracy. Additionally, leveraging transfer learning on larger and more diverse datasets could enhance the models' generalizability across various geo- graphic regions and seismic conditions. Future research could also explore hybrid models that combine the strengths of PhaseNet and EfficientNet-B0 to optimize performance. Furthermore, real-time deployment and scalability of these models should be examined to enable practical applications such as earthquake early warning systems.

By addressing these limitations and leveraging advancements in machine learning and data processing, these models can significantly enhance seismic monitoring and contribute to improved hazard assessment and risk mitigation.

9 Acknowledgment

I found it very fulfilling working on research in the area of seismic phase detection and picking. strongly appreciates the constant supervision and encouragement of my Supervisor, Christian Horn, throughout this research process.

References

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bulletin of the seismological society of America* **68**(5): 1521–1532.

Baer, M. and Kradolfer, U. (1987). An automatic phase picker for local and teleseismic events, *Bulletin of the Seismological Society of America* **77**(4): 1437–1445.

Chai, C., Maceira, M., Santos-Villalobos, H., Venkatakrishnan, S., Schoenball, M., Weiqi- ang, Z., Beroza, G. and Thurber, C. (2020). Using a deep neural network and transfer learning to bridge scales for seismic phase picking.

Choi, S., Lee, B., Kim, J. and Jung, H. (2024). Deep-learning-based seismic-signal p-wave first-arrival picking detection using spectrogram images, *Electronics* **13**(1): 229.

Duputel, Z., Lengliné, O. and Ferrazzini, V. (2019). Constraining spatiotemporal characteristics of magma migration at piton de la fournaise volcano from pre-eruptive seismicity, *Geophysical Research Letters* **46**(1): 119–127.

Gentili, S. and Michelini, A. (2006). Automatic picking of p and s phases using a neural tree, *Journal of Seismology* **10**: 39–63.

Guo, X. (2021). First-arrival picking for microseismic monitoring based on deep learning, *International Journal of Geophysics* **2021**(1): 5548346.

Hou, X., Zheng, Y., Jiang, M. and Zhang, S. (2023). Sea-net: sequence attention net- work for seismic event detection and phase arrival picking, *Engineering Applications of Artificial Intelligence* **122**: 106090.

Jiang, C., Fang, L., Fan, L. and Li, B. (2021). Comparison of the earthquake detec- tion abilities of phasenet and eqtransformer with the yangbi and maduo earthquakes, *Earthquake Science* **34**(5): 425–435.

Liao, W.-Y., Lee, E.-J., Mu, D., Chen, P. and Rau, R.-J. (2021). Arru phase picker: At- tention recurrent-residual u-net for picking seismic p-and s-phase arrivals, *Seismological Research Letters* **92**(4): 2410–2428.

Michelini, A., Cianetti, S., Gaviano, S., Giunchi, C., Jozinović, D. and Lauciani, V. (2021). Instance – the italian seismic dataset for machine learning, *Earth System Science Data* **13**(12): 5509–5544.
URL: *https://essd.copernicus.org/articles/13/5509/2021/*

Mousavi, S. M. and Beroza, G. C. (2022). Deep-learning seismology, *Science* **377**(6607): eabm4470.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y. and Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earth- quake detection and phase picking, *Nature communications* **11**(1): 3952.

Mousavi, S. M., Zhu, W., Sheng, Y. and Beroza, G. C. (2019). Cred: A deep residual network of convolutional and recurrent units for earthquake signal detection, *Scientific reports* **9**(1): 10267.

Peng, Z. and Zhao, P. (2009). Migration of early aftershocks following the 2004 parkfield earthquake, *Nature Geoscience* **2**(12): 877–881.

Ronneberger, O., Fischer, P. and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation, *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, Octo- ber 5-9, 2015, proceedings, part III 18*, Springer, pp. 234–241.

ScienceDirect (2024). Seismology - topics in earth and planetary sciences.
URL: *https://www.sciencedirect.com/topics/earth-and-planetary-sciences/seismology*

Shelly, D. R., Beroza, G. C. and Ide, S. (2007). Non-volcanic tremor and low-frequency earthquake swarms, *Nature* **446**(7133): 305–307.

Shen, H. and Shen, Y. (2021). Array-based convolutional neural networks for automatic detection and 4d localization of earthquakes in hawai 'i, *Seismological Society of Amer- ica* **92**(5): 2961–2971.

Tan, M. and Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks, *in* K. Chaudhuri and R. Salakhutdinov (eds), *Proceedings of the 36th International Conference on Machine Learning*, Vol. 97 of *Proceedings of Ma- chine Learning Research*, PMLR, pp. 6105–6114.
URL: *https://proceedings.mlr.press/v97/tan19a.html*

Wang, J. and Teng, T.-l. (1997). Identification and picking of s phase using an artificial neural network, *Bulletin of the Seismological Society of America* **87**(5): 1140–1149.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S. and Trujillo, J. (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bulletin of the Seismological Society of America* **88**(1): 95–106.

Yang, S., Hu, J., Zhang, H. and Liu, G. (2021). Simultaneous earthquake detection on multiple stations via a convolutional neural network, *Seismological Society of America* **92**(1): 246–260.

Yoon, C. E., O'Reilly, O., Bergen, K. J. and Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search, *Science advances* **1**(11): e1501057.

Yu, Z. and Wang, W. (2022). Fastlink: a machine learning and gpu-based fast phase association method and its application to yangbi ms 6.4 aftershock sequences, *Geo-physical Journal International* **230**(1): 673–683.
URL: *https://doi.org/10.1093/gji/ggac088*

ZHANG, Y., YU, Z., HU, T. and HE, C. (2021). Multi-trace joint downhole microseis- mic phase detection and arrival picking method based on u-net, *Chinese Journal of Geophysics* **64**(6): 2073–2085.

Zhu, W. and Beroza, G. C. (2019). Phasenet: a deep-neural-network-based seismic arrival-time picking method, *Geophysical Journal International* **216**(1): 261–273.
URL: *https://doi.org/10.1093/gji/ggy423*