Seismic Phase Detection & Picking using EfficientNet

MSc Research Project
Data Analytics

Arunprasath Ramakrishnan Arularasan
Student ID: 20207417

School of Computing
National College of Ireland

Supervisor: Dr. Vladimir Milosavljevic

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Arunprasath Ramakrishnan Arularasan		
Student ID:	20207417		
Programme:	Data Analytics	**Year:**	2021-2022
Module:	MSc Research Project		
Supervisor:	Dr. Vladimir Milosavljevic		
Submission Due Date:	August 15, 2022		
Project Title:	Seismic Phase Detection & Picking using EfficientNet		
Word Count:	8367 **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: Arunprasath Ramakrishnan Arularasan

Date: August 15, 2022

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):	

Seismic Phase Detection & Picking using EfficientNet

Arunprasath Ramakrishnan Arularasan
20207417

Abstract

The monitoring of seismic waves for the detection of earthquakes and the picking of the arrival of P and S waves has been a challenging task in the field of observational seismology. While the use of deep learning techniques has led to improved performance, the models tend to suffer from poor generalizability and poor picking performance with the S waves. This research proposes the use of EfficientNet architecture on the spectrogram and waveform plots of the signals for phase detection and phase picking. The performance of the model was evaluated using the Italian Seismic dataset. The model was also trained using the Stanford Earthquake dataset for comparing it with the existing models. The EfficientNet baseline model was outperformed by a simple CNN architecture on the Italian Seismic dataset. When trained on the Stanford dataset, the EfficientNet model had an F1 score of 0.95 which is slightly lesser than the existing models. In phase regression, the EfficientNet architecture had poor picking performance compared to the existing models. While the models trained on the Stanford dataset had a higher accuracy compared to the Italian dataset, both the models suffered with poor generalizability when tested with waveforms from different regions.

1 Introduction

1.1 Background and Motivation

For years, the signals from the seismometer have been monitored by a group of analysts for the detection of earthquakes and the arrival of P and S waves. The increase in the number of sensors on the seismic station led to huge volumes of data and hand-picking of signals became difficult and automated methods were considered (Mousavi, et al., 2020). The detection of earthquake signals is called phase detection and the prediction of arrival times of the P and the S waves is called phase picking. Phase detection is important as it helps in the early prediction of earthquakes and phase picking helps in estimating the magnitude, hypocenter location, and spectral analysis. The earliest automated detection method came as early as 1978 with the use of the Short-time Average (STA)/ Long-time Average (LTA) method for phase detection (Allen, 1978) and later the Akaike information criterion (AIC) was used for phase picking (Leonard & Kennett, 1999). While these methods are still widely used in a few seismic stations, the traditional models tend to suffer when the signal-to-noise ratio (SNR) is very low and also had trouble picking the S waves (Saad & Chen, 2022).

The development of machine learning found its usage in the field of observational seismology for phase detection and phase picking. The use of machine learning models like K-Means, fuzzy logic and SVM improved the performance of the detection model compared to the traditional methods. The models still suffered in picking the signals at low SNR and the unsupervised methods took a long time while loading the data (Saad & Chen, 2022). The use of deep learning algorithms drastically improved the performance of classification of seismic signals and phase picking. While the development of state-of-the-art algorithms like PhaseNet (Zhu & Beroza, 2019) and EQTransformer (Mousavi, et al., 2020) improved the accuracy of phase detection and the picking of the arrival of P waves they still performed considerably worse while picking the arrival of S waves. Most successful approaches for phase detection and phase picking have used ground motion signals with neural networks. Some approaches have also plotted the signals as waveforms and spectrograms and used CNN architectures for identifying the earthquakes and phase picking. The CNN architectures are usually built at a fixed cost budget and then scaled for better accuracy based on the available resource budget. To overcome this, EfficientNet was developed that scaled the image in a more principled way along its depth, width, and resolution (Tan & Le, 2019). EfficientNet has gained immense popularity in the field of image segmentation tasks like the autonomous vehicle trajectory detection and classification of medical diseases. The motivation of this research was to use the EfficientNet architecture on the spectrogram and waveform plots of the seismic signals to check if it could address the drawbacks of the existing models.

The recent advancements in observational seismology were due to the availability of curated datasets like the Stanford Earthquake Dataset (STEAD) and Southern California Seismic Network Dataset (SCSN). One of the drawbacks of the models trained on these datasets was the poor generalizability while predicting waveforms from outside the regions they were trained in (Jiang, et al., 2021). The INSTANCE dataset is the Italian seismic dataset that has been curated from the EIDA node of the Italian seismic bulletin (INGV) for machine learning (Michelini, et al., 2021). This dataset has been used in this research for analyzing the performance of the EfficientNet architecture for phase classification and picking. The research also assesses the overall generalizability of the model trained on this dataset compared to the STEAD dataset.

1.2 Research Question

1. How EfficientNet Classification model perform while detecting the seismic waves?
2. How EfficientNet Regression model perform while picking the P and S waves?

1.3 Research Objectives

- To evaluate the performance of the EfficientNet model for Phase Classification using the INSTANCE dataset
- To evaluate the performance of the Efficient Regression model for Phase picking using the INSTANCE dataset
- To compare the performance of the EfficientNet models with the other existing models on the STEAD dataset
- To assess the generalizability of the models trained with the INSTANCE dataset.

2 Related Work

2.1 Phase detection and phase picking

The need to process the huge data gathered from the seismic stations led to the use of the Short-Time Average and Long-Time Average (STA/LTA) method for classification of earthquake signals (Allen, 1978). The research compared the amplitude ratio of the wave in a short-time window and long-time window. The model predicted an earthquake when the ratio exceeded the defined threshold. In the (Leonard & Kennett, 1999) research, the authors have used the Akaike Information Criterion (AIC) for detecting the phase arrival of the earthquake signals. This method used autoregressive modelling on the earthquake and the noise signals to calculate the AIC. Earthquake signals were classified based on the minimum AIC value assigned for the arrival time. While these models are still used in many seismic stations, the model's performance dropped when the SNR was low and while picking the S waves.

The authors in (Chen, 2018) used a unsupervised learning method for clustering the microseismic data time samples into a waveform and non-waveform points. The major advantage of this model was its flexibility as it didn't need huge amounts of data. (Li, et al., 2018), proposed the use of a "local similarity" detection function for identifying the seismic signals. This function compared the traces with the neighbouring stations and eliminated the high-frequency spikes that were contaminated in the STA/LTA method. This approach led to identifying earthquakes that even had weak SNR in the frequency range < 1Hz. However, this approach didn't show improved performance in the frequency range 5-10Hz. While the state-of-the-models that were developed during this period using machine learning performed better than the traditional algorithms, their performance in low SNR was still not significant. The use of unsupervised models like clustering also resulted in slower prediction time as the model had to check for similarity every time before making a prediction (Saad & Chen, 2022).

The use of deep learning models in the field of seismology helped in the development of robust techniques. PhaseNet was one of the earliest deep learning architectures that had immense success in earthquake detection (Zhu & Beroza, 2019). It was based on the image semantic segmentation network Unet, which can outline the target object from the image precisely. PhaseNet was trained using 600,000 waveforms and it had 5 layers and 268,000 parameters recorded by the Northern California Earthquake Data Centre. It created gaussian spots on the target phase with a 0.1s of standard deviation. PhaseNet model while picking P waves had precision and recall of 96% and 93% while picking S waves. Despite PhaseNet being widely used in many regions like the United States, Ridgecrest, and Central Apennines, Italy, they have some drawbacks. Since this model creates small gaussian spots, the model had difficulty making positive predictions. The model also relied on end-to-end segmentation networks and hence was computationally heavy. The drawbacks of the PhaseNet model were discussed in (Pardo, et al., 2019) and the authors proposed a three-step approach using a convolutional neural network architecture. A rough segmentation mask was created around the targeted phase followed by a distance map around the peak mask value. Finally, using Hough voting both the networks were combined. This model outperformed other methods like the PhaseNet with a higher F1 score while predicting the picks that were close to 0.1s.

While most deep learning models used the waveform signals directly for monitoring the seismic activities, the authors in (Curilem, et al., 2018) proposed the usage of spectrograms for monitoring the volcanic seismic signals. This research converted the seismic signals into a spectrogram and used a CNN architecture to make predictions. The results show that representing the signals in 2D increased the classification performance. The authors in (Calderon, et al., 2020) used grayscale spectrogram images for the classification of seismic signals using a similar technique. This research showed improved performance while classifying seismic signals using a modified CNN classifier with three convoluted layers.

Earthquake Transformer (EQT Transformer) was another state-of-the-art model that was developed for phase detection and phase picking (Mousavi, et al., 2020). In this research, the authors combined both the CNN and RNN layers for encoding the temporal relationships of the seismic time-domain signals into high-level representations. The model then applied an attention mechanism using a transformer for retaining the local and the global features of the seismic signals. Based on the resultant information, three separate decoders were used for calculating the probabilities of p-phase picking, s-phase picking, and earthquake detection separately. The EQT Transformer was trained using the 850,000 waveforms in the STEAD dataset and had 378,000 parameters. The EQT Transformer had a precision of 99% while picking p-waves and s-waves and had a recall of 99% and 96% for p and s waves respectively. EQT Transformer gained huge popularity due to its ability to simultaneously perform phase detection and phase picking. The authors in research (Van Der Laat, et al., 2021), incorporated the EQT Transformer into a deep learning pipeline OVSICORI Kabre Seismological Pipeline (OKSP) for monitoring the earthquakes in Costa Rica. The model was given continuous input from 10 nearby seismic stations consisting of 24hr three-component waveforms. These inputs were then converted into the required input format for the EQT Transformer which was then used for phase detection and phase picking. The outliers were removed from the output of the EQT Transformer to build the final catalog for earthquake monitoring. This helped in identifying 1100 more earthquakes than the previous methods.

The success and wide usage of the PhaseNet and the EQT Transformer led to comparative research on these models. The authors in (Jiang, et al., 2021) research compared both these models to the seismic signals from the Qinghai Maduo and Yunnan Yangzi earthquakes. Both the model’s picking performance was reduced considerably as the waveform signals from different regions were used. PhaseNet had a lower error while picking arrival time and a higher recall than the EQT Transformer. The missed detection rate was also less in PhaseNet, but the model predicted more earthquakes than the actual count which could be due to the absence of noise in the training dataset. Although the EQT Transformer had fewer false positives and a higher precision, they missed more earthquakes than PhaseNet. Though EQT Transformer was trained using waveforms in the STEAD dataset that were taken from around the world, they had poor generalizability than PhaseNet. The author states that a higher volume training dataset may not increase the generalizability and that future research should focus on the quality of the dataset. They also suggest that testing of the seismic model in the future should be done on waveforms from different regions.

Other approaches were also considered for seismic phase picking and classification. The authors of the (Jozinović, et al., 2021) paper proposed convolutional neural network architecture for the rapid classification of earthquakes. The model input was the raw multi-station waveform, to simulate the real-time data from the epicentre. It was trained with 40,443 noise waveforms and 33,855 earthquake waveforms and the model gave a satisfactory performance while detecting earthquakes. The model’s performance degraded while classifying earthquakes greater than 5 which could be cited to the lack of waveforms with a higher magnitude in the dataset. The author states that this model will not be useful in real time prediction, but it can be useful in estimating the ground motion signals. Another model that is being used in recent times is Crowd Quake (Huang, et al., 2020), which uses low-cost acceleration sensors for monitoring earthquakes. This model used a Convolutional-Recurrent Neural Network (CRNN) and had a precision of 99% with false alarms being less than 1%. CrowdQuake has been deployed in South Korea with 300 low-cost acceleration sensors and till now the model has been able to detect 5 earthquakes near the area.

The authors in (Choubik, et al., 2021) used Fully convoluted Neural network classifiers (FCN) that were widely successful in fields like medical image analysis and image segmentation for detecting earthquakes. This classifier consists of 4 Convolutional layers, a batch normalization layer, followed by a pooling layer, and a SoftMax function to classify the seismic signals. The balanced STEAD dataset was used to train the classifier. Each sample was normalized separately, which had shown to improve the classifier's accuracy by up to 16 percent. Capsule neural networks that have been recently gaining popularity in replacing the CNN have been used in this research (He, et al., 2021) for picking the P-wave arrival. The authors of this research developed PickCapsNet consisting of two convolutional layers and a capsule network. The major advantage of the PickCapsNet was its ability to scale and the model had higher accuracy in different SNRs. Hence, PickCapsNet was more reliable and stable than the traditional models. Since this model used a lot of pooling layers like CNN, some important information was lost during training.

To overcome the drawbacks of PickCapsNet, the authors of the research (Saad & Chen, 2021) proposed the use of CapsNet model for predicting the arrival times. In this method, the earthquake signals containing both P and S waves were identified, and the P wave arrival time was extracted. The CapsNet architecture uses three convolutional layers for extracting the significant features along with a primary and digit capsule layer for producing combinations of the features and outputting the probability of the classification respectively. The Southern California Seismic Network (SCSN) Data was used for training this model which consists of 4.5 million seismograms. The model was tested with 222,395 seismic waveforms from the US, Europe, and Japan and it was able to detect 217,305 waveforms. In 97% of the seismic signals, the model predicted the arrival time with an error below 0.2s. Although CapsNet performed well when picking P waves, it performed poorly when picking S waves. This was because the model performed a binary classification considering the P and the S waves as same. In the research (Saad & Chen, 2022), the authors have developed CapsPhase to address this issue. The architecture of CapsPhase was like the CapsNet with the addition of another convolutional layer and a dynamic routing strategy in the digital capsule layer to output the probabilities for the earthquake, P-wave, and S-wave. This model was trained using the SCSN dataset and the performance was tested using the Japanese seismic dataset. This model outperformed the CapsNet and traditional methods like the STA/LTA and had a less mean absolute error while picking P-waves. The only limitation of the model was the picking accuracy of the S-waves which hadn't improved much.

2.2 EfficientNet

Convolutional Neural Network architectures are generally developed at an initial cost and then scaled up for better accuracy in the future. The scaling up of these architectures is done along the width, depth, and resolution for improving the accuracy of the model. The drawback of scaling up an architecture was the performance got saturated easily. In the research (Tan & Le, 2019), the authors proposed the use of EfficientNet which uses compound scaling to principally scale the dimensions along the depth, width, and resolution. The EfficientNet B0 is the baseline network architecture and the scaled-up architectures from B1 to B7 provide better accuracy. This architecture achieved a state-of-the-art accuracy of 91% on the CIFAR-100 dataset, 98% on the Flowers dataset, and three other transfer learning datasets. EfficientNet also used fewer parameters than the other architectures like DenseNet, ResNet, and Inception.

The authors in (Sgibnev, et al., 2020) research used the EfficientNet architecture for addressing the problem of the semantic image in an autonomous vehicle. The authors used the EfficientNet-B0 architecture along with other architectures like the ResNet, MobileNet, and ShuffleNet. The EfficientNet had an accuracy of 90% when used with the DeepLab decoder on the pre-trained cityscapes dataset. In the research (Tang, et al., 2021) the authors have also used EfficientNet for autonomous vehicle trajectory prediction. The authors used the vehicle data, nearby moving objects data, and traffic data for predicting the vehicle trajectory. The research compares the performance of the EfficientNet model along with VGG16, and ResNet34 and finds that the EfficientNet models had a minimal loss.

The superior image segmentation performance of the EfficientNet led to its usage in classification of medicinal diseases. The authors in the (Atila, et al., 2021) research used it for detecting the plant cassava disease due to its balance in accuracy, parameters, and speed. The architecture chosen for this research was the EfficientNet-B0 and the model outperformed all the previous models for cassava image classification. In the research (Vong & Dinh, 2021), the authors used EfficientNet architecture for the image segmentation of Pneumothorax in the Chest X-Rays. Radiologists check the Chest X-Rays manually for diagnosing Pneumothorax and the researchers used EfficientNet for automatic and quick detection. They used the EfficientNet-B4 architecture for this research and got a dice coefficient of 0.85 indicating the effectiveness of the EfficientNet architecture in the image segmentation tasks.

3 Research Methodology

This research analyses the performance of the EfficientNet architecture in classifying the earthquake and noise signals. The EfficientNet regression model is also used for predicting the arrival of P and S waves. This research follows the CRISP-DM data mining methodology. The research methodology discusses the steps involved in the classification of seismic signals with the use of the spectrogram and waveform plots of the raw seismic signals.

3.1 Data Collection and Data pre-processing

The dataset used for this research was the INSTANCE dataset (Michelini, et al., 2021) consisting of raw seismic waveforms and their associated metadata. This dataset was curated from the Italian Seismic data centers that are available in the INGV Bulletin online. The dataset consists of over 1,000,000 earthquake waveforms and over 100,000 noise waveforms. The waveforms selected were collected between January 2005-2020 as they coincided with the time, the sensors were upgraded in the National Seismic network. This ensured consistency across the data for all the waveforms. The selected earthquake waveforms were ensured they had the P-wave arrival time and the S-wave arrival time when applicable. This is important as it helps the EfficientNet regression model is used for predicting the onset of the P and S waves. The waveforms signal data were made available online as HDF5 files along with their associated metadata. The metadata file for the seismic events waveforms consists of around 115 columns that provide details about the source, path, station, and trace.

Fig 1 shows the distribution of the station channels and the networks that contributed to the waveforms. The channels with a higher gain (HH & EH) contribute to over 70% of the waveforms. This is significant as most of the waveforms used in the research have a higher quality. The HH channels are associated with high gain seismometers (51%) and the EH channels are associated with the extremely short period channels (19.5%). Fig 1b shows that the IV station network that represents the Italian National Seismic Network contributes over 78% of the waveforms for this research.

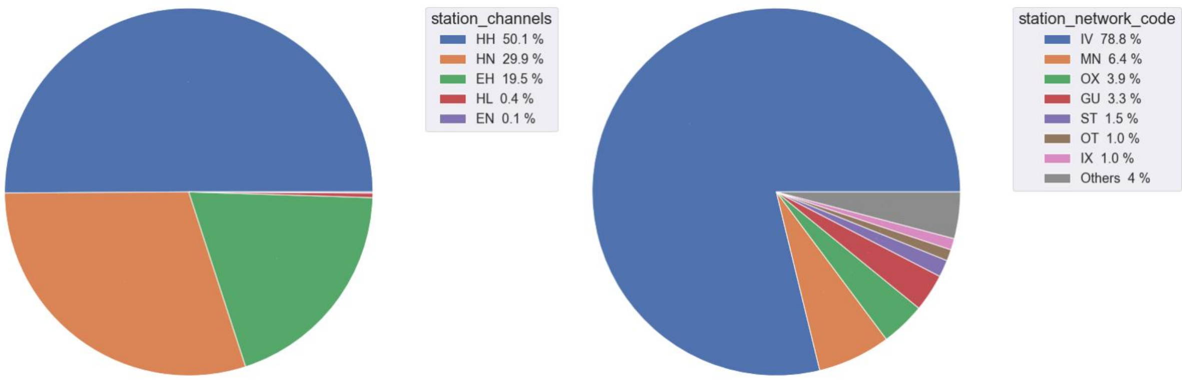


Figure 1: Distribution of traces by channels and station network (Michelini, et al., 2021)

The target variable for this research from the metadata columns is the "*source_type*" for the classification of signals, and the "*trace_P_arrival_time*" & "*trace_S_arrival_time*" for the phase picking. Fig 2 shows the distribution plot of the arrival time of the P and S waves. As observed the P arrival time and S arrival time of most of the samples were distributed around 2000 which indicates that the average onset time for the P and S waves has been around 20 seconds.

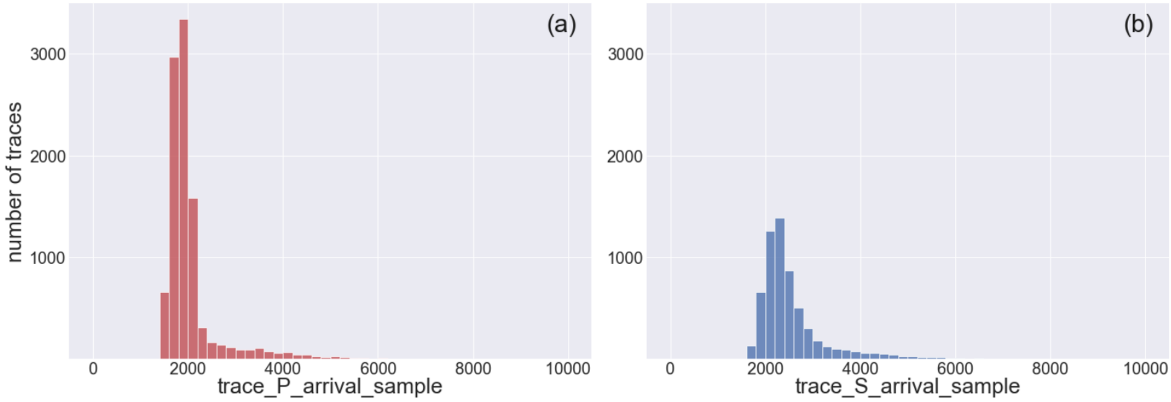


Figure 2: Distribution of P & S arrival times (Michelini, et al., 2021)

The complete dataset consists of 90% earthquake waveforms and 10% of noise waveforms. For this research, 20,000 earthquake and noise waveforms were sampled. The ratio of the earthquake and the noise waveforms were kept equal to ensure the model didn't have any bias towards either of the classes. Since in this research the EfficientNet architecture will be used on the waveform and the spectrogram plots, the raw signals need to be converted. The raw waveform signals were plotted using the "pyplot" function in the "matplotlib" library. The spectrograms were plotted using the "specgram" function. While converting the signals into grayscale will help in lower computation power, the EfficientNet architecture expects the number of channels in the image to be 3 and hence plasma was chosen as the colormap for this research.

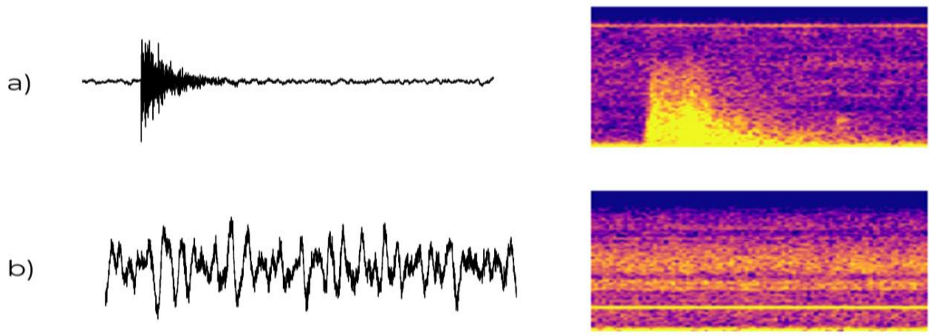


Figure 3: Sample waveform & spectrogram for a) Earthquake b) Noise

3.2 Modelling techniques

3.2.1 EfficientNet architecture

This research uses the EfficientNet architecture for the classification and the regression for the seismic waves. The EfficientNet architecture (Tan & Le, 2019) developed by Google, addresses the most common issue with the CNN architectures which was scaling it up. For this research as the model is based on training the images EfficientNet architecture has been chosen. The baseline EfficientNet-B0 model has been used in this research as it will give the base performance of the model which can then be scaled up in the future. The

EfficientNet B0 expects the images in the shape of (224,224,3) where the 224 represents the size of the image and the 3 represents the number of channels (RGB). The waveform images and the spectrogram images were ensured they were resized before training the model.

3.2.2 CNN architecture

Due to their convolutional ability, CNNs have gained immense success in the field of image processing. While this research focuses on the use of EfficientNet architecture on the INSTANCE dataset, a simple CNN architecture was also developed for benchmarking process. The CNN architecture was designed separately for the classification model and the regression model. The CNN architecture for this research used convolutional layers, pooling, dropout, and the dense layer. While the convolutional layers are the building blocks of CNN, the pooling layer was used for reducing the number of computations. The dropout layer was used to prevent the overfitting of the model and the dense layer was used to get the output for the classification and the regression models.

3.3 Evaluation methodology

The research proposes the use of the EfficientNet architecture on the INSTANCE dataset. To compare the performance of this architecture on the newly curated INSTANCE dataset, a small CNN architecture was used. Also, to compare the performance of the EfficientNet model with the other works discussed in the literature review, this model was also trained on the STEAD (STanford EArthquake Dataset). The metrics used to evaluate the classification model are Precision, Recall, and the F1 score. The precision helps in identifying the percentage of correct predictions while predicting an earthquake. The recall or the sensitivity helps in identifying the percentage of earthquakes that were detected from all the total earthquakes. Since while identifying earthquakes, false negatives must be minimized and hence recall serves as a better evaluation metric than the precision. The overall performance of the model can be evaluated based on the F1 score as it combines the precision and the recall as a single measure.

For the regression model, the metrics that were used to evaluate were the Mean Absolute Percentage Error (MAPE) and the Mean Squared Error (MSE). The mean errors indicate how close the predicted values were to the actual values. The baseline MSE and MAPE was calculated by predicting the mean values of the arrival times and compared with the model's results. The F1 score, precision, and recall were also calculated for the research by assuming a true positive if the prediction lay within 0.5s. This helps in evaluating the model performance against the other existing methods like in (Mousavi, et al., 2020).

4 Design Specification

The design architecture of the research is shown in Fig 4. The raw signals from the INSTANCE dataset are pre-processed and converted to the spectrogram images and the waveform images required for this research. The research uses the spectrogram images for training the classification model and the waveform images for the regression model. The Spectrogram image dataset was used as it provides the harmonic and temporal representations of the signals. This helps in better prediction even with lower image sizes and smaller datasets. For the phase picking, the spectrograms can't be used as they lose the phase information. Hence for the phase picking model, the waveforms are used. Different architectures were used for the EfficientNet classification and the regression model. The architectures of these EfficientNet models and the CNN are discussed below.

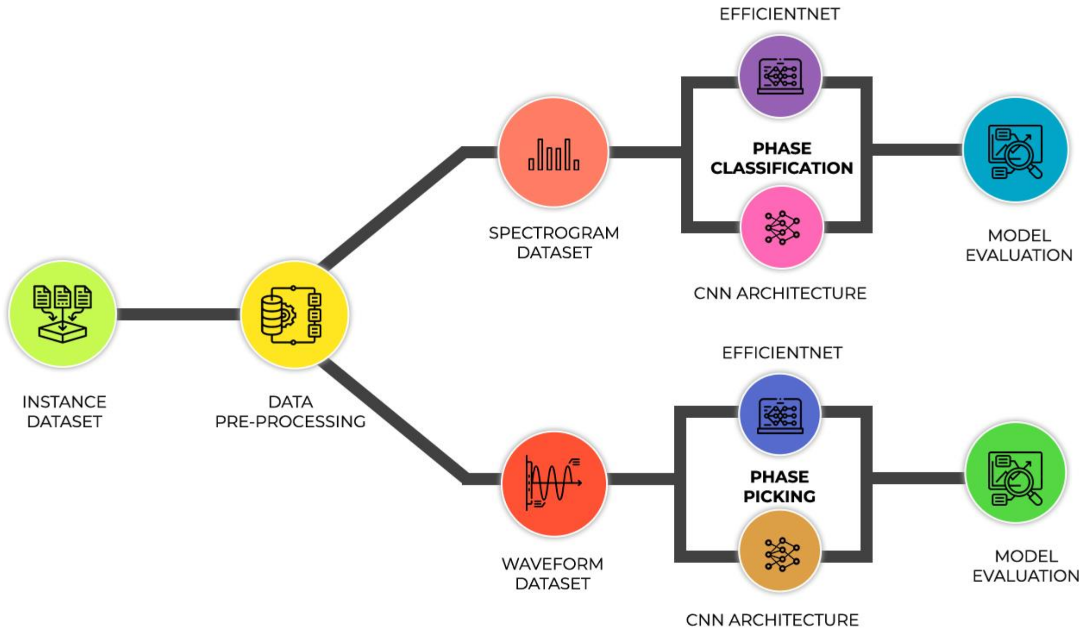


Figure 4: Design Architecture

4.1 EfficientNet architecture

The architecture chosen for this research was the EfficientNet architecture. The EfficientNet model has 8 architectures B0 to B7. The B0 is the baseline architecture, and the other architectures are the scaled-up versions with more parameters for better accuracy. For this research, the baseline network EfficientNet B0 was chosen. The EfficientNet B0 model expects the images in the shape of (224, 224, 3). The architecture of the EfficientNet classification model is shown in Fig. The output from the EfficientNet architecture is sent to a Global Max pooling layer and then to a dense layer that outputs the predictions if the signal was an earthquake or noise.

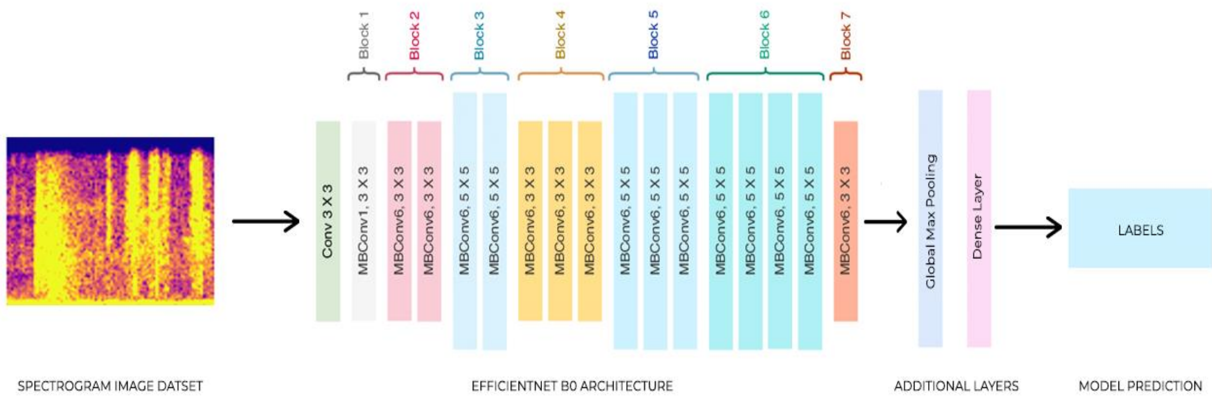


Figure 5: EfficientNet Classification architecture (Tan & Le, 2019)

The Fig shows the EfficientNet regression architecture used for the picking of the P and S arrival waves. The EfficientNet B0 architecture is also used for this model, along with a global max pooling layer, normalization layer, dropout layer, and a dense layer. The normalization layer is used as it helps to settle the learning process and thereby reduces the training epochs. The dropout layer is used to avoid overfitting the model. The dense layer outputs the prediction for the arrival time of the P and S waves.

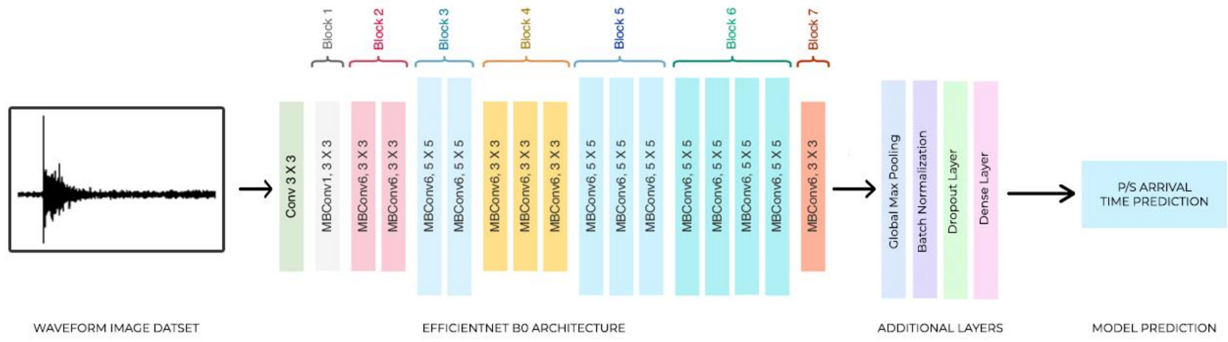


Figure 6: EfficientNet Regression architecture (Tan & Le, 2019)

4.2 CNN architecture

For comparing the performance of the EfficientNet architecture on the dataset, a simple CNN architecture was also used for benchmarking process. The architecture used for the CNN classification model is shown in Fig. The CNN architecture consisted of a convolutional layer, max pooling layer, flatten layer, and dense layer. The convolutional layers and the max pooling layers are used to extract the important features after downscaling the image. The Flatten layer converts the output from the pooling layer to a 1D output which is then passed through the dense layers to get the prediction for earthquake and noise.

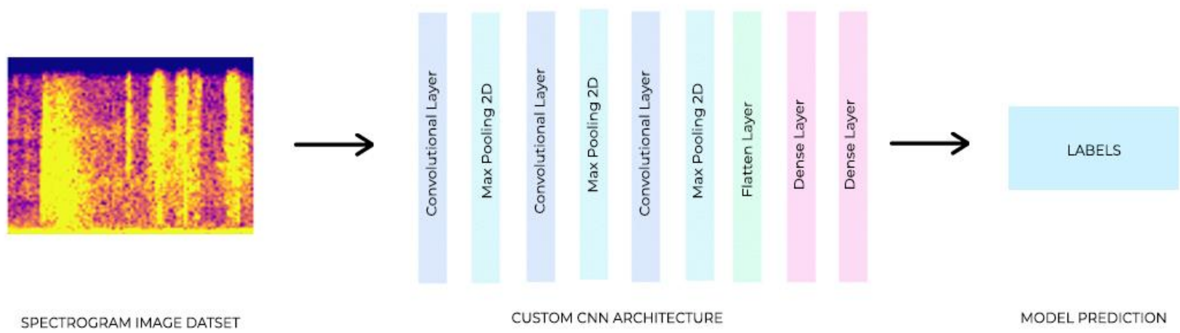


Figure 7: CNN Classification architecture

The CNN regression architecture used for the prediction of the arrival time of the P and S waves has been shown in Fig. The architecture used consists of three Convolutional layers, two Max Pooling layers, a flatten layer, and a couple of dense layers. The output from the convolutional and max pooling layers is passed to the flatten layer which converts it into 1D output which is then passed through the dense layers to get the prediction for the arrival time of the P/S waves.

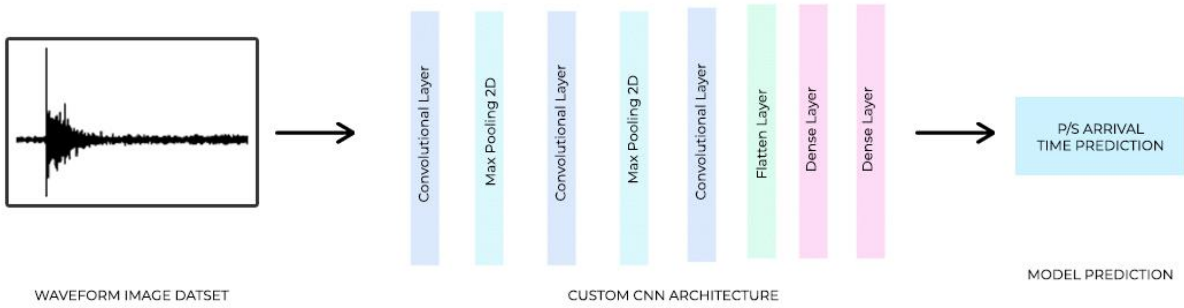


Figure 8: CNN Regression architecture

5 Implementation

The Implementation of the research was done on Jupyter notebooks using Python language. For using the EfficientNet-B0 architecture the TensorFlow was updated to the latest version and was imported from the TensorFlow Keras applications. The research followed different implementations for the phase classification model and the phase picking model that are explained below.

5.1 Phase Classification

The Spectrogram image dataset used for the research was split into three directories train, test, and validation in the ratio of 60:20:20. To process the images for the machine learning model, the Keras image generator class is used. The "flow_from_directory" helps in processing the image from the train, test, and validation directories into separate generators. The images in the generator are read in batches of 32 and are resized into (224,224,3) which is the input shape that is required for the EfficientNet-B0 architecture. The images are shuffled to prevent the model from having any bias. The EfficientNet-B0 classification architecture discussed above is compiled with the "Adam" optimizer to optimize the input weights by comparing the loss function with the prediction. The metrics for the model is chosen as "accuracy" and the loss function is chosen as "categorical cross entropy". The model used two callback functions, the early stopping callback, and the model checkpoint callback. The model checkpoint callback was used to save only the model with the best accuracy at every epoch so that it can be loaded later for training. The early stopping callback helped to monitor the model if the loss function didn't improve much over multiple epochs. The patience for the model was set at 10 to ensure the model stops early only when the loss has saturated over 10 epochs. The model was trained for 50 epochs and then evaluated with the test dataset. For the CNN architecture, a sequential model was created using the Keras library. The layers were added as described in the above architecture. The model was also compiled with the same parameters given for the EfficientNet architecture. The evaluation metrics were calculated for both the EfficientNet architecture and the CNN architecture.

5.2 Phase Picking

The waveform dataset is used for the phase picking model as the phase information gets lost while plotting the spectrogram. For the phase picking, the EfficientNet Regression architecture is used to predict the arrival times of P and S waves. Since it is a regression model, the "flow_from_directory" function cannot be used as the folder structure will be applicable only for the categorical values. For reading the images for this model, the path of the image files and their associated labels for the P and S values are stored in a panda DataFrame. The labels for the respective images are found using the metadata file. The pandas DataFrame was then split into train, test, and validation using the "train_test_split" function under the "sklearn" library. The "flow_from_dataframe" under the Keras Image generator library is used to convert the train, test, and validation DataFrames into their respective generators. The generator also resizes the image in the shape of (224,224,3) required for the EfficientNet-B0 architecture. The callbacks function for early stopping and model checkpoint were used similar to the classification architecture. The model was compiled with the rectified adam unit optimizer and the loss function chosen was the mean absolute error. The metrics chosen for the model were the mean absolute error and mean absolute percentage error and the model was trained for 50 epochs. The CNN Regression architecture discussed above was used to predict the arrival times. The CNN architecture was also compiled with the same functions and the model was trained with the train generator. The performance of both the architectures was evaluated using the test generator.

6 Evaluation

The motive of this study was to evaluate the performance of EfficientNet architecture in seismic phase detection and phase classification. The first section evaluates the performance of the EfficientNet architecture and the CNN architecture that was trained using the INSTANCE dataset. The second section evaluates the performance of both the model on the STEAD dataset to compare their performance with some of the models discussed in the literature review.

6.1 INSTANCE Dataset

6.1.1 Phase Classification Model

Since the phase classification model is used for binary classification, the confusion matrix of the prediction of test data helps in evaluating the model. The confusion matrix for the EfficientNet-B0 architecture is shown in Fig 9(a). The model rightly classified 1652 earthquakes. While the model wrongly classified 151 noise signals as earthquakes the model classified 348 earthquakes as noise signals. The model had an overall accuracy of 87.5% with a precision of 91.6% and a recall of 82.6%. In the case of classification of earthquakes having a higher sensitivity is more important than accuracy, but in this case, the model had trouble classifying earthquakes more than the noise signals. The ROC curve (Receiver Operating Characteristic) assesses the performance of a classification model at all the thresholds. The ROC curve is shown in Fig 9(b). The AUC (Area under the ROC Curve) for this classification model is 88% which indicates the overall ability of the model to distinguish between the noise and the earthquake signals.

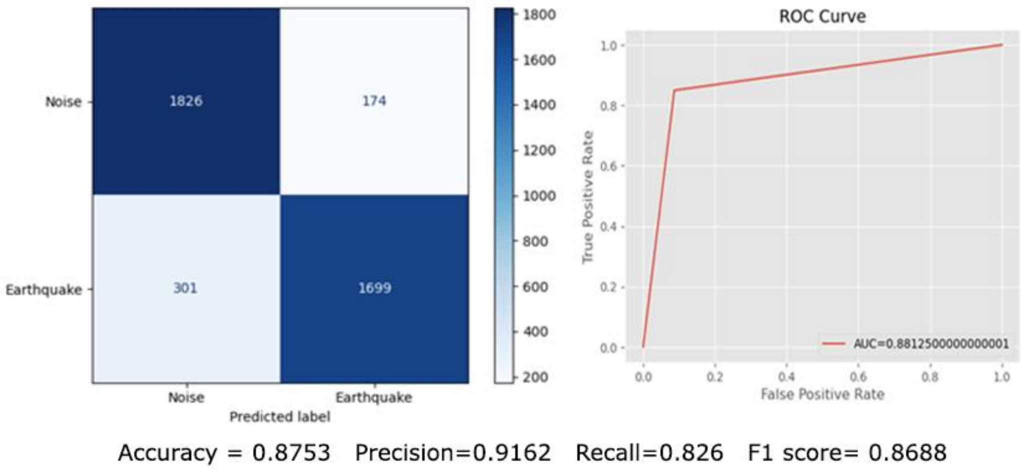


Figure 9: a) Confusion Matrix b) ROC curve for EfficientNet - INSTANCE

The Fig 10(a) shows the loss and accuracy curve of the training model. The loss curve for every epoch indicates the model training process and the direction in which it learns. The training loss curve indicates the model had a good learning rate. The validation loss curve is around the training loss curve which indicates the good fit of the model. The training accuracy curve and the validation accuracy curve indicate the model was not overfitted. The Fig 10(b) shows the model accuracy of the train and the test dataset after every epoch. Both the validation dataset and the test dataset experience oscillations over the training curve. While the model of the testing and validating accuracy doesn't fall below 0.8 it is important to diagnose it. The possible reasons for the oscillations could be due to the presence of noisy data (spectrogram images with low SNR) that might lead to the model making assumptions for the predictions. Another possible reason could be the size of the neural network architecture. Though the research uses the baseline EfficientNet-B0 network the size of the network is still huge which could have caused the oscillations. This can be verified with the plots of the CNN architecture.

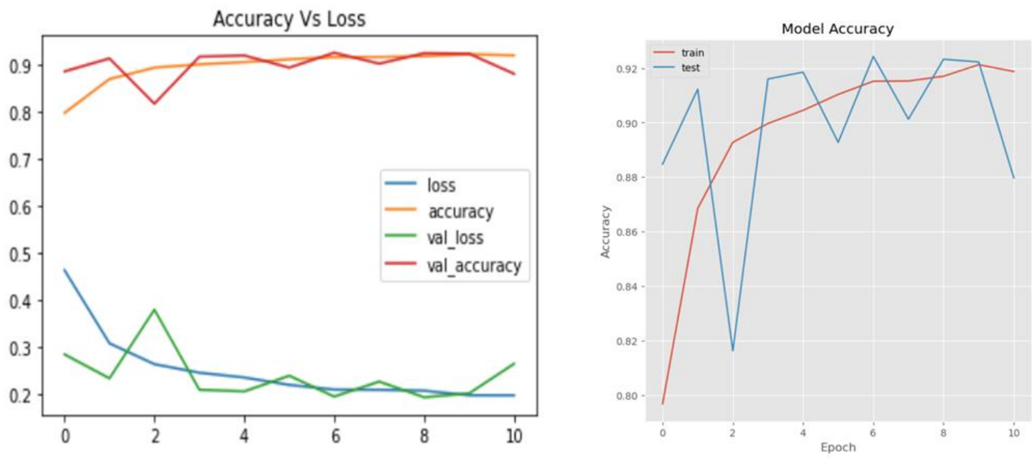


Figure 10: Plots - a) Accuracy vs Loss b) Test accuracy over epoch

Fig 11 shows the confusion matrix and the ROC curve for the CNN architecture. The CNN architecture was able to classify the 1899 earthquakes from the 2000 earthquake signals. The CNN architecture had an overall accuracy of 94.7% while classifying the earthquake and the noise signals. The model had a sensitivity of 94.9% which was very higher than the EfficientNet architecture. The AUC curve was also 94.7% which indicates the model was a good fit for classifying the earthquake and the noise signals.

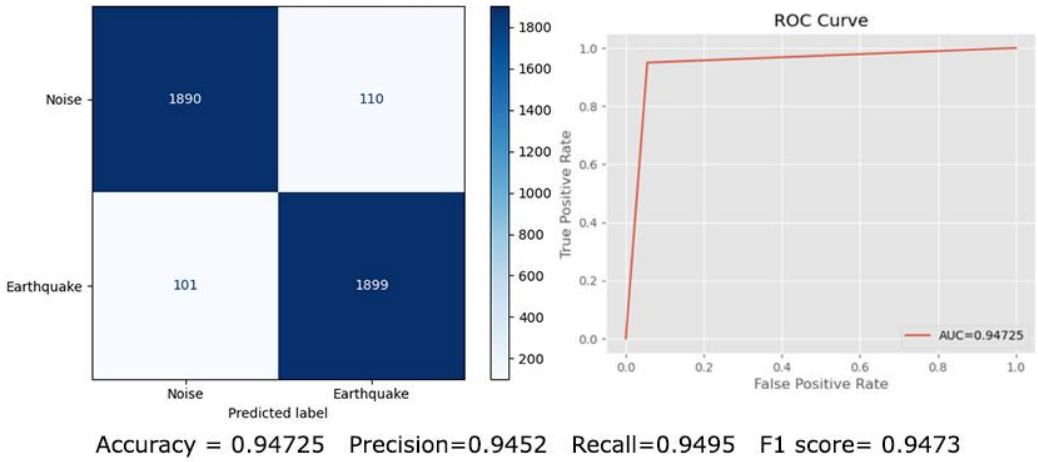


Figure 11: a) Confusion Matrix b) ROC curve for CNN - INSTANCE

The accuracy and the loss plot for the loss and the accuracy for the CNN architecture is shown in the Fig 12(a). The validation accuracy curve is slightly below the training curve indicating a slight overfitting. The accuracy vs loss plot doesn't have many oscillations compared to the EfficientNet architecture indicating that the complexity of the architecture could have been the cause. The plot of the train and the test accuracy in Fig 12(b) still has few oscillations which could be attributed to the noise in the dataset. Overall, considering all the metrics the CNN architecture has performed better than the EfficientNet architecture.

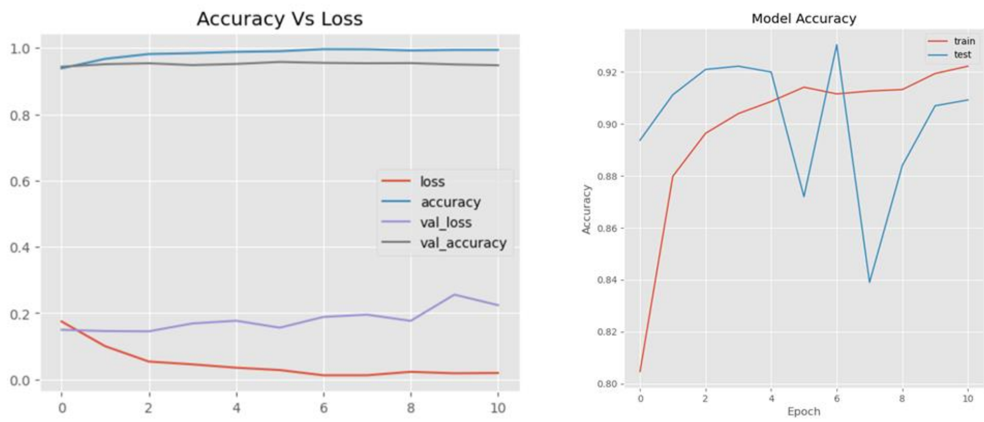


Figure 12: Plots - a) Accuracy vs Loss b) Test accuracy over epoch

6.1.2 Phase Regression Model

The Phase Regression model was used for predicting the arrival of the P and the S waves. The Mean Absolute error (MAE) and the Mean Absolute percentage error (MAPE) were diagnosed for the prediction of the arrival times. The baseline MAE and the MAPE for the prediction of P arrival waves that were calculated using the mean were 358 and 15.78% respectively. The test dataset was used to evaluate the MAE and the MAPE for both architectures. The EfficientNet architecture had an MAE of 308 with a MAPE of 12%. Similarly, the CNN Regression architecture had an MAE 291 of and MAPE of 12% on the test data. The true picks of the models were calculated if the predicted value lay within a range. The EfficientNet architecture was able to pick only 48.2% of the data within 1.5s while the CNN architecture was able to pick only 48.4%. The Fig 13 shows the history plot for the MAPE of both the architectures for both the training and the validation dataset. The CNN architecture was able to train faster compared to the EfficientNet architecture.

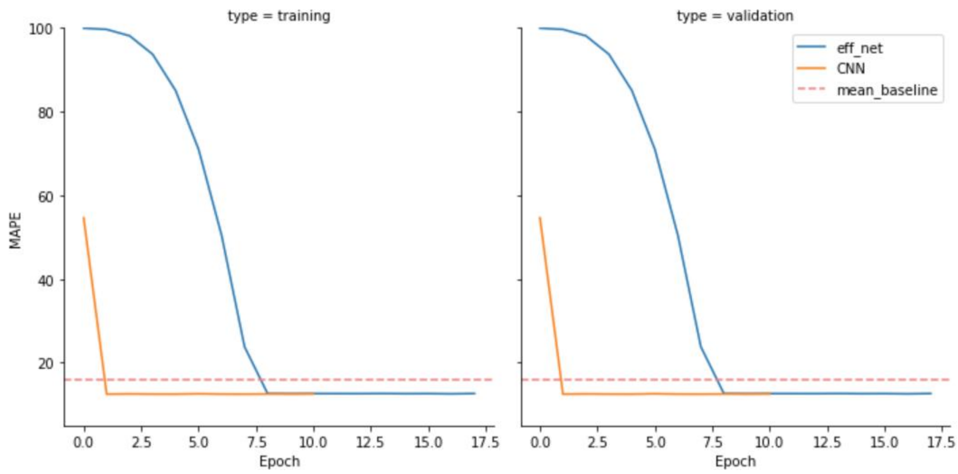


Figure 13: MAPE for picking P waves

For picking the S arrival waves, the same architectures were used with the labels changed. The baseline MAPE and the MAE for the S arrival waves of the INSTANCE dataset were 18% and 466 respectively. The EfficientNet architecture had a MAPE and MSE of 15% and 488 respectively while the CNN architecture had MAPE and MAE of 9% and 278. The CNN had lower error rate than the EfficientNet architecture while picking the S waves. The Fig 14 also indicates the CNN model took fewer epochs to train than the EfficientNet architecture. The true picks for both the models were almost similar.

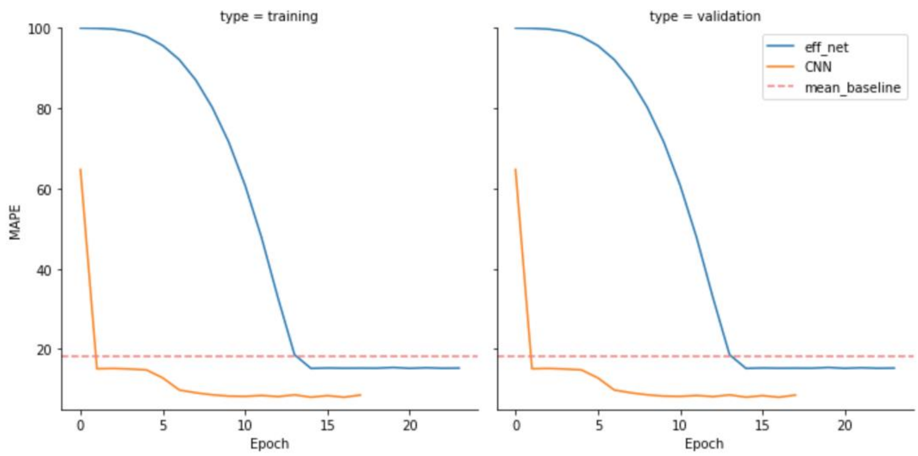
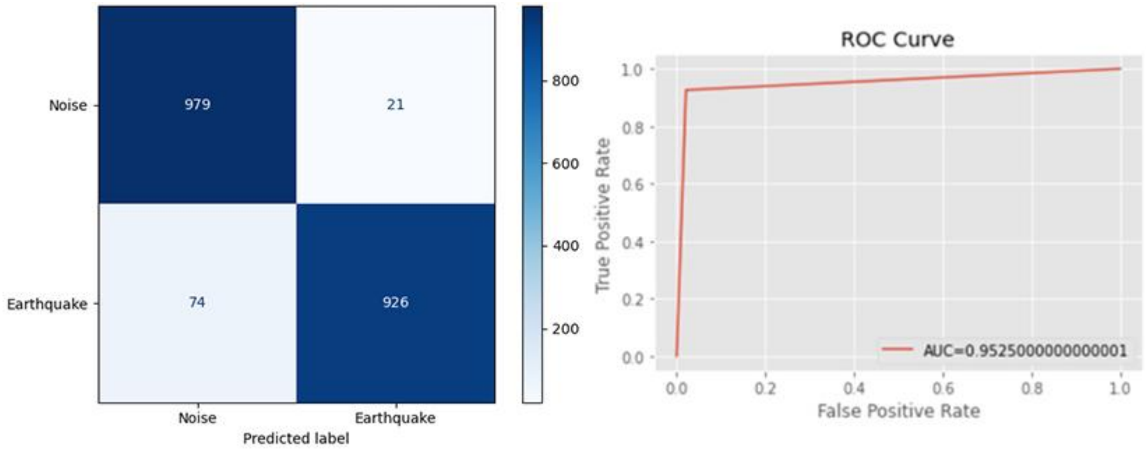


Figure 14: MAPE for picking S waves

6.2 STEAD Dataset

6.2.1 Phase Classification Model

The same EfficientNet classification architecture was used for training with the STEAD dataset. The confusion matrix and the ROC curve for the model prediction on the test data are shown in Fig 15(a). The EfficientNet architecture has performed much better on this dataset compared to the INSTANCE dataset. The model had an overall accuracy of 95% with a sensitivity of 92.6%. Similar to the model trained on the INSTANCE dataset, the model had a lower recall than the precision. This indicates the EfficientNet architecture used for this phase classification had a slight issue while identifying earthquakes. The AUC and F1 scores were 95% which indicates the overall good fitness of the model for both the classes.



Accuracy = 0.9525   Precision=0.9728   Recall=0.926   F1 score= 0.9512

Figure 15: a) Confusion Matrix b) ROC curve for EfficientNet - STEAD

The accuracy vs loss curve for the CNN architecture shows that the validation data was a good fit for the model. The test curve for every epoch also indicates the performance doesn't oscillate much like the EfficientNet model trained on the INSTANCE dataset. This indicates that most of the oscillations in the validation could have been caused due to the noise in the dataset.

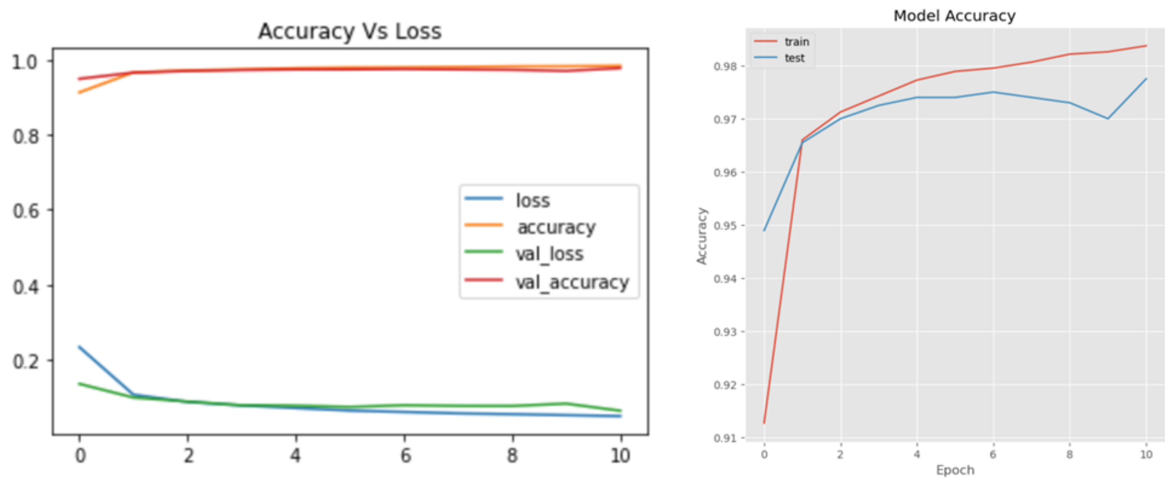


Figure 16: Plots - a) Accuracy vs Loss b) Test accuracy over epoch

The CNN architecture was also trained with the STEAD to check if it outperforms the EfficientNet architecture. The model had an accuracy and a recall of 98% on the test data indicating the model's suitability for the STEAD dataset. The AUC was also 98% showing it outperformed the EfficientNet architecture. The plot showing the training accuracy and test accuracy for every epoch shown in Fig 17(b) shows there were fewer oscillations compared to the other models. Overall, the CNN architecture outperformed the EfficientNet-B0 architecture in both the INSTANCE and the STEAD dataset.

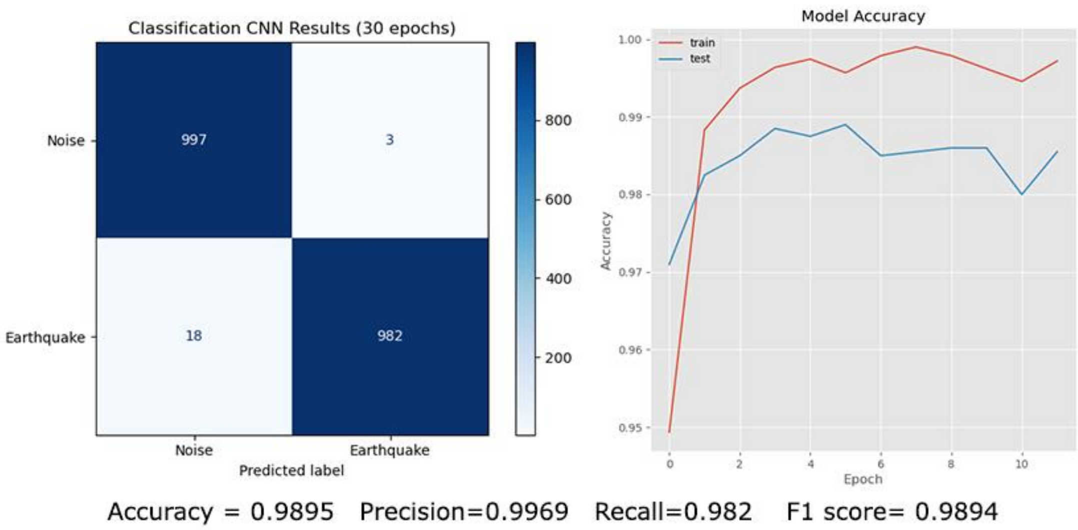


Figure 17: a) Confusion Matrix b) Test accuracy curve for CNN - STEAD

6.2.2 Phase Regression Model

The same EfficientNet and the CNN regression architectures were also tested on the STEAD dataset. The baseline MAPE and the MAE for the P waves arrival were 27% and 158. The EfficientNet model had a MAPE and MAE of 24% and 156 on the test data while the CNN architecture had 25% and 186. The Fig 18 indicates that both the models performed only slightly better than the baseline value indicating the poor fit of the model. The CNN took a longer time to train than the EfficientNet architecture which could possibly be due to the higher learning rate used while building the model.

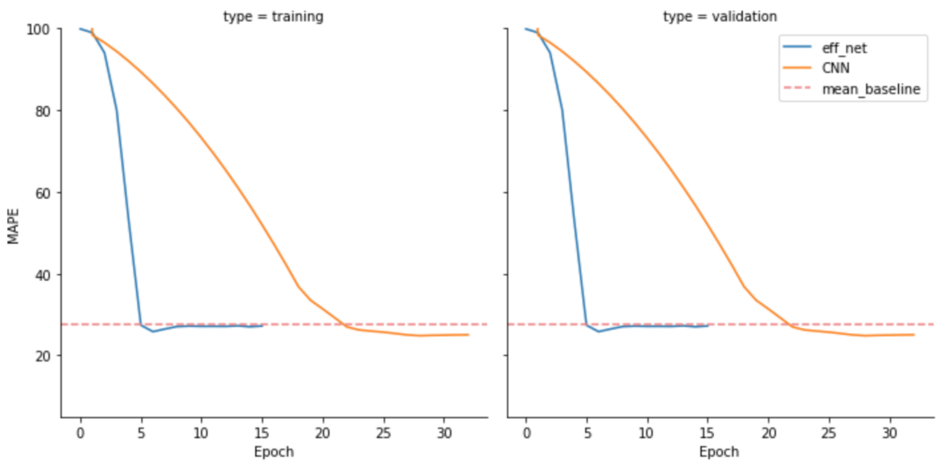


Figure 18: MAPE for picking P waves

Similarly, the performance was analyzed for predicting the arrival time of the S waves. While the baseline MAE and the MAPE were 343 and 31%, the EfficientNet had 157 and 24% and the CNN model had 69 and 6% respectively. As observed from the Fig 19, the MAPE of the CNN is very low compared to the EfficientNet model but it doesn’t translate into better predictions. While the EfficientNet had a true pick rate of 30%, the CNN had a pick rate of 25%. This could possibly be because MAPE penalizes the undercasting more than overcasting of the predicted value which could be the case with a lower MAPE for the

CNN architecture, Overall considering all the evaluation metrics, both the architectures in the STEAD dataset performed quite poorly.

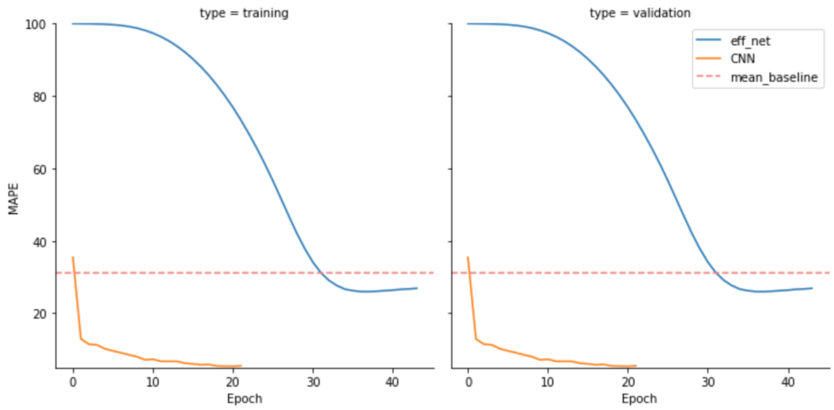


Figure 19: MAPE for picking S waves

7 Discussion

In the Phase classification model, the CNN architecture outperformed the EfficientNet model in both the INSTANCE dataset and the STEAD dataset with a higher F1 score and recall. Though the EfficientNet model used was the baseline network B0 which could have been scaled for higher accuracy, the research shows that the model with a few layers designed for a particular dataset can outperform the more complex architectures. One of the major issues with the model trained on the INSTANCE dataset was the oscillations that were present during the validation dataset. The oscillations reduced considerably while using the STEAD dataset indicating the presence of potential noise in the INSTANCE dataset. The performance of the model increased drastically while training with the STEAD dataset. While an F1 score of 0.95 and 0.98 for the EfficientNet model and the CNN model seems high, comparing it with the other methods shown in Table 1, the performance is almost on par with the other models. Both the models were also trained with only 20,000 waveform signals compared to the other methods that were trained with more waveforms. The main drawback with using the EfficientNet architecture was its poor recall value compared to the other existing methods. While these metrics are evaluated on the test split of the STEAD dataset, it doesn't necessarily translate into waveforms outside the test dataset as seen in the research (Jiang, et al., 2021).

Table 1 Comparison with the other models

Model	***F1***	***Precision***	***Recall***	***Reference***
STA/LTA	0.95	0.91	1	(Allen, 1978)
EQTransformer	1	1	1	(Mousavi, et al., 2020)
CapsPhase	0.97	0.94	0.99	(Saad & Chen, 2022)
EfficientNet-B0	0.95	0.97	0.92	This paper
CNN architecture	0.98	0.99	0.98	This paper

To test the generalizability of the EfficientNet-B0 model trained on the STEAD dataset, the model was tested with the test dataset of the INSTANCE dataset and the model had an overall accuracy of 83%. Similarly, the model trained on the INSTANCE dataset had only an accuracy of 70% on the STEAD dataset. This shows both the models suffered from poor generalizability. The INSTANCE dataset consists of waveforms only from the seismic centers around Italy, which could be the reason it failed on the STEAD dataset consisting of waveforms curated from all over the world. As the prediction of earthquake is a highly sensitive issue, the model's interpretability is necessary to understand the edge cases where it will fail. The Fig 20 shows a random sample of wrong predictions made by the EfficientNet model on the STEAD dataset. Most of the actual noise and earthquake signals are very similar which led to the model making wrong predictions. The model also seems to suffer when the signal-to-noise ratio is very low. For the phase regression model, both the architectures were only slightly lower than the baseline values. The EfficientNet model was able to slightly outperform the CNN architecture while calculating the true picks. The EfficientNet model was only able to classify 48% of the P waves and 30% of the S waves within 1.5s, compared to the EQTransformer and PhaseNet which had an F1 score of 98% and 90% while classifying the P and S waves with a smaller threshold. Both these models had used the raw signals, while in this research the waveform plot images were used which performed quite poorly compared to the existing models.

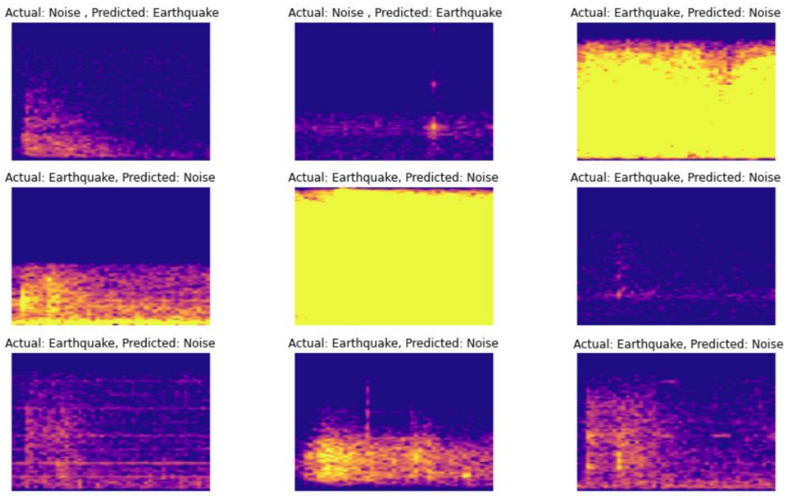


Figure 20: EfficientNet Wrong Predictions

8 Conclusion and Future Work

The motive of this research was to evaluate the performance of the EfficientNet-B0 architecture in phase classification and phase picking of seismic waves. The model was trained on the INSTANCE dataset along with a CNN architecture. For phase classification, the CNN outperformed the EfficientNet architecture with a higher F1 score and recall. Both the models were trained on the benchmark STEAD dataset and had an F1 score of 0.98 for the CNN and 0.95 for the EfficientNet architecture. While this is slightly lower than the current state-of-the-art models, these models were trained using only 20,000 waveforms from the dataset. The major drawback of generalizability in the STEAD dataset was also present in the models trained with the INSTANCE dataset. In phase regression, both the EfficientNet and the CNN architectures performed considerably weaker than the other models like PhaseNet, and EQTransformer. Phase picking has always been a challenging task in the field of observational seismology as they are more prone to errors and the use of regression architectures on the waveform plots didn't help in improving the performance of the model.

In future work, the phase classification models can be used to monitor the live seismic signals from the Incorporated Research Institutions for Seismology (IRIS) SeedLink client. The whole solution can be deployed online in a cloud environment which will send alerts when it predicts an earthquake. This might help in understanding the overall performance of the model and will allow testing the model with waveforms from around the world. The EfficientNet model can also be scaled up to check if the performance of the model improves for both phase classification and regression.

Acknowledgement

I would like to thank my research supervisor Dr. Vladimir Milosavljevic for his constant feedback and support during my course work.

References

Allen, R. V., 1978. Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America,* 68(5).

Atila, Ü., Uçar, M., Akyol, K. & Uçar, E., 2021. Plant leaf disease classification using EfficientNet deep learning model. *Ecological Informatics,* Volume 61.

Calderon, F., Perez, N., Benitez, D. S. & Riofrio, D., 2020. Volcanic seismic event classification based on CNN architectures. *2020 IEEE ANDESCON, ANDESCON 2020.*

Chen, Y., 2018. Automatic microseismic event picking via unsupervised machine learning. *Geophysical Journal International,* 212(1).

Choubik, Y., Mahmoudi, A. & Himmi, M. M., 2021. Fully Convolutional Networks for Local Earthquake Detection. *International Journal of Advanced Computer Science and Applications,* 12(2).

Curilem, M., Canário, J. P., Franco, L. & Rios, R. A., 2018. Using CNN to Classify Spectrograms of Seismic Events from Llaima Volcano (Chile). *Proceedings of the International Joint Conference on Neural Networks,* Volume 2018-July.

He, Z., Peng, P., Wang, L. & Jiang, Y., 2021. PickCapsNet: Capsule Network for Automatic P-Wave Arrival Picking. *IEEE Geoscience and Remote Sensing Letters,* 18(4).

Huang, X., Lee, J., Kwon, Y. W. & Lee, C. H., 2020. CrowdQuake: A Networked System of Low-Cost Sensors for Earthquake Detection via Deep Learning. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

Jiang, C., Fang, L., Fan, L. & Li, B., 2021. Comparison of the earthquake detection abilities of PhaseNet and EQTransformer with the Yangbi and Maduo earthquakes. *Earthquake Science,* 10, 34(5), pp. 425-435.

Jozinović, D., Lomax, A., Štajduhar, I. & Michelini, A., 2021. Rapid prediction of earthquake ground shaking intensity using raw waveform data and a convolutional neural network. *Geophysical Journal International,* 222(2).

Leonard, M. & Kennett, B. L., 1999. Multi-component autoregressive techniques for the analysis of seismograms. *Physics of the Earth and Planetary Interiors,* 113(1-4).

Liao, X. et al., 2020. First Arrival Time Identification Using Transfer Learning with Continuous Wavelet Transform Feature Images. *IEEE Geoscience and Remote Sensing Letters,* 17(11).

Li, Z. et al., 2018. High-resolution seismic event detection using local similarity for Large-N arrays. *Scientific Reports,* 8(1).

Michelini, A. et al., 2021. INSTANCE-the Italian seismic dataset for machine learning. *Earth System Science Data,* 13(12).

Mousavi, S. M. et al., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications,* 11(1).

Mousavi, S. M., Sheng, Y., Zhu, W. & Beroza, G. C., 2019. STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. *IEEE Access,* Volume 7.

Pardo, E., Garfias, C. & Malpica, N., 2019. Seismic phase picking using convolutional networks. *IEEE Transactions on Geoscience and Remote Sensing,* 57(9).

Saad, O. M. & Chen, Y., 2021. Earthquake Detection and P-Wave Arrival Time Picking Using Capsule Neural Network. *IEEE Transactions on Geoscience and Remote Sensing,* 59(7).

Saad, O. M. & Chen, Y., 2022. CapsPhase: Capsule Neural Network for Seismic Phase Classification and Picking. *IEEE Transactions on Geoscience and Remote Sensing,* Volume 60.

Sgibnev, I., Sorokin, A., Vishnyakov, B. & Vizilter, Y., 2020. DEEP SEMANTIC SEGMENTATION for the OFF-ROAD AUTONOMOUS DRIVING. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives,* 43(B2).

Tang, H., Wang, Y., Yuan, W. & Sun, Y., 2021. An Efficientnet Based Method for Autonomous Vehicles Trajectory Prediction. *2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology, CEI 2021.*

Tan, M. & Le, Q. V., 2019. EfficientNet: Rethinking model scaling for convolutional neural networks. *36th International Conference on Machine Learning, ICML 2019,* Volume 2019-June.

Van Der Laat, L. V., Baldares, R. J., Chaves, E. J. & Meneses, E., 2021. OKSP: A Novel Deep Learning Automatic Event Detection Pipeline for Seismic Monitoring in Costa Rica. *3rd IEEE International Conference on BioInspired Processing, BIP 2021 - Proceedings.*

Vong, K.-M. & Dinh, T. B., 2021. Pneumothorax Segmentation In Chest X-Rays Using UNet++ And EfficientNet.

Zhu, W. & Beroza, G. C., 2019. PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1)