HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 1 (2024)

A Deep Semi-Supervised Domain Generalization Approach for Epileptic Seizure Prediction using Electro Encephalo Graphy (EEG)

Sunwoo An[1] and Koby Osei-Mensah[#]

[1]The Loomis Chaffee School, Republic of Korea
[#]Advisor

ABSTRACT

According to the World Health Organization, nearly 50 million people suffer from epilepsy, one of the most common neurological disease. Epilepsy is characterized by abnormal brain activity, leading to recurrent seizures. Each seizure manifests as sudden, uncontrolled bursts of electrical activity in the brain, and the injuries and restrictions on daily life underscores the urgency of finding effective methods for epileptic seizure prediction. With the use of deep learning techniques, early predictions of epileptic seizures, an unsolved problem, are attempted in this paper. Previous research has limitations of being sensitive to noise as it is dependent on specific electroencephalogram (EEG) devices and datasets, a serious issue this paper solves. In this paper, a semi-supervised based domain generalization method to develop an accurate seizure prediction system is proposed. It consists of two phases: representation learning and transfer learning phase. To achieve high precision, the proposed method utilizes a representation learning approach. Here, a feature-swapping mechanism that effectively disentangles seizure-related features is introduced. During transfer learning, the pre-trained network is trained to output the probability of whether the input EEG indicates a seizure or not. The proposed model achieves state-of-the-art performance, with an accuracy of 90.53% and 94.88% on the NICU and Epileptic Seizure Recognition datasets respectively in within-dataset evaluations. It outperforms the previous methods by 19.35% in cross-dataset evaluations. This robust improvement opens up promising possibilities for real-world clinical applications. The proposed feature disentangling method is also expected to contribute to developing reliable medical tools.

Introduction

Problem Definition

Background

According to the World Health Organization, nearly 50 million people suffer from epilepsy worldwide, making it one of the most common neurological diseases globally (World Health Organization, 2023). Only migraine, stroke, and Alzheimer's disease are ahead in the list before epilepsy. Epilepsy is a neurological disorder characterized by abnormal brain activity, leading to recurrent seizures, and can affect individuals of any age. Each seizure manifests as sudden, uncontrolled bursts of electrical activity in the brain, resulting in a wide range of symptoms depending on the affected regions such as the stiffening of the body, loss of consciousness or breathing problems (Pruhiti et al., 2023). They are caused by a sudden abnormal, self-sustained electrical discharge that occurs in the cerebral networks and usually lasts for a few minutes (Rasheed et al., 2020). The unpredictable nature of seizures pose a significant threat to the life of those who suffer from epilepsy. The potential for injury

and the restrictions it imposes on daily life underscores the urgency of finding effective methods for epileptic seizure prediction.

Seizure Detection from Electroencephalograms

Seizures are detected using electroencephalograms (EEG), a test that detects abnormalities in the brain waves or in the electrical activity of the brain using small, metal discs (electrodes) attached to the scalp (Johns Hopkins University, 2023). During the recording, the electrode detects any electrical charges that result from the activity of the brain cells. The charges are then amplified and graphed for the healthcare provider to interpret the readings. When epilepsy is present, seizure activity will appear as rapid spiking waves on the EEG recordings. Out of the several types of EEG tests, ambulatory EEG (aEEG) are commonly used to predict epileptic seizures in daily life as brain activity can be recorded throughout a few days, up to 72 hours. The electrodes are attached to a small portable EEG recorder, and patients can continue with most of their normal daily activities while the recording is being taken (National Health Service, 2023).

The early detection of epileptic seizures in their pre-ictal states is crucial as seizures can have devastating consequences, including physical injuries, loss of consciousness or even death if not treated. The ability to predict seizures in advance holds immense potential for minimizing these risks and improving patient outcomes. Early detection allows individuals with epilepsy to take precautionary measures, such as moving to a safe environment, alerting a caregiver, or taking medication like anti-epileptic drugs. These kinds of timely administration of medication or other interventions can help mitigate the severity and duration of seizures. Early detection also offers individuals with epilepsy a sense of control over their condition, reducing their anxiety and thus empowering them to have the confidence to lead a normal life.

Objectives and Significance of Research

While significant progress has been made in the field of epileptic seizure prediction using EEG signals, further research is essential for several reasons. First of all, the accuracy of predictions needs to be enhanced. Although some existing prediction models have shown acceptable results, there is still room for improvement in terms of accuracy and reliability. In the medical field and the prediction of illnesses, reducing false negative errors are pivotal as the worst case is the illness not being detected properly.

Secondly, two crucial problems that can be identified from previous research is that the models are sensitive to noise and are biased to specific datasets and EEG devices. However, as real-life EEG devices are not identical in number of nodes and sensitivity, the noise and datasets will not be uniform. Further research is needed to improve these problems, and I will prove how my proposed method is insensitive to noise and is independent from one dataset or EEG device.

Previous Method

This chapter presents a comprehensive review of the existing literature related to the detection of epileptic seizures via EEG signals. Throughout the chapter, key findings, methodologies, outcomes and limitations of selected studies will be analyzed.

To address the aforementioned problem, various researches have been conducted. Before we analyze each research, a brief overview of the history in predictions of seizures with EEG signals and machine learning will be given.

After the first international collaborative workshop on seizure prediction was held in 2004, numerous researchers attempted to predict epileptic seizures using EEG signals. The first research that showed the feasibility of seizure predictions was “Identification of Epilepsy Seizures Using Multi-resolution Analysis and Artificial Neural Networks" written by Pilar Gómez-Gil. The study presented an accuracy of 99.26 ± 0.26 %, a sensitivity of 98.93 % and a specificity of 99.59 %, using data provided by the University of Bonn (Andrzejak

RG et al., 2001). Although the accuracy was mentioned to be 99.26%, this paper has low external validity; thus practically can not be used in real life. After this study, epilepsy datasets were formed properly and techniques such as feature extraction in time and frequency and neural networks developed. Several selected developed researches will be analyzed in this chapter.

Fang et al. proposed a seizure detection model using the Spatio-Temporal Gated Recurrent Unit (ST-GRU) and Convolutional Neural Networks (CNN). They employed time-domain features and used the CHB-MIT dataset. The accuracy achieved was 77.30%, making the biggest limitation low level of accuracy.

Bizopoulos et al. utilized SoftMax activation and standard neural networks to detect seizures. They used the concept of 2D and 3D phase space presentations to capture the intrinsic mode and functions of EEG signals. It was conducted on the BONN dataset (Andrzejak RG et al., 2001), yielding an accuracy of 85.30%. Again, the biggest limitation was low detection accuracy.

Yao et al. proposed a SoftMax and Long Short-Term Memory (LSTM) based model, using independent Recurrent Neural Networks (RNNs). The CHB-MIT dataset was used, and the accuracy was 88.80%. The study indicated limitations in terms of low sensitivity and precision.

Yuan et al. applied the SoftMax and Sparse Deep Autoencoder (SpDAE) model, using frequency-time domain and Continuous Wavelet Transform (CWT) for feature extraction. The CHB-MIT dataset was used, and the study achieved an accuracy of 90.82%. However, there were limitations again in low detection accuracy.

Turk and Ozerdem proposed a model that uses Softmax activation and 2D Convolutional Neural Networks (2DCNN), with frequency-time domain and Continuous Wavelet Transform (CWT) feature extraction techniques. The model used the Freiburg dataset and demonstrated an accuracy of 93.60%. However, the study had clear limitations of having low specificity for multi-class classification.

Talathi and Vartak utilized Recurrent Neural Networks (RNNs), specifically Gated Recurrent Units (GRU) for seizure detection. The BONN dataset was used, and the accuracy was 94.00%. However, the model had limitations of high time complexity.

San-Segundo et al. applied the SoftMax and 2D Convolutional Neural Networks (2DCNN) models, using Discrete Wavelet Transform (DWT) for feature extraction. The CHB-MIT dataset was used, resulting in an accuracy of 96.10%. The study had limitations of having a long training time.

Overall, all studies evaluated had significant limitations. The biggest problem is that all models are extremely sensitive to noise as it is dependent on specific devices and datasets used to train the model. This is a serious issue as the EEG device is not identical in number of nodes and sensitivity, and the datasets as well as real-life EEG signals are not uniform, but all the models will not be able to predict epileptic seizures accurately. Moreover, the accuracy was quite low in most studies; this is unacceptable especially when predicting seizures as the models should predict nearly every single one and alert the patient.

In this research paper, in order to solve the aforementioned problem, I propose a novel seizure detection framework and the methodologies will be further explained in the following chapter.

Proposed Method

To solve the problem of models being sensitive to noise and its bias on specific datasets, I propose a novel representation learning for seizure detection framework. The proposed framework is composed of the representation learning and the transfer learning phase.

There are three main contributions of the proposed method. First of all, the problems such as models being sensitive to noise and being biased on specific devices and datasets are considered. Secondly, they are solved using innovative methods such as representation learning and transfer learning. Third, experiments are conducted to prove that the proposed method is verified and accurate.

Moving on to the first key part of the model's architecture, the Autoencoder, used in the representation learning to extract the optimal features from the entangled latent features is explained. Originally, the features

extracted from the CNN included a combination of features extracted from the inputted Electroencephalogram (EEG) signals, often entangled. To increase the accuracy of seizure classification, the disentanglement and extraction of seizure-related features are pivotal as training on one feature will enhance the performance of the model.

In order to evaluate if the feature extraction is performing well, the L1 loss function is used. This compares the original input data and data reconstructed using only the extracted features. If the feature extraction is performing well, the two data will be similar. The function will be explained further in Chapter 3.1.

Another key method proposed is the Seizure Classifier. It uses transfer learning, a machine learning technique where knowledge gained from solving one problem is applied to a different but related problem. In other words, it will use a pre-trained model from the previous step. The Seizure Classifier will receive an input of EEG signals, which the pre-trained Convolutional Neural Network (CNN) including the Autoencoder will extract the seizure-related features from. The neural network then trains using these features, outputting the probability of the input will lead to a seizure.

In order to evaluate the performance of the Seizure Classifier, the Cross-Entropy loss function is used. It measures the difference between predicted probabilities and the true labels of a given dataset. The functions will be explained further in Chapter 3.2.

This research paper comprises five chapters. In Chapter 2, background knowledge is explained. In Chapter 3, the proposed methods are explained. In Chapter 4, the experimental results are explained. Lastly, in Chapter 5, a conclusion is given to summarize the research.

Related Work

In this chapter, the details regarding the background knowledge related to the research topic are presented and discussed.

Electroencephalograms

Electroencephalograms (EEG) is a non-invasive neurophysiological technique that detects abnormalities in the brain waves or in the electrical activity of the brain using small, metal discs (electrodes) attached to the scalp. Neurons, which make up the brain's electrical charge, are other. When the wave reaches the electrodes on the scalp, they can push or pull the electrons located in the metal of the electrodes. Since metal conducts the interaction of electrons easily, the difference is electrically charged through membrane transport proteins that pump ions across their membranes. They incessantly exchange ions with the extracellular milieu, pushing out countless ions simultaneously. This causes volume conduction, which is where ions released push their nearest ions, and those ions the next bordering ones, and as such in a wave as oppositely charged ions repel each in push or pull voltages between any two electrode are measured by voltmeters; making up the EEG signals as time passes (Tatum Wo et al., 2008). The charges are then amplified and graphed for the healthcare provider to interpret the readings. When epilepsy is present, seizure activity will appear as rapid spiking waves on the EEG recordings (Johns Hopkins University, 2023).

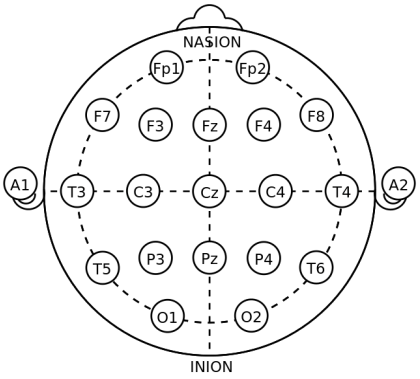


Figure 1. <Electrode locations of International 10-20 system for EEG (electroencephalography) recording>

The International 10-20 system is used to place the nodes on the scalp. The 10-20 system is based on the relationship between the location of an electrode and the underlying area of the cerebral cortex. As shown in Figure 1, each point on the scalp indicates a possible electrode position. Each site has a letter that identifies the lobe and a number or letter to identify the hemisphere location. The letters representing the lobes stands for pre-frontal (Fp), frontal (F), temporal (T), parietal (P), occipital (O), central (C), and mastoid process (A). The electrodes with even numbers (2,4,6,8) refer to the placement on the right side of the head, whereas odd numbers refer to those on the left. The letter 'z' represents the center vertical line of the scalp (Journal of Clinical Neurophysiology, 1991).

Out of the several types of EEG tests such as Routine EEG, Video EEG, Prolonged EEG, Sleep EEG or intracranial EEG, ambulatory EEG (aEEG) are commonly used to predict epileptic seizures in daily life as brain activity can be recorded throughout a few days, up to 72 hours. The electrodes are attached to a small portable EEG recorder, and patients can continue with most of their normal daily activities while the recording is being taken.

Conversion of Time Domain into Frequency Domain by Fourier Transform

In machine learning, data can be represented and analyzed in various domains depending on the nature of the problem and the characteristics of the data. Out of the various domains such as the time, frequency, spatial, image or graph domains, this chapter will discuss the conversion of time domain into the frequency domain using a method called the Fourier Transform, utilized in the proposed model.

In a time domain, signals are represented as a function of time. In this domain, the signal is typically plotted on a time graph with time on the x-axis and signal amplitude on the y-axis. This domain is useful for studying phenomena that are time-dependent, such as transient events or dynamic changes in a signal.

Similarly, in a frequency domain, signals are analyzed based on their frequency content. The frequency domain representation provides information about the different frequency components present in a signal and their respective amplitudes. The resulting frequency graph representation is often displayed on a graph with frequency on the x-axis and signal magnitude (or power) on the y-axis.

EEG signals are given as a one-dimensional time domain, and this is converted to the frequency domain in this research for several reasons. The advantages include noise suppression as certain types of noise, such as high-frequency noise or specific frequency interference, can be easier to identify and remove in the frequency domain. By filtering out abnormal frequencies, the quality of the data and thus the quality of the model will be increased. Moreover, frequency domain transformation helps in the process of feature extraction.

The frequency domain transformation may reveal useful features of the data that can be trained by the machine learning model. These features play a key role in capturing important attributes and help enhance the performance of the model.

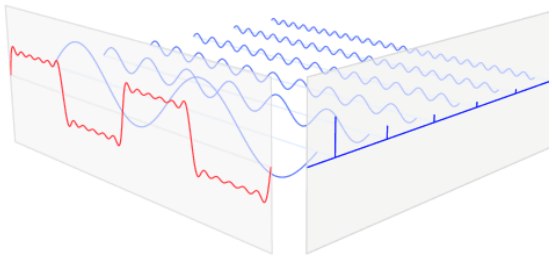


Figure 2. <Example of the Fourier Transform, with the red signal being decomposed into a sum of sinusoidal components, represented by the blue signals>

The conversion from the time domain into the frequency domain is done by a method called the Fourier Transformation, which converts a function into a form that decomposes the original data into a sum of sinusoidal components with different frequencies, outputting a complex-valued function of frequency (Bailey, David H et al., 1994) as shown in Figure 2. The Fourier Transform can be broken down into three steps. First, in the decomposition process, the Fourier Transform breaks down the time-domain signal into its individual frequency components. It then analyzes the amplitudes and phases of these components in order to determine their contributions to the overall signal. Secondly, in the process of outputting the frequency spectrum, the Fourier Transform results in a representation of the original data in the frequency domain, called the frequency spectrum. It contains a set of frequency bins, with each representing the amplitude of a particular frequency component. Lastly, after the signals are converted into the frequency spectrum, the presence and relative strengths of diverse frequency components are revealed, making it significantly easier for the machine learning model to train with the data inputted.

The Fourier Transform will be used to convert the time domain into the frequency domain in the proposed model as a pre-processing step before the model is trained using the data.

Method

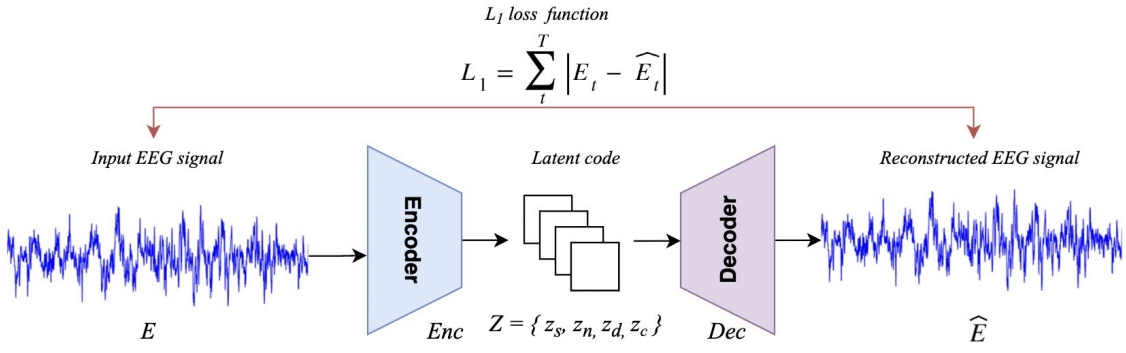


Figure 3. Vanilla Autoencoder

In this chapter, the proposed model will be explained thoroughly. In Chapter 3.1, I will elaborate on how the proposed representation learning method effectively disentangles the seizure-related features.

In Chapter 3.2, the seizure classifier made by transfer learning will be presented. Figure 3 shows the architecture of a Vanilla Autoencoder (Hinton et al., 2006), the most basic form of an autoencoder, applied to this research before a new method is proposed. The Encoder takes in the input as a two-dimensional EEG signal, and the latent code is unentangled. z_s, z_d, z_n, and z_c represent the seizure-related, device-related, noise-related and constant features of the inputted EEG signal. These are called latent features as they all have potential to contribute in finding a meaningful pattern. Finally, the decoder uses disentangled features and reconstructs the EEG signal based on it. The L1 loss function then compares the original inputted EEG signal and reconstructed EEG signal in order to evaluate if the features have been disentangled properly.

The proposed method uses some of this logic, but instead includes a Feature Swapping Mechanism and extracts one necessary feature for training. This will be explained further in Chapters 3.1 and 3.2.

Representation Learning (Autoencoder)

The goal of representation learning is to effectively extract the necessary features from the entangled latent features for the machine learning model to train on. The features extracted from the CNN include a combination of frequency, amplitude, noise, and seizure-related features, which are often entangled.

In order to increase the accuracy of seizure classification, it is important to disentangle only the seizure-related feature. Compared to when the model trains on multiple features including unnecessary noise, only training on important features will enhance the performance of the model.

Thus, the Autoencoder extracts only the necessary, seizure-related features. To disentangle the necessary features, a new feature swapping mechanism will be proposed. This will be further explained in Chapter 3.2.

In order to evaluate if the feature extraction is performing well, a loss function is used. After the optimal features are extracted from the compressed data in the Autoencoder, the Decoder reconstructs the EEG signals using only the extracted, compressed data. If the feature extraction is performing well, the original data and reconstructed data will be similar.

As shown in Equation 1, the L1 loss function is calculated as follows.

Equation 1: L1 loss function

$$L_1 = \sum_{t}^{T} \left| E_t - \widehat{E_t} \right|$$

Here, T denotes the total time of the dataset, and E_t and $\widehat{E_t}$ denotes a point in the original data and reconstructed data, respectively. For each second of the EEG signals, the difference between the original data and reconstructed data is calculated and summed. Theoretically, the L1 loss function can be 0, meaning feature extraction has been done perfectly.

The methods proposed will be evaluated and verified in Chapter 4.

Feature Swapping Mechanism

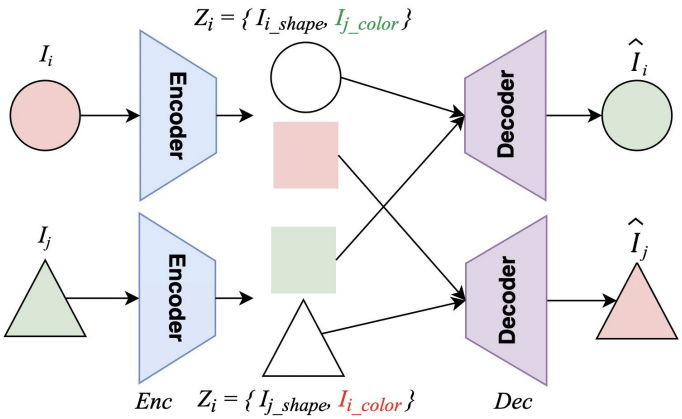


Figure 4. Feature Swapping Mechanism

The ultimate goal of the proposed Feature Swapping Mechanism is to disentangle necessary features from the seizure-related, device-related, noise-related and common features from the inputted Electroencephalogram (EEG) signals. This will enhance the performance of the Seizure Classifier explained in Chapter 3.3.

In Figure 4, the general mechanism of the Feature Swapping Mechanism is presented. Two inputs, I_i and I_j are processed by the Encoder (Enc) which extracts the entangled features. Next, one specific feature is swapped. In this example, the color-related features are swapped. The swapped features are then reconstructed by the Decoder (Dec). The loss function, which compares the outputs, $\hat{I}_i$ and $\hat{I}_j$, with the inputs, will be explained thoroughly in Chapter 3.2.1, 3.2.2 and 3.2.3 for each specific scenario.

EEG Feature Extractor (EFE(.)) is defined as following: ***EFE***: $E \rightarrow Z$, $Z = \{ z_s, z_d, z_n, z_c \}$.
The Decoder (DEC(.)) is defined as following: ***DEC***: $Z \rightarrow \hat{E}$.

The Feature Swapping Mechanism swaps specific elements of the four features. The seizure-related, device-related and noise-related features will each be swapped separately, and the loss of each result will be calculated. As the machine learning model trains, the accuracy of the extractions and swaps will be improved.

The underlying assumption is that when specific features are swapped, the results will change accordingly. For instance, when seizure-related or device-related features are swapped, the reconstructed output of EEG signals will be identical theoretically as these features are common for the two inputted samples of EEG signals. On the other hand, when noise-related features are swapped between one input of an EEG signal with noise and the other without noise, the outputs will also be swapped theoretically.

Unlike the Vanilla Encoder explained in Chapter 3.1, DEC in the proposed Feature Swapping Method takes feature swapped latent code (Z) as an input as following: ***DEC***: $Z_{swap} \rightarrow \hat{E}_{swap}$.
Here, Z_{swap} represents the swapped latent code, and $\hat{E}_{swap}$ represents the outputted reconstructed EEG signals built with the swapped features. In the proposed Feature Swapping Method, three scenarios of swapping features (seizure-related, device-related, noise-related) will be presented.

The effectiveness of the proposed Feature Swapping Mechanism will be evaluated and proved in Chapter 4.

Seizure-Related Feature Swapping

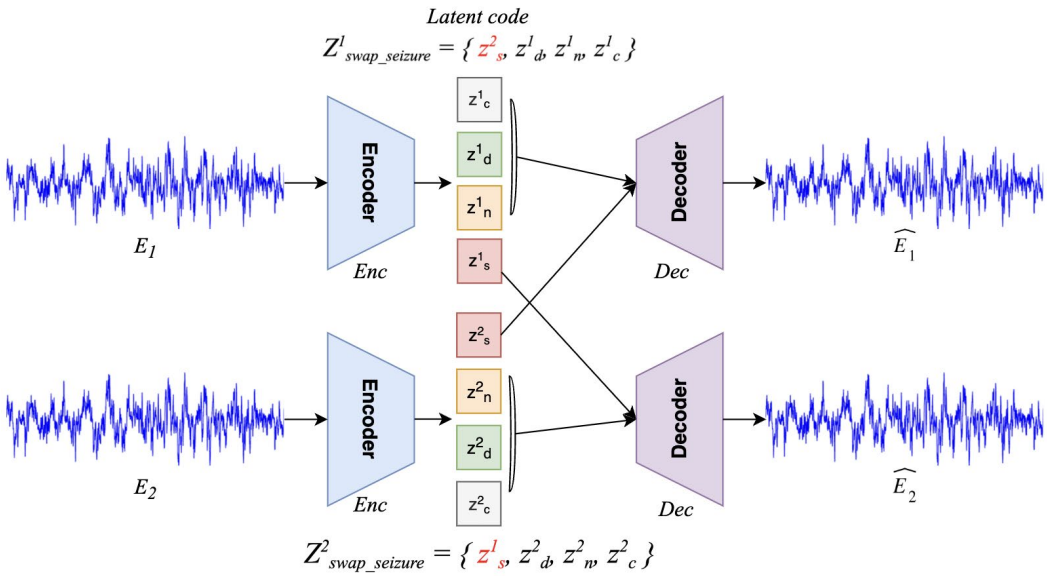


Figure 5. Seizure-related feature swapping

In this chapter, I will elaborate on the comprehensive procedure for disentangling seizure-related features. Figure 5 demonstrates the architecture of the proposed Feature Swapping Mechanism for seizure-related features. The encoder takes two EEG signals as input and extracts the latent features, and the seizure-related features are swapped as shown above in Figure 5. Note that the input EEG signals are of the same seizure category. This is due to the fact that the feature being swapped needs to be consistent. The Decoder then reconstructs the EEG signals based on the swapped features.

As shown in Equation 2, the Feature Swapping Mechanism for seizure-related features is defined as follows.

Equation 2: Feature swapping (seizure-related)

$$\boldsymbol{Z}^1_{swap\_seizure} = \{ z^2_s, z^1_d, z^1_n, z^1_c \}$$
$$\boldsymbol{Z}^2_{swap\_seizure} = \{ z^1_s, z^2_d, z^2_n, z^2_c \}$$

Here, $\boldsymbol{Z}^1_{swap\_seizure}$ and $\boldsymbol{Z}^2_{swap\_seizure}$ represent two inputs of EEG signals of the same seizure category measured with different devices. $z^1_s, z^1_d, z^1_n, z^1_c$ represents the seizure-related, device-related, noise-related and common features for $\boldsymbol{Z}^1_{swap\_seizure}$ respectively. $z^2_s, z^2_d, z^2_n, z^2_c$ represents the seizure-related, device-related, noise-related and common features for $\boldsymbol{Z}^2_{swap\_seizure}$ respectively.

z^1_s, originally an element of $\boldsymbol{Z}^1_{swap\_seizure}$, is swapped with z^2_s, originally an element of $\boldsymbol{Z}^2_{swap\_seizure}$. When the two seizure-related features are swapped, the DEC will reconstruct the changed features into EEG signals. As the assumption is that the seizure-related features for two EEG signals of the same seizure category are identical, the output of each is optimal when it is identical with its inputs. Thus, the loss function should be calculated between $\boldsymbol{Z}^1$ and $\boldsymbol{Z}^1_{swap\_seizure}$ after it has been processed by the DEC, as well as $\boldsymbol{Z}^2$ with $\boldsymbol{Z}^2_{swap\_seizure}$ after it has been processed by the DEC. The loss functions will inform us if the seizure-related features are being extracted accurately. Equation 3 shows how the proposed $\boldsymbol{L}_{seizure}$ loss is calculated.

Equation 3: $L_{seizure}$ loss function

$$L_{seizure} = \sum_{t}^{T} \left| E_t - \widehat{E_t} \right|$$

Here, T denotes the number of samples in the inputted EEG signal, and E_t and $\widehat{E_t}$ denotes the inputted EEG signal and its reconstructed EEG signals based on the swapped features, respectively. Similarly to the Equation 1, the proposed $L_{seizure}$ function measures the element-wise difference between the inputted and reconstructed EEG signals.

Device-Related Feature Swapping

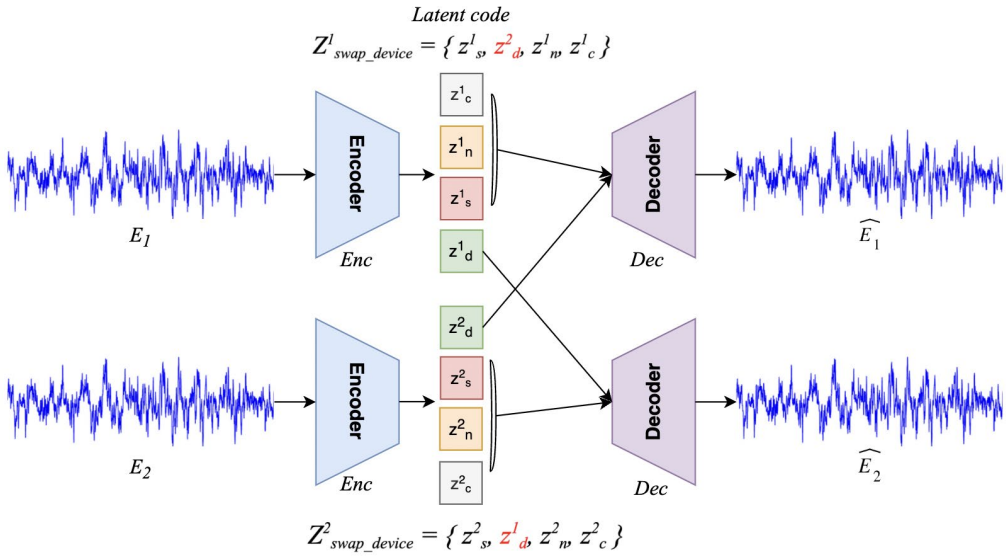


Figure 6. Device-related feature swapping

Similarly to Chapter 3.2.1, this chapter elaborates on the comprehensive procedure for disentangling device-related features. Figure 6 demonstrates the architecture of the proposed Feature Swapping Mechanism for device-related features. The mechanism is similar to the Feature Swapping Mechanism for seizure-related features. However, note that the input EEG signals need to be samples measured from the same device instead of seizure categories. This is due to the fact that the feature being swapped needs to be consistent.

As shown in Equation 4, the Feature Swapping Mechanism for device-related features is defined as follows.

Equation 4: Feature Swapping Mechanism (device-related)

$$\boldsymbol{Z}^1_{swap\_device} = \{ z^1_s, z^2_d, z^1_n, z^1_c \}$$
$$\boldsymbol{Z}^2_{swap\_device} = \{ z^2_s, z^1_d, z^2_n, z^2_c \}$$

Here, $\boldsymbol{Z}^1_{swap\_seizure}$ and $\boldsymbol{Z}^2_{swap\_seizure}$ represent two inputs of EEG signals of the same seizure category measured with different devices. $z^1_s, z^1_d, z^1_n, z^1_c$ represents the seizure-related, device-related, noise-related and common features for $\boldsymbol{Z}^1_{swap\_seizure}$ respectively. $z^2_s, z^2_d, z^2_n, z^2_c$ represents the seizure-related, device-related, noise-related and common features for $\boldsymbol{Z}^2_{swap\_seizure}$ respectively.

z^1_d, originally an element of $\boldsymbol{Z}^1_{swap\_seizure}$, is swapped with z^2_d, originally an element of $\boldsymbol{Z}^2_{swap\_seizure}$. When the two device-related features are swapped, the DEC will reconstruct the changed features into EEG signals. As the assumption is that the device-related features for two EEG signals measured with the same device are identical, the output of each is optimal when it is identical with its inputs. Thus, the loss function should be calculated between $\boldsymbol{Z}^1$ and $\boldsymbol{Z}^1_{swap\_seizure}$ after it has been processed by the DEC, as well as $\boldsymbol{Z}^2$ with $\boldsymbol{Z}^2_{swap\_seizure}$ after it has been processed by the DEC. The loss functions will inform us if the seizure-related features are being extracted accurately. Equation 5 shows how the proposed $\boldsymbol{L}_{seizure}$ loss is calculated.

Equation 5: L_{device} loss function

$$L_{device} = \sum_{t}^{T} \left| E_t - \widehat{E_t} \right|$$

Here, T denotes the number of samples in the inputted EEG signal, and E_t and $\widehat{E_t}$ denotes the inputted EEG signal and its reconstructed EEG signals based on the swapped features, respectively. Similarly to the Equation 1, the proposed L_{device} function measures the element-wise difference between the inputted and reconstructed EEG signals.

Noise-Related Feature Swapping

In this chapter, the comprehensive procedure for disentangling noise-related features is elaborated. Figure 7 demonstrates the architecture of the proposed Feature Swapping Mechanism for noise-related features. Note that there are differences compared to Chapter 3.2.1 and 3.2.2, which were quite alike. The Feature Swapping Mechanism for noise-related features takes two inputs: a sample of EEG signal and the same EEG signal sample but with its noise suppressed. The wavelet-inverse wavelet transformation will be used to suppress the noise of the EEG signals.

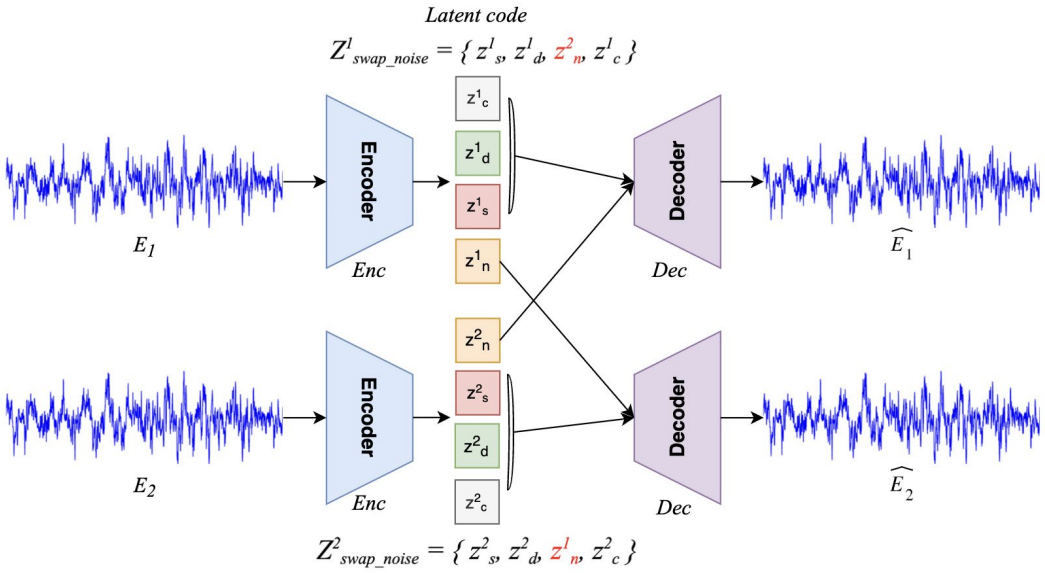


Figure 7. Device-related feature swapping

For example, E_1 is the original EEG signal and E_2 is the EEG signal generated from E_1 with its noise suppressed through the wavelet - inverse wavelet transformation in Figure 7.

The outputs should theoretically be swapped as well. $\hat{E}_1$ should be identical to E_2, and $\hat{E}_2$ should be identical to E_1. This is because the features swapped are different. The noise-related features were swapped, and one had noise-related features while the other had suppressed noise-related features. Accordingly, the loss function should be calculated between $\hat{E}_1$ with E_2, and $\hat{E}_2$ with E_1.

As shown in Equation 6, the Feature Swapping Mechanism for noise-related features is defined as follows.

Equation 6: Feature Swapping Mechanism (noise-related)

$$\boldsymbol{Z}^1_{swap\_noise} = \{ z^1_s, z^1_d, z^2_n, z^1_c \}$$
$$\boldsymbol{Z}^2_{swap\_noise} = \{ z^2_s, z^2_d, z^1_n, z^2_c \}$$

Here, $\boldsymbol{Z}^1_{swap\_seizure}$ and $\boldsymbol{Z}^2_{swap\_seizure}$ represent two inputs of EEG signals, with one sample of EEG signal and the other the same signal but with the noise suppressed. $z^1_s, z^1_d, z^1_n, z^1_c$ represents the seizure-related, device-related, noise-related and common features for $\boldsymbol{Z}^1_{swap\_seizure}$ respectively. $z^2_s, z^2_d, z^2_n, z^2_c$ represents the seizure-related, device-related, noise-related and common features for $\boldsymbol{Z}^2_{swap\_seizure}$ respectively.

z^1_n, originally an element of $\boldsymbol{Z}^1_{swap\_seizure}$, is swapped with z^2_n, originally an element of $\boldsymbol{Z}^2_{swap\_seizure}$. When the two device-related features are swapped, the DEC will reconstruct the changed features into EEG signals. As the assumption is that the noise-related features are different for the two EEG signals, with one unsuppressed and the other suppressed, the outputs should be identical with the other signal instead of its original signal after being processed by the DEC. Thus, the loss function should be calculated between $\boldsymbol{Z}^1$ and $\boldsymbol{Z}^2_{swap\_seizure}$ after it has been processed by the DEC, as well as $\boldsymbol{Z}^2$ with $\boldsymbol{Z}^1_{swap\_seizure}$. The loss functions will inform us if the seizure-related features are being extracted accurately. Equation 7 shows how the proposed $\boldsymbol{L}_{noise}$ loss is calculated.

Equation 7: L_{noise} loss function

$$L_{noise} = \sum_{t}^{T} \left| E_t - \widehat{E_t} \right|$$

Here, T denotes the number of samples in the inputted EEG signal, and E_t and $\widehat{E_t}$ denotes the inputted EEG signal and its reconstructed EEG signals based on the swapped features, respectively.

In conclusion, the total loss function is a weighted linear combination of the three previously mentioned loss functions, as shown in Equation 8.

Equation 8: L_{total} loss function

$$\boldsymbol{L} = \boldsymbol{L}_{seizure} + \alpha * L_{device} + \beta * L_{noise}$$

Here, α and β are variables that assign weights to each loss function. Through extensive experiments, I have discovered that setting α to 0.8 and β to 0.7 produces the most optimal results.

Transfer Learning (Seizure Classifier)

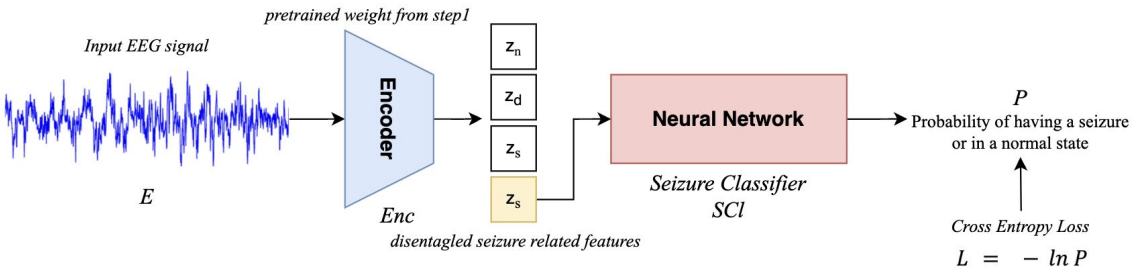


Figure 8. <Transfer Learning (Seizure Classifier)>

Another method proposed in this research is transfer learning, as shown in Figure 8. Transfer learning is a machine learning technique where knowledge gained from solving one problem is applied to a different but related problem. The ultimate goal is to build an accurate seizure classification network.

The Seizure Classifier will receive a one dimensional input of EEG signals, which the pre-trained Convolutional Neural Network (CNN) from Chapter 3.1 will extract the seizure related features from. The neural network trains using these features, outputting the probability of the input will lead to a seizure.

The features are disentangled when processed through the encoder (ENC) as following: *ENC:* $E \rightarrow Z$. The features will be separated into Z_s, Z_d, Z_n, Z_c, seizure-related, device-related, noise-related and common features, respectively.

The extracted features are finally processed into a probability as follows: *SCL:* $z_s \rightarrow P_{seizure}$. After the features are extracted, Z_s will be extracted. Z_s, the seizure-related features, are then taken as an input and processed by the Seizure Classifier (SCL). SCL is a neural network that takes the seizure-related features as an input and outputs the probability, $P_{seizure}$, whether a seizure is about to take place or not.

In order to evaluate the performance of the SCL, the Cross-Entropy loss function is used. It measures the difference between predicted probabilities and the true labels of a given dataset.

Before the Cross-Entropy loss is computed, an activation function called the Softmax function is applied to the scores. As shown in Equation 9, it normalizes the raw scores to probabilities between 0 and 1.

Equation 9: Softmax function

$$S(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$$

Here, x_i denotes the ith raw score. The equation applies exponential functions, an increasing function only including positive numbers in order to change the scores into positive numbers without changing the strict inequality of them since probabilities can not be negative. Then, the scores are normalized into probabilities that sum to 1 by dividing each score by the sum of all the scores.

After the raw scores are converted into probabilities, the Cross-Entropy loss (L_{CE}) is computed as shown in Equation 10 below.

Equation 10: Cross-Entropy (L_{CE}) Loss

$$L_{CE} = -\sum_{i=1}^{output\ size} y_i \times \log \hat{y}_i$$

Here, y_i and $\hat{y}_i$ denotes the true output and predicted output, respectively. The log function is utilized because log values increase rapidly as the predicted probability approaches 0, the wrong prediction. This helps increase accuracy as the loss will be much greater as the prediction reaches a completely wrong value.

Implementation Details

I utilized the ResNet-50 (He et al. 2016) architecture as the encoder, and I modified the downsampling layers with upsampling layers to develop the decoder. For training, I used Adam optimizer with an initial learning rate set to 0.0001 over the course of 120 epochs. Additionally, I implemented learning rate reductions by a factor of 0.1 at the 60th and 100th epochs. In transfer learning, I constructed the Seizure Classifier using a two-layer neural network and proceeded to train it for 40 epochs, maintaining a constant learning rate of 0.0001 without any decay. It was observed that this transfer learning approach rapidly converges towards the optimal point, attributable to the enhanced representations during the process of the proposed feature swapping mechanism based representation learning.

Experimental Results

Dataset

In this chapter, I will provide a comprehensive description of the two datasets used in the research, the NICU dataset (Stevenson et al. 2019) and the Epileptic Seizure Recognition dataset (Wu et al. 2017).

The first dataset used is the NICU dataset, a dataset of neonatal electroencephalography (EEG) recordings with seizure annotations. It consists of 46,640 samples of EEG signals in total collected from 79 individuals. Out of the samples, 17,940 were collected during seizures (positive), and 28,700 were collected during normal states (negative).

The second dataset used is the Epileptic Seizure Recognition dataset from the UC Irvine Machine Learning Repository. It contains 11,500 EEG signal samples collected from 500 individuals. 2,300 samples were collected during seizures (positive), and 9,200 were collected from normal states (negative).

Both datasets will be used for within-dataset evaluation and cross-dataset evaluation.The two datasets will be split into a total of 5 folds, with a 8:2 ratio of training data and test data respectively for the 5-fold cross validation.

Protocol

In this research, various evaluations are conducted to verify the effectiveness of the proposed approach. Comparison with state-of-the-art (SOTA) methods, including within dataset and cross dataset tests, will allow the comparison between the efficiency of the proposed method compared to the existing ones. The cross dataset test especially will prove how the proposed method solves the bias problem on datasets and Electroencephalogram (EEG) devices. Both comparisons with SOTA evaluation will be conducted through the 5-fold validation. These evaluations will be explained further in Chapter 4.3.

The evaluations will be organized into evaluation metrics. Each column represents Accuracy, Recall, Precision and the F1 score. Before moving on, the confusion matrix needs to be described. The confusion matrix consists of four types: True Positive, the correctly predicted positive results, True Negative, the correctly predicted negative results, False Positive, when the correct prediction is negative but was incorrectly predicted to positive, and False Negative, when the correct prediction is positive but was incorrectly predicted to positive. In the medical field, the False Negative value should be especially lower as this means no predictions were not made.

First of all, as shown in Equation 11, Accuracy is the number of correctly predicted data, True Positive and True Negative, divided by the total number of data.

Equation 11: Accuracy

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + False\ Negative}$$

Secondly, as shown in Equation 12, Recall is the number of data that the model recognizes as positive, True Positive, when the data is actually positive, True Positive and False Negative over the number of total data.

Equation 12: Recall

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

Thirdly, as shown in Equation 13, Precision is the number of data that the model predicts to be positive, True Positive, that is actually positive, True Positive and False Positive.

Equation 13: Precision

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Lastly, as shown in Equation 14, the F1 Score is an evaluation metric that measures a model's accuracy in general. It is the harmonic average of precision and recall. The harmonic mean is calculated instead of the normal mean because the closer the precision and recall are to zero, the lower the F1 score should be.

Equation 14: F1 Score

$$F1\ Score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

After the tests are conducted and evaluated, the feature maps will be visualized through the t-sne visualization, a statistical method for visualizing high-dimensional data by compressing it to dots on a two-dimensional graph. This method will prove if the latent features are extracted properly. The visualization of the feature maps will be elaborated in Chapter 4.4.

For the comparison methods, 5-fold validations were chosen. The ultimate goal when comparing the proposed framework with state-of-the-art (SOTA) methods is to prove the enhanced accuracy of the model and to show how the proposed framework, unlike the existing ones, will not be dependent on specific datasets and devices. As shown in Figure 9, 5-fold validation is when all data is randomly split into 5 folds, and the model is trained on the 4 folds, while one fold is left to test the model. This procedure is repeated 5 times.

Comparison with SOTA

Within Dataset Evaluation

The first evaluation conducted is the 5-fold validation within the dataset. The goal of this evaluation is to prove the enhanced accuracy of the proposed model. The two datasets, NICU dataset and Epileptic Seizure Recognition dataset are used. As explained in Chapter 4.2, 5-fold validation is when all data is randomly split into 5 folds, and the model is trained on the 4 folds, while one fold is left to test the model. This procedure is repeated 5 times.

Table 2. Performance comparison on NICU dataset (within dataset)

	Accuracy	Recall	Precision	F1-score
VGG19 (Simonyan et al. 2014)	0.8417 (±0.0004)	0.8611 (±0.0010)	0.7785 (±0.0005)	0.7689 (±0.0012)
MobileNetV2 (Sandler et al. 2018)	0.8692 (±0.0005)	0.8918 (±0.0016)	0.7894 (±0.0009)	0.7907 (±0.0011)
DenseNet121 (Huang et al. (2017)	0.9144 (±0.0009)	0.9589 (±0.0011)	0.8312 (±0.0008)	0.8928 (±0.0007)
HRNet32 (Sun et al. 2019)	0.9006 (±0.0013)	0.9481 (±0.0007)	0.8208 (±0.0004)	0.8794 (±0.0010)
Proposed model	0.9053 (±0.0011)	0.9508 (±0.0008)	0.8284 (±0.0006)	0.8854 (±0.0009)

Table 3. Performance comparison on Epileptic Seizure Recognition dataset (within dataset)

	Accuracy	Recall	Precision	F1-score
VGG19 (Simonyan et al. 2014)	0.8804 (±0.0006)	0.8903 (±0.0010)	0.8774 (±0.0008)	0.8915 (±0.0007)
MobileNetV2 (Sandler et al. 2018)	0.8851 (±0.0004)	0.8887 (±0.0008)	0.8680 (±0.0009)	0.8779 (±0.0006)
DenseNet121 (Huang et al. (2017)	0.9465 (±0.0007)	0.9662 (±0.0008)	0.9307 (±0.0004)	0.9551 (±0.0006)

HRNet32 (Sun et al. 2019)	0.9352 (±0.0011)	0.9477 (±0.0009)	0.9187 (±0.0006)	0.9405 (±0.0010)
Proposed model	0.9488 (±0.0006)	0.9504 (±0.0007)	0.9215 (±0.0010)	0.9513 (±0.0009)

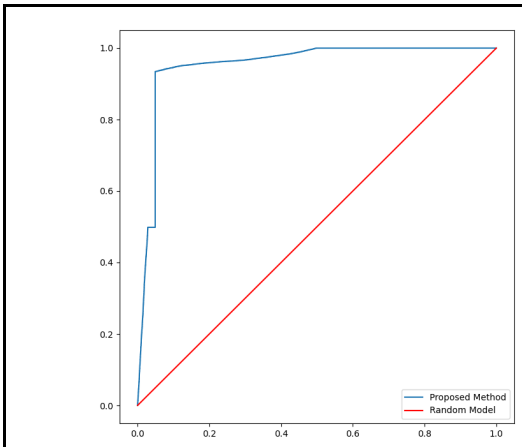
Table 2 and Table 3 shows the performance of five models including the proposed model on the within dataset evaluation on the NICU dataset and the Epileptic Seizure Recognition dataset, respectively.

The proposed method outperforms the previous models by an accuracy of 0.0238 (2.38%) and 0.0370 (3.70%) when using the NICU dataset and Epileptic Seizure Recognition dataset respectively. In total, the proposed method has a 0.0304 (3.04%) increased accuracy. Although the difference is not great, it still outperforms the four models on average.

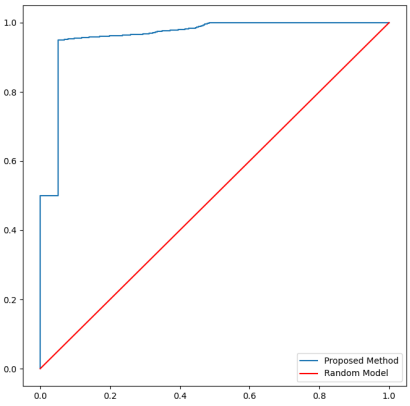
There are several noticeable points from the results. The first observable point is that the performance of the VGG19 (Simonyan et al. 2014) and MobileNetV2 (Sandler et al. 2018) are remarkably low. This is due to the fact that the two models have comparably shallow networks.

The second noticeable point is that the DenseNet121 (Huang et al. (2017) and HRNet32 (Sun et al. 2019) have comparable performance as the two models consist of deeper layers. However, as these two models train on a supervised approach, it tends to be biased on the dataset used during training. Consequently, their performance experiences a significant performance drop during cross dataset evaluation. A more detailed explanation will be given in Chapter 4.3.2.

Another noticeable point is that the proposed method shows comparably high performance in the within dataset evaluation, as well as the cross dataset evaluation as it is not biased on the certain dataset. The proposed method effectively disentangles seizure-related features without exhibiting bias
towards specific datasets. I attribute this generalization ability of the proposed method to the feature swapping mechanism based representation learning. This will also be explained thoroughly in Chapter 4.3.2.



(a)



(b)

Figure 9. ROC curve
(a): ROC curve of the within-dataset evaluation using the NICU dataset
(b): ROC curve of the within-dataset evaluation using the Epileptic Seizure Recognition dataset

An ROC curve (receiver operating characteristic curve) is a graph that shows the performance of a classification model at all classification thresholds. The curve plots two parameters: True Positive Rate as the y-axis and False Positive Rate as the x-axis. In other words, it depicts the trade-off between sensitivity and specificity. The red line represents a random model, and the blue line the proposed method.

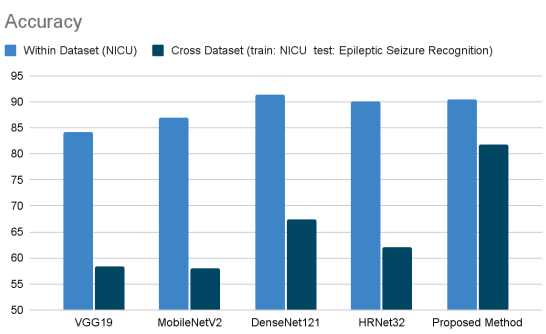
When the AUC (Area Under the ROC Curve) is greater than 0.5 and closer to 1, it has a higher chance that the model can classify positive and negative values. The graphs in Figure 9 (a) and (b) are the ROC curves for the within-dataset evaluation using the NICU dataset and Epileptic Seizure Recognition dataset respectively. They both have AUC comparably close to 1, meaning that the two models have satisfactory performance in terms of sensitivity and specificity.

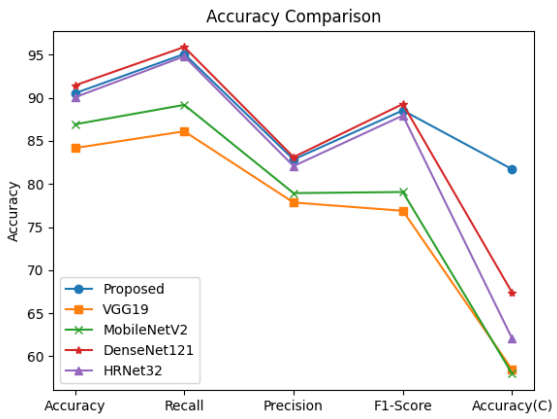
Cross Dataset Evaluation

The second evaluation conducted is the cross dataset 5-fold validation test. The goal of this evaluation is to prove that the proposed model is unbiased to any datasets or Electroencephalogram (EEG) devices. Two datasets, NICU and Epileptic Seizure Recognition datasets are used. Similarly to the previous evaluation, the two datasets are both randomly split into 5 folds. Unlike the within dataset evaluation which is trained and tested using the same dataset, the cross dataset evaluation trains with the 4 folds of one of the two datasets, then is tested using one fold of the other dataset. This procedure is repeated 5 times.

Table 4. Performance comparison on model trained on the NICU dataset and tested on the Epileptic Seizure Recognition Dataset (cross-dataset)

	Accuracy	Recall	Precision	F1-score
VGG19 (Simonyan et al. 2014)	0.5846 (±0.0006)	0.5894 (±0.0004)	0.5785 (±0.0006)	0.5881 (±0.0005)
MobileNetV2 (Sandler et al. 2018)	0.5798 (±0.0007)	0.5901 (±0.0006)	0.5562 (±0.0005)	0.5711 (±0.0008)
DenseNet121 (Huang et al. (2017)	0.6739 (±0.0010)	0.6884 (±0.0011)	0.6585 (±0.0011)	0.6701 (±0.0009)
HRNet32 (Sun et al. 2019)	0.6207 (±0.0012)	0.6511 (±0.0009)	0.6067 (±0.0012)	0.6294 (±0.0010)
Proposed model	0.8171 (±0.0008)	0.8588 (±0.0009)	0.7892 (±0.0007)	0.8357 (±0.0005)





(a) **(b)**

Figure 10. Performance evaluation graphs comparing state-of-the-art methods and models trained on the NICU dataset
(a): Bar graph on accuracy of within-dataset test and cross-dataset test
(b): Line graph on evaluation of within-dataset test and cross-dataset test

Table 4 demonstrates the evaluation results for the cross dataset 5-fold validation test. The NICU dataset is used to train the model, and the Epileptic Seizure Recognition dataset is used for testing . The accuracy, recall, precision and f1-score of the proposed method is also higher than most of the values of the state-of-the-art methods.

When the accuracy of the within-dataset tests on the previous methods and proposed method are compared, they are fairly similar. As explained in DenseNet121 (Huang et al. (2017) and HRNet32 (Sun et al. 2019) show comparable accuracy due to the fact that they both consist of deeper layers. The proposed method, which also has a deep layer, shows high accuracy of 0.9053 (±0.0006).

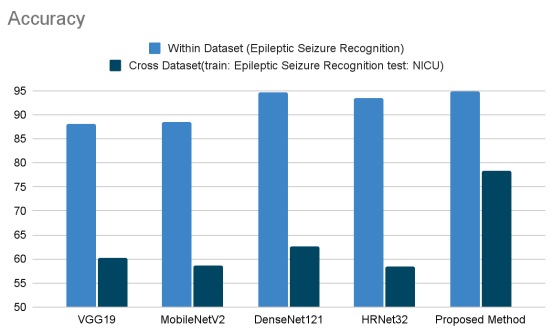
However, when the accuracy of the cross-dataset tests on the models are compared, the proposed method has impressively high accuracy. As shown in Figure 10 (b), in the last column, which compares the accuracy of the cross-dataset evaluation results between the five models, the proposed method shows significantly high accuracy. This can be observed in Figure 10 (a) as well.

The two models that showed high accuracy during the within-dataset evaluation, DenseNet121 and HRNet 32, train on a supervised approach, tending to be biased on the dataset used during training. Consequently, their performance experiences a significant performance drop during cross dataset evaluation. On average, the proposed method performs 0.2024 (20.24%) better than the previous models.

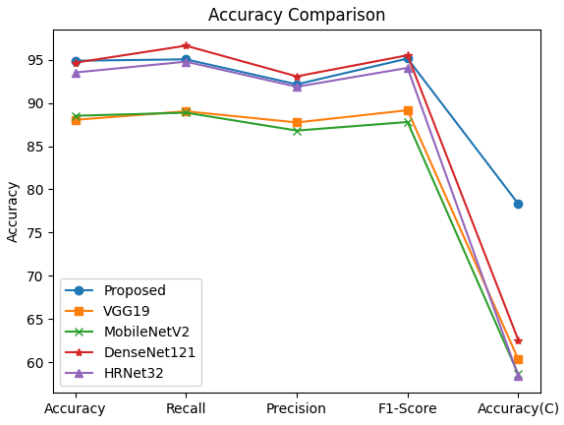
In other words, the proposed model is not biased on a certain dataset. As the feature swapping mechanism based representation learning disentangles and extracts the necessary features effectively, the generalization ability is higher than previous models that do not disentangle features. The disentanglement of features will be evaluated further in Chapter 4.4. When features are not extracted before used for training, the model trains with unnecessary information as well. This causes the previous models to be biased on certain datasets. On the other hand, the proposed method only uses the extracted seizure-related features for training. Thus, the model performs effectively on other datasets.

Table 5. Performance comparison on model trained on the Epileptic Seizure Recognition Dataset dataset and tested on the NICU dataset (cross-dataset)

	Accuracy	Recall	Precision	F1-score
VGG19 (Simonyan et al. 2014)	0.6033 (±0.0007)	0.6242 (±0.0008)	0.5798 (±0.0011)	0.5873 (±0.0012)
MobileNetV2 (Sandler et al. 2018)	0.5860 (±0.0004)	0.5911 (±0.0006)	0.5718 (±0.0007)	0.5850 (±0.0010)
DenseNet121 (Huang et al. (2017)	0.6256 (±0.0006)	0.6447 (±0.0008)	0.5918 (±0.0008)	0.5994 (±0.0005)
HRNet32 (Sun et al. 2019)	0.5838 (±0.0009)	0.6218 (±0.0005)	0.5597 (±0.0014)	0.5784 (±0.0012)
Proposed model	0.7832 (±0.0011)	0.8041 (±0.0007)	0.7706 (±0.0009)	0.7881 (±0.0010)



(a)



(b)

Figure 11. Performance evaluation graphs comparing state-of-the-art methods and models trained on the Epileptic Seizure Recognition Dataset
(a): Bar graph on accuracy of within-dataset test and cross-dataset test
(b): Line graph on evaluation of within-dataset test and cross-dataset test

Similarly to Table 4, Table 5 demonstrates the evaluation results for the cross dataset 5-fold validation test on the Epileptic Seizure Recognition Dataset. The accuracy, recall, precision and f1-score of the proposed method is also higher than most of the values of the state-of-the-art methods.

The results are analogous to the tests based on models trained on the NICU dataset. When the accuracy of the within-dataset tests on the previous methods and proposed methods are compared, the results of the DenseNet121, HRNet32 and the proposed method all show high performance since they consist of deep layers. The proposed model has an accuracy of 0.9488 (±0.0006).

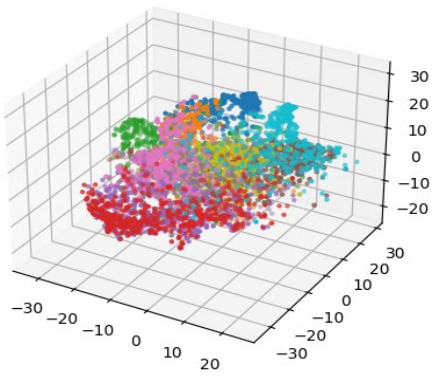
However, when the accuracy of the cross-dataset tests on the models are compared, the proposed method has significantly high accuracy. This can be observed in Figure 11 (a) and (b). As the previous models train on a supervised approach, they all tend to be biased on the dataset used during training. Consequently, their performance experiences a significant performance drop during cross dataset evaluation. On average, the proposed method performs 0.1835 (18.35%) better than the previous models.

Unlike the previous methods, the proposed model is not biased on a certain dataset as it has high generalization ability due to the feature swapping mechanism based representation learning that extracts the necessary features effectively as explained previously. Thus, the proposed model performs effectively in all datasets.

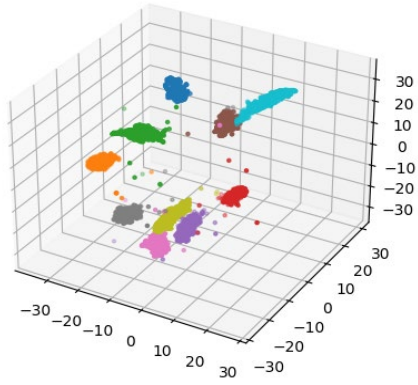
One difference is that the drop between the accuracy of the within-dataset evaluation and cross-dataset evaluation of the proposed method is slightly greater when trained on the Epileptic Seizure Recognition Dataset compared to when the model is trained on the NICU dataset. The reason for this is that the Epileptic Seizure Recognition Dataset has a smaller number of sample data than the NICU dataset. As a wide range of dataset distribution is not provided for the model to train on, the performance will decrease slightly as an effect.

In brief, the within-dataset and cross-dataset evaluations were conducted using the NICU dataset and Epileptic Seizure Recognition Dataset, and the proposed method performed better than the state-of-the-art methods, especially in the cross-dataset evaluation. The differences in accuracy is as follows: the proposed model outperformed the previous methods by 0.0304 (3.04%) during the within-dataset evaluation, and 0.1935 (19.35%) in the cross-dataset evaluation. The considerable difference in accuracy in the cross-dataset evaluation shows that the proposed model is unbiased on the datasets used for training unlike the previous methods. Thus, the problem mentioned in Chapter 1.2 that the models are biased on devices and datasets has been solved and proved with the proposed model.

Feature Maps Visualization



(a)



(b)

Figure 12. t-SNE (Van der Maaten et al. 2008) visualization of feature maps
(a): t-SNE visualization of the original data (NICU dataset)
(b): t-SNE visualization after the features have been disentangled by the Autoencoder

The purpose of the visualization of the feature maps is to evaluate if the latent features are extracted properly through the proposed Encoder (ENC) network. The t-sne (Van der Maaten et al. 2008) visualization is used, which projects high-dimensional feature vectors to low-dimensional plots so that the mutual distance relationships between the points in the initial high-dimensional space are captured in the low-dimensional space. In other words, if two points were close in the initial high-dimensional space, they stay close in the resulting graph, but should stay far apart if the points were far from each other in the original space.

In the visualization, the data of one person from the dataset will be plotted as a dot using the same color. This will allow the evaluation if the features from the same person are grouped well. As shown in Figure 12 (a), when the original data from the NICU dataset is visualized, the dots with different colors are all mixed with each other. However, as shown in Figure 12 (b), where the feature map is visualized after the features have been disentangled, the dots with the same colors are clustered in groups. This clearly shows that the data points of the same category have been grouped together correctly in the Autoencoder. Thus, the t-SNE visualization shows that the disentanglement of features are working properly, proving that the method extracting necessary features will also be effective.

Conclusion

In this paper, I proposed a semi-supervised based domain generalization method to develop an accurate seizure prediction system. The proposed method consists of two phases: the representation learning and the transfer learning phase. The Feature Swapping Mechanism was implemented in the representation learning phase to disentangle and extract seizure-related features, enhancing the accuracy of the model. In the transfer learning phase, pre-trained networks and neural networks were utilized, outputting the probability there will be a seizure in the inputted electroencephalogram (EEG) signal.

Overall, the experimental results demonstrated that the proposed method outperformed the state-of-the-art methods in the within dataset 5-fold evaluation by an average of 0.0304 (3.04%). It outperformed the state-of-the-art methods in the cross dataset 5-fold evaluation by 0.1935 (19.35%) as well. This proves how the several problems detected from previous methods such as the sensitivity of the model to noise and the bias on specific devices and datasets were solved.

The results highlight the feasibility of applying the proposed method to real-life situations when predicting epileptic seizures. For the future, I intend to create a portable, lightweight EEG device that can be worn in daily life, and alerts the patients and healthcare before a seizure occurs through mobile devices.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

Reference

Ahmad, I., Wang, X., Zhu, M., Wang, C., Pi, Y., Khan, J., Khan, S., Samuel, O. W., Chen, S., & Li, G. (2022). EEG-Based Epileptic Seizure Detection via Machine/Deep Learning Approaches: A

Systematic Review. *Computational Intelligence and Neuroscience*, *2022*, 1–20. https://doi.org/10.1155/2022/6486570
Bailey, D. H., & Swarztrauber, P. N. (1994). A fast method for the numerical evaluation of continuous fourier and laplace transforms. *SIAM Journal on Scientific Computing*, *15*(5), 1105–1110. https://doi.org/10.1137/0915067
CHB-MIT Scalp EEG Database V1.0.0. (2010, June 9). https://physionet.org/content/chbmit/1.0.0/
Downloads. (n.d.). UKB | Epileptologie. https://www.ukbonn.de/epileptologie/arbeitsgruppen/ag-lehnertz-neurophysik/downloads/
EEG Database — Seizure Prediction Project Freiburg. (n.d.). https://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database
Electroencephalogram (EEG). (2021, August 8). Johns Hopkins Medicine. https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/electroencephalogram-eeg
He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
Huang, G. (2016, August 25). *Densely connected convolutional networks*. arXiv.org. https://arxiv.org/abs/1608.06993
Ikeda, Y., Fusauchi, T., & Samura, T. (2021). Balance between performance and robustness of recurrent neural networks brought by Brain-Inspired constraints on initial structure. In *Lecture Notes in Computer Science*. https://doi.org/10.1007/978-3-030-92270-2_15
MobileNetV2: Inverted residuals and linear bottlenecks. (2018, June 1). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/8578572
&Na, N. (1991). American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature. *Journal of Clinical Neurophysiology*, *8*(2), 200–202. https://doi.org/10.1097/00004691-199104000-00007
Rasheed, K. (2020). *Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A Review*.
San-Segundo, R., Gil-Martín, M., D'Haro, L. F., & Pardo, J. M. (2019). Classification of epileptic EEG recordings using signal transforms and convolutional neural networks. *Computers in Biology and Medicine*, *109*, 148–158. https://doi.org/10.1016/j.compbiomed.2019.04.031
Seizures - Symptoms and causes - Mayo Clinic. (2023, February 4). Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/seizure/symptoms-causes/syc-20365711
Shoeibi, A., Khodatars, M., Ghassemi, N., Jafari, M., Moridian, P., Alizadehsani, R., Panahiazar, M., Khozeimeh, F., Zare, A., Hosseini-Nejad, H., Khosravi, A., Atiya, A. F., Aminshahidi, D., Hussain, S., Rouhani, M., Nahavandi, S., & Acharya, U. R. (2021a). Epileptic Seizures Detection Using Deep Learning Techniques: A review. *International Journal of Environmental Research and Public Health*, *18*(11), 5780. https://doi.org/10.3390/ijerph18115780
Shoeibi, A., Khodatars, M., Ghassemi, N., Jafari, M., Moridian, P., Alizadehsani, R., Panahiazar, M., Khozeimeh, F., Zare, A., Hosseini-Nejad, H., Khosravi, A., Atiya, A. F., Aminshahidi, D., Hussain, S., Rouhani, M., Nahavandi, S., & Acharya, U. R. (2021b). Epileptic Seizures Detection Using Deep Learning Techniques: A review. *International Journal of Environmental Research and Public Health*, *18*(11), 5780. https://doi.org/10.3390/ijerph18115780
Shoeibi, A., Khodatars, M., Ghassemi, N., Jafari, M., Moridian, P., Alizadehsani, R., Panahiazar, M., Khozeimeh, F., Zare, A., Hosseini-Nejad, H., Khosravi, A., Atiya, A. F., Aminshahidi, D., Hussain, S., Rouhani, M., Nahavandi, S., & Acharya, U. R. (2021c). Epileptic Seizures Detection Using Deep Learning Techniques: A review. *International Journal of Environmental Research and Public Health*, *18*(11), 5780. https://doi.org/10.3390/ijerph18115780

Simonyan, K. (2014, September 4). *Very deep convolutional networks for Large-Scale image recognition*. arXiv.org. https://arxiv.org/abs/1409.1556
Stevenson, N. J., Tapani, K., Lauronen, L., & Vanhatalo, S. (2019). A dataset of neonatal EEG recordings with seizure annotations. *Scientific Data*, *6*(1). https://doi.org/10.1038/sdata.2019.39
Sun, K. (2019, February 25). *Deep High-Resolution Representation learning for human pose estimation*. arXiv.org. https://arxiv.org/abs/1902.09212
Tatum, W. O. (2008). *Handbook of EEG Interpretation*. Demos.
Türk, Ö., & Özerdem, M. S. (2019). Epilepsy Detection by Using Scalogram Based Convolutional Neural Network from EEG Signals. *Brain Sciences*, *9*(5), 115. https://doi.org/10.3390/brainsci9050115
UCI Machine Learning Repository. (n.d.). https://archive.ics.uci.edu/dataset/388/epileptic+seizure+recognition
Website, N. (2023, May 17). *Electroencephalogram (EEG)*. nhs.uk. https://www.nhs.uk/conditions/electroencephalogram/
World Health Organization: WHO. (2023). Epilepsy. *www.who.int*. https://www.who.int/news-room/fact-sheets/detail/epilepsy
Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. Journal of machine learning research, 9(11).