Potential benefits of explicitly preserving the locations of the electrodes in the structure of EEG data for EEG classification

Yige Yang

Master of Science
School of Informatics
University of Edinburgh
2023

Abstract

The electroencephalogram (EEG) is a non-invasive way of assessing brain activity that records electrical impulses produced by neurons. EEG data categorization is critical in cognitive neuroscience, clinical diagnostics, and brain-computer interfaces (BCI), which allow humans to control equipment with their thoughts. Before sending raw EEG data into classification algorithms, most researchers transform it to a two-dimensional time-domain matrix. This method, however, ignores the spatial information between electrodes, which might be highly useful for EEG categorization. As a result, this work presents both three-dimensional and four-dimensional formats for presenting EEG data that explicitly maintain the spatial information of electrodes in the EEG data structure. Several classification models will be trained and evaluated on diverse datasets utilising EEG data in the formats of two dimensions, three dimensions, and four dimensions to highlight the potential benefit of preserving spatial information about electrodes in the structure of EEG data for EEG classification.

Research Ethics Approval

This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee. The electroencephalography (EEG) data utilized in this study exclusively originated from published studies, thus negating the requirement for any additional ethical authorization.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Yige Yang*)

Acknowledgements

I want to start by expressing my gratitude to my supervisor, Douglas Armstrong, for his consistent support, insightful observations, and constant direction throughout the study process. His knowledge, patience, and helpful criticism were crucial in helping to shape and improve this dissertation. I am very appreciative of his mentoring and the chance to work under his supervision.

For their constant love, support, and tolerance throughout this tough period, I want to express my gratitude to my family and friends. My motivation to overcome obstacles and achieve this important academic milestone came from their unwavering support and belief in my talents.

Last but not least, I want to genuinely thank everyone who has supported and encouraged me during this trip. My Master's dissertation's academic success was greatly influenced by their ongoing support and belief in my talents.

Table of Contents

Chapter 1

Introduction

The human brain has been a subject of scientific study for centuries; it is a marvel of complexity and nuance. Its exceptional abilities in vision, cognition, and emotion have attracted researchers, inspiring extensive research aimed at unlocking its mysteries[13]. A powerful tool for studying the electrical impulses produced by the brain is electroencephalography (EEG), a non-invasive method that records and analyses these signals[18]. Consider it in this manner. EEG signals are frequently used to detect and monitor various neurological disorders in therapeutic settings[17]. Understanding the changes in brain function that occur as people age and develop is made easier with the help of EEG[29].

EEG classification is a technique for identifying patterns in brain waves acquired by EEG and applying these patterns to classify the person's cognitive or emotional state. In neuroscience, medicine, and human-computer interaction, EEG categorization has a wide range of uses[15, 21].

However, the nature of EEG signals present a challenge for EEG classification. EEG signals are typically recorded from multiple electrodes placed on the scalp, resulting in a high-dimensional data representation. The large number of channels and time points in EEG recordings leads to a high-dimensional feature space[8]. Besides, EEG signals are susceptible to various sources of noise and artifacts, which can distort the underlying neural activity. These noise sources include muscle activity (electromyographic noise), eye movements (ocular artifacts), environmental electromagnetic interference, and electrode artifacts[31]. As a result, many feature extraction approaches have been developed, such as time-domain and frequency-domain analysis, to minimise data dimensionality and identify relevant attributes for classification[30].

There has been an increase in interest in utilising machine learning algorithms to

identify EEG in recent years since they can automatically discover subtle patterns in data and give high accuracy. For EEG classification, a range of machine learning models such as support vector machines (SVMs), artificial neural networks (ANNs), decision trees, and random forests were utilised[15, 28, 1, 23]. At the same time, constructing 2D time-domain EEG signals has become the dominant way of representing EEG signals for machine learning models.

After reviewing others' works, we believe that it is beneficial to retain the spatial information between electrodes more explicitly for machine learning models in EEG classification. To prove this hypothesis, new structures of EEG data were created and the same or similar machine learning models were trained and tested on the newly structured data to compare with their performance on 2D time-domain data. In the end, the results showed models trained with location-considered time-domain data had a better performance than models trained with pure time-domain data. Therefore, it successfully proved the potential of containing locations of electrodes in the structure of time-domain EEG data.

The study's findings may have important consequences for improving the quality of life of persons with disabilities by making more accessible and inclusive technology available to them. Furthermore, by exposing a new approach for representing and classifying EEG data, this work hopes to enhance the discipline and stimulate further research in this area. The remaining chapters of this dissertation consist of the following sections:

1. "Background": This chapter provides an in-depth analysis of various structures of EEG signals.

2. "Methodology": This chapter focuses on the proposed structures of data and provides detailed information about the experimental setup.

3. "Results & Evaluation": This chapter presents the findings of the experiments conducted and provides an evaluation of the obtained results.

4. "Conclusion": The final chapter of the dissertation summarizes the key findings, draws conclusions based on the results, and offers insights into the future work.

Chapter 2

Background

2.1 Literature Review

The field of machine learning and data analytics has witnessed tremendous advancements in recent years, revolutionizing various domains and industries. One crucial aspect of this progress lies in the ability to develop accurate classification models that can extract meaningful insights and make informed decisions. However, the efficacy of these models heavily relies on the quality and representation of the input data.

In the context of classification models, one area that has received considerable attention is the representation of raw data before feeding it into the model. Effective representation techniques aim to capture the essential characteristics and patterns of the data, enhancing the model's ability to discern meaningful distinctions and make accurate predictions.

Because of its intrinsic richness and multidimensionality, raw EEG data offers unique obstacles for representation. Raw data collected from sensors or other measuring equipment may contain a variety of properties and features important for categorization tasks. However, not all qualities contribute equally to the model's prediction capacity, and some may provide noise or irrelevant data that might impede proper categorization.

As a result, it is critical to preprocess and describe raw EEG data in a way that optimises the model's performance. We intend to explore the existing approaches and methods used for representing raw EEG data in the context of classification tasks in this literature review, with the goal of identifying current practises and uncovering novel strategies that could bring a new representation of the data to potentially improve classification model performance.

2.1.1 Approaches using subset of channels

2.1.1.1 Channel Selection

Channel selection refers to the process of identifying a subset of EEG channels that are most useful or informative for the analysis or modelling task. This strategy aims to reduce computer complexity while increasing interpretability by focusing on channels that collect the most relevant information. There are various options available. Choosing channels with the highest variance or discriminatory power, for example, can improve the signal-to-noise ratio and feature quality[11].

In the study of Yong et al., the authors focused on automatic feature selection for motor imagery EEG pattern recognition[33]. They proposed a method based on the difference in variance between different classes of motor imagery tasks. By selecting channels with significant variance differences, they were able to identify the most discriminative channels for classification[33]. Their results demonstrated that this variance-based channel selection method improved the accuracy of motor imagery classification compared to using all channels[33].

Another is selecting based on scalp topography, which focuses on identifying channels that exhibit prominent spatial distribution patterns related to specific cognitive or neural processes[14, 25]. These patterns can provide insights into the functional organization of the brain. However, there are potential drawbacks to consider. Excluding certain channels may result in the loss of important information, neglecting less prominent but still relevant neural activity. The subjective selection of channels can introduce bias based on prior knowledge or assumptions, and the optimal channels for selection may vary depending on the specific research question, experimental design, or population under investigation. For example, in the scenario of scalp topography-based selection, scalp topography can vary across individuals, making it challenging to generalize the selected channels to different populations or experimental contexts.

On the other hand, scalp topography-based selection is not suitable for classifying tasks when the associated potential differences on the scalp are very small or the spread of the differences have no obvious pattern. Therefore, electrodes cannot be selected effectively. For example, this approach may not perform well on classification tasks related to subjects with certain neurological conditions, such as Alzheimer's disease, Parkinson's disease, or multiple sclerosis, which can cause significant change in EEG activity[24, 2, 6]. Furthermore, these changes could increase the difficulty to identify the related scalp area.

2.1.2 Approaches using all channels

2.1.2.1 EEG Signals to 2D Matrix in Time Domain

A sample of EEG signals will be converted into a 2D matrix using this method. Each row contains the values from one electrode in the time sequence, and each column contains the values from all electrodes at the same time point. For example, Zheng et al (2015) suggested a deep learning model for EEG categorization based on a deep belief network (DBN). The model was trained on raw EEG signals organised in a 2D matrix and obtained 89.21% accuracy in a two-class classification task[35].

Chen et al. suggested a CNN-based model for emotion identification from EEG data the same year. They also subjected these 2D raw EEG signals to a series of extraction techniques such as combined wavelet transform, principal component analysis (PCA), and autoregressive modelling, before feeding the retrieved features into a CNN. For a four-class classification task, the model attained an accuracy of 87.4%[10].

Schirrmeister et al. developed a CNN-based deep learning model for EEG decoding and visualisation in 2017. They employed 2D EGG signals as inputs and obtained 87.8% accuracy on a four-class classification task[23]. Song et al.'s hybrid model, which consists of a convolutional component and a transformer component, attained an average accuracy of more than 80% for motor imaging tasks in 2022[26].

Overall, the efficacy of organizing raw EEG signals into two-dimensional structures has been substantiated through the utilization of diverse machine learning models, ranging from CNN to transformers. These models have provided compelling evidence regarding the broad applicability of this approach. Furthermore, the classification outcomes attained through the adoption of this methodology have consistently demonstrated remarkable levels of accuracy and this has become the popular way of representing EEG signals among researches in recent years.

Nevertheless, it is noticeable that the performance of models trained on certain subjects is not sufficiently satisfactory. For instance, in the experiment conducted by Song et al. [26], the subject 2 exhibited an accuracy that was 14% lower than the average accuracy across all nine subjects. In the study conducted by Zheng et al. [35], it was observed that the standard deviation of classification accuracies across six subjects was approximately 8.5. This finding indicates that the performance of the poorest-performing model trained on the subject 1 significantly deviated from the average accuracy.

Given that EEG classification is directly employed in the diagnosis of diseases or

the facilitation of disability, it is imperative to strive for models that exhibit high average accuracy and low standard deviation across different subjects. For instance, consider a scenario where an EEG-based classification model reports high average accuracy in diagnosing a neurological condition. A practitioner, relying on this information, proceeds to diagnose a patient based on the model's results. However, due to the limitations of the model and the potential variations in individual subjects, a false diagnosis is made, leading to inappropriate treatment or unnecessary interventions. Despite the high average accuracy reported by a model, the model performance on individual cases can still vary a lot. Consequently, there exists ample room for improvement in either enhancing the existing models or exploring alternative approaches for representing EEG signals.

2.1.2.2 EEG Signals in Frequency Domain

Bashivan et al. (2015) suggested a method for producing a feature vector by combining spectral readings from all electrodes over the same time period[8]. They employed the Azimuthal Equidistant Projection (AEP) method to project the position of EEG electrodes onto a 2-D surface while keeping the relative distance between electrodes[8]. The spectral strength within three important frequency bands is computed for each site, and this information is used to produce topographical pictures for each time frame[8]. Finally, the pictures were fed into a deep recurrent-convolutional neural network (RCNN) for classification[8].

Although the study found that this method of organising EEG data produced superior outcomes than non-deep learning models, the efficacy of this strategy is only partially demonstrated due to a lack of comparison with machine learning approaches using different structures of EEG signals.

On the other hand, this approach increased the complexity of data by involving the locations of electrodes and generating spectral power according to different frequency bands. The selections of frequency bands here can be a problem when encountering different classification tasks. Nonetheless, this unique technique illuminated the concept that retaining the data's underlying spatial organisation might be advantageous to EEG categorization.

2.1.2.3 EEG Signals in Time-Frequency Distribution

In this approach, EEG signals are converted to time-frequency images namely spectrograms, which are a visual representation of the spectrum of frequencies of a signal over time[32]. To create a spectrogram in general, we typically start with a time-domain signal, such as an audio waveform. This signal is divided into short overlapping segments, and for each segment, a Fourier Transform is applied to obtain its frequency spectrum[34]. The result is a series of spectra for each time segment, forming a three-dimensional representation with time on the x-axis, frequency on the y-axis, and magnitude or intensity represented by the color or brightness[19].

In 2013, Bajaj et al. focused on classifying sleep stages from EEG signals using the smoothed pseudo Wigner-Ville distribution (SPWVD) for obtaining the time-frequency representation of the signals[7]. Their proposed method achieved 88.47% of accuracy[7]. However, a sample data is in the length of 30 seconds, which is quite long compared to the deep learning methods mentioned in the last section. For example, Song et al. only used 4-sec samples to train their model and achieved similar results in terms of accuracy.

In 2014, Fu et al. proposed a seizure detection method using a combination of the Hilbert-Huang Transform (HHT) and Support Vector Machine (SVM) for analyzing EEG signals[12]. They decomposed the EEG signals before using Hilbert-Huang Transform to obtain related spectrogram[12]. Their method showed promising results in accurately classifying seizure and nonseizure EEG signals[12]. Their method of using spectrograms may not be transferable to other more complicated classification tasks. For example, since a seizure is a transient event characterized by an abnormal and excessive electrical discharge among brain cells[20], their method may not work well when classifying events when we need to identify the pattern of all abnormal electrical discharge.

In 2020, Aslan et al. proposed a method to automatically diagnose individuals with schizophrenia (SZ) using EEG recordings[5]. By transforming raw EEG signals into spectrogram images and feeding them to a CNN model, they achieved 95% and 97% of accuracy in Dataset A and B irrespectively[5]. This proposed method with other deep learning methods using time-frequency EGG data stands out for its simplicity and effectiveness, as it does not require manual feature extraction or extensive preprocessing[5]. However, it has been noticed that, although Time-Frequency Distribution (TFD) methods allow the analysis of relatively long continuous segments of EEG data with rapidly

changing dynamics, they require a balance between time and frequency resolution and may not always provide detailed information compared to pure frequency domain methods or pure time domain methods[3].

2.2 Summary & Inspiration

The methods covered in the preceding section centre on using various EEG signal representations for analysis and classification tasks. To simplify and make interpretation easier, the first set of methods chooses a subset of EEG channels. Although these methods can increase interpretability and computational effectiveness, they have drawbacks include the possible loss of useful data and subjectivity in channel selection.

The second set of approaches involves using all channels of EEG signals. The most promising one is to transform EEG signals into a 2D matrix in the time domain, where columns represent signals from all electrodes at the same time point, and rows represent signals from one electrode across all time points. This representation has been successfully utilized in various deep learning models, such as CNNs and transformers, achieving high accuracy in EEG classification tasks. However, the performance of models trained on individual subjects can vary, indicating a need for improvement to ensure robustness across different subjects.

Representing EEG signals in the frequency domain while keeping the spatial information of electrode placements is another method that is interesting. This method entails placing electrodes in their exact locations on a 2D surface, and then producing feature vectors for each electrode by combining spectral readings from all electrodes taken over a predetermined period of time. The time-domain 2D technique, which structures the data while taking electrode placements into account, might be improved with the use of this method. Comparing this approach to the frequency-domain approach, this allows a reduction in the complexity of the data structure. It also avoids the information loss that could happen when choosing a certain time frame to compute spectral power.

Comparing to the 2D time-domain approach, incorporating the spatial information of electrodes allows models to process the data with additional contextual information. This additional information has the potential to improve the average accuracy of the models and reduce the associated standard deviation across different subjects. Consequently, the resulting models may offer enhanced reliability and performance. Therefore, this literature review inspires us to examine the potential of integrating the spatial information of electrodes into time-domain EEG data.

Chapter 3

Methodology

The experiments involved two EEG datasets, which underwent several preprocessing steps. First, the qualified trails were selected and labelled. Then, they were subjected to band-pass filtering and normalized. Subsequently, the data was transformed into 2D, 3D, and 4D structures. For the baseline models, state-of-the-art models, namely a CNN model and a conformer model, were selected from previous research papers (Schirrmeister et al., 2017; Song et al., 2022). To investigate the advantages of using 3D and 4D data, modifications were made to the CNN model and the conformer model and then train them on the respective 3D and 4D data for each subject. The results obtained from these modified models were then compared with the baseline results.

To ensure a fair and valid comparison, the hyperparameters, including learning rates, were kept the same as those used in training the baseline models. Moreover, the modifications made to the models were limited to the addition of new structures on top of the existing baseline models.

3.1 Datasets

Two datasets were utilised in this study. One is the Graz dataset 2a, which was collected by Graz University of Technology and features EEG recordings from nine subjects undergoing a cue-based Brain-Computer Interface paradigm with four motor imagery tasks[9]. The EEG signals were from 22 electrodes. These tasks include left-hand movement (class1), right-hand movement (class2), both feet movement (class3), and tongue movement (class4)[9].

The other is EEG Motor Movement/Imagery Dataset[22], which were collected on 64 electrodes from a hundred subjects with four tasks (considering the amount time for

training models, we randomly selected 50 subjects). These are task 1 (open and close left or right fist), task 2 (imagine opening and closing left or right fist), task 3 (open and close both fists or both feet) and task 4 (imagine opening and closing both fists or both feet)[22].

Within the framework of our study, the open accessibility of both datasets presents a significant advantage, enabling us to conduct comprehensive comparisons of our model's outcomes against those of other researchers. Specifically, the Graz BCI Competition Dataset 2a has been leveraged in the studies where our baseline models were introduced[9]. Notably, this dataset provides meticulous descriptions of each stage of the recording process down to the second, enhancing our potential to achieve refined model results[9]. The EEG Motor Movement/Imagery Dataset, in contrast, boasts a more extensive subject pool compared to the Graz BCI Competition Dataset 2a, augmenting our ability to substantiate the robustness and broad applicability of our approach[22]. Divergences in electrode count, sample length, and classification labels between the two datasets collectively serve as a rigorous testbed to assess the resilience of our theoretical framework[9, 22].

3.2 Data preprocessing

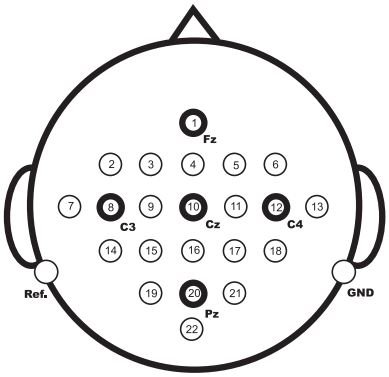
The EEG recordings underwent several preprocessing stages. Initially, individual trials corresponding to specific task periods of interest were extracted and isolated from the continuous EEG data for each run. This extraction refines the models' focus on relevant task-related intervals. Subsequently, these extracted trials were standardized to a fixed duration. This adjustment is essential due to variations in the length of certain recording segments. A fixed duration aligns with the input requirements of machine learning models, which need uniform input sizes. Moreover, a band-pass filter spanning from 4 Hz to 40 Hz, as detailed in [27], was applied to the isolated trials. The selection of this frequency range aligns with established motor imagery task-related frequency bands, as described in [4], and aims to eliminate extraneous signals outside this designated range. This standardized approach is a prevalent technique in EEG classification studies and also mentioned in the studies where our baseline models were originated. By adhering to this process, we ensure our baseline results can be a valid representation of the performance of the 2D EGG structure. After the completion of the preprocessing steps, both datasets were partitioned into training and testing sets, allocating 70% of each class for training purposes and reserving the remaining 30% for testing evaluation.

3.3 Data Construction

Constructing 2D data is straightforward. EGG data will be stored as a 2D matrix, with columns representing all voltages at a given time point and rows representing voltages from a single electrode over time. Taking the electrodes of the Graz dataset in the upper graph of Figure 3.1 as an example, the values from electrodes 1-22 at the same time point will be in the same column.

In contrast, with 3D data, the voltages from all electrodes at each time point create a 2D matrix, with the position of each electrode respected. Using the middle graph of Figure 3.1 as an illustration, the values from electrodes 1-22 at the same time point will be in a matrix with shape (6, 7). The electrode 10 will be at the coordinate (2, 3) with the electrode 11 at (2, 4). Stacking the 2D matrix by time order results in a 3D matrix. To show the absence of EEG signals, the empty values in the 3D matrix will be assigned a value of 0, since the voltage range for EEG data can not be zero since the brain is functioning constantly [16]. Others might inquire about the possibility of encoding missing data as null values within the matrix. Nevertheless, incorporating null values into the input for machine learning models poses programming challenges. This approach directly leads to running error in the training of the models, when using Python ML libraries. The two-dimensional matrices can then be stacked along the time axis to generate a three-dimensional matrix. The reason to do this is to represent the relative locations between electrodes on a 2D surface.

In the 4D construction, instead of the 2D matrix discussed in the 3D approach, all voltage values at one time point are placed in a 3D matrix. If the source of signals are like the ones in the upper graph of Figure 3.1, the 3D matrix will be in the shape of (6,7,4) with the electrode 1 at (2,3,0), the electrode 11 at (2,4,1) and the electrode 12 at (2,5,2). In other words, as indicated in the lower graph of Figure3.1., electrode 10 will be the centre of a two-dimensional matrix at the top level of the three-dimensional matrix, and the four neighbouring values to the central electrode will appear in an a two-dimensional matrix at the second top level of the three-dimensional matrix. As the distance from the centre electrode grows, the level of 2D matrices for voltage values drops, and 0 is utilised to represent empty values. As a result, a 3D matrix for a single time point is produced. When the time domain is included, the 3D data is transformed to 4D data by stacking the 3D matrices along the time axis, each 3D matrix representing a different time point. This is done to depict the relative positions of the electrodes in 3D space. As a result, EEG data provides more spatial information.



	0	0	0	1	0	0	0
	0	2	3	4	5	6	0
	7	8	9	10	11	12	13
	0	14	15	16	17	18	0
	0	0	19	20	21	0	0
	0	0	0	22	0	0	0

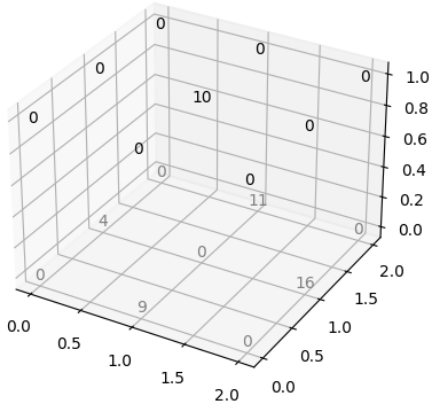


Figure 3.1: The upper graph is an illustration for the locations of electrodes for the Graz dataset; the middle graph is an illustration of how the electrodes locate in a 2D matrix of the 3D structure; the lower graph is an illustration of how the electrodes locate in a 3D matrix of the 4D structure

3.4 Model Structures

3.4.1 Baseline 2D Models

- CNN: Schirrmeister et al. introduced a convolutional neural network (CNN) design tailored for EEG decoding. Their CNN design is made up of several layers, including convolutional, pooling, and fully linked layers[23]. The study

demonstrates that this CNN architecture attains exceptional performance on multiple EEG datasets, establishing the efficacy of deep learning in EEG decoding and positioning it as a leading approach in the field.

- Conformer: A new framework has been introduced for end-to-end EEG classification, which integrates Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) [27]. This framework comprises three main elements: a convolutional component, a self-attention component, and a fully-connected classifier. The final stage involves a concise classifier with multiple fully-connected layers to generate the classification outcomes. The primary objective of this proposed framework is to harness the advantages of both CNNs and Transformers while addressing their respective limitations [27].

3.4.2 3D Models

Initially, our approach involved building models by leveraging existing video classification models. This choice was motivated by the fact that these models are specifically designed to handle 3D inputs, and the field of video classification is well-established, providing ample resources for learning. However, we soon realized that video classification models tend to be considerably larger compared to our baseline models or other models used in EEG classification. The feasibility of training such large models with limited EEG data became a matter of concern.

Furthermore, the video classification model can identify certain regions for classification and neglect the area beyond the regions. However, because the signals from all electrodes are all interconnected, it is best for our model to focus on the full data. Furthermore, rather than researching innovative model designs, our major goal is to illustrate the possibilities of the new data structures. To accord with this purpose, it would be more helpful to adjust the baseline models and minimise model differences, resulting in better support for our desired goals.

3.4.2.1 autoencoder + baseline models

The autoencoder is a neural network architecture designed to compress and reconstruct input data, effectively learning a compact representation of the original data. In our specific implementation, the autoencoder takes a 2D matrix at each time point in the structured 3D data and aims to compress it.

Table 3.1: Encoder and Decoder Layers for Graz Dataset

Part	**Layer**	**Kernel Size**
Encoder	nn.Conv2d+nn.ReLU	(3, 2)
Decoder	nn.ConvTranspose2d+CustomActivation	(3, 2)

The structure of our autoencoder as shown in Table 3.1 involves a single 2D convolutional layer for both the encoder and decoder. We experimented with different models and found that adding more layers increased the loss, indicating poorer performance. It is possible that the small size of the 2D input and the presence of many zero values led to this result. Therefore, we opted for a simpler architecture with a single layer for each component.

Regarding the choice of the custom activation function for the decoder layer, we found that our EEG data typically ranges from approximately -20 to 20, after normalization. Standard activation functions would not cover this range adequately, resulting in unmanageable model loss. To address this, we introduced a custom activation function m * torch.tanh(x). This approach allows a range between -m and m. The parameter m is the larger absolute value between the maximum and minimum values from a certain subject. By employing this custom activation function in the decoder layer, our autoencoder can better handle the EEG data's value range, leading to improved reconstruction performance and mitigating issues with model loss.

The autoencoder is trained to compress a 2D matrix from a 3D sample into a $2, x$ matrix. This compressed matrix is then flattened into a 1D array with 2x elements. These flattened arrays from the same 3D sample are then stacked together in the order of time, forming a 2D matrix. This transformed data can be processed by baseline models. The decision to flatten the compressed matrix into a 2D array instead of a 1D array is to minimize the loss in the autoencoder, as compressing directly to a 1D array would result in a larger loss. At the same time, we aware of the increased training time from the autoencoder and the possible loss of information in the flattening step.

3.4.2.2 3D convolution + baseline models

These modified models consist of additional layers in Table 3.2, which will be placed on the top of two baseline models. The input 3D data from the Graz dataset for these models have a shape of (batchsize, 1, 6, 7, 1000). After passing through the convolutional layers described in Table 3.2, the output shape will be (batchsize, 40, 1, 976). This

output is obtained by applying the convolutional operations on the input data, resulting in feature maps with 40 channels and dimensions of 1x976. Next, the data is squeezed along dimension 2, resulting in a shape of (batchsize, 40, 976). Squeezing eliminates the singleton dimension and retains the feature maps with 40 channels and a length of 976.

Table 3.2: Description of 3D Convolutional Layers

Layer	Input Channels	Output Channels	Kernel Size	Padding
1	1	40	(1, 1, 25)	(0, 0, 0)
2	40	40	(6, 7, 1)	(0, 0, 0)

The rest of the models then process the compressed data. These baseline models use the processed input shape of (batchsize, 40, 976) for further training. In comparison to the semi-supervised technique, we can prevent information loss when training the encoder and the flattening phase. As a result, we have a greater chance of outperforming the baselines. However, there is fear that the 3D convolutional layers will not be adequately trained since the model structure grows intricate and the quantity of samples remains constant.

3.4.3 4D Model

The 4D model is likewise semi-supervised, with its encoder immediately converting 4D input data to 2D data. The autoencoder's encoder, as illustrated in Table 3.3, consists of two 3D convolutional layers. The input data has the following format: (batchsize, 1000, 6, 7, 4), with 1000 indicating the time axis.

Table 3.3: Encoder Layers

Layer	**Input Channels**	**Output Channels**	**Kernel Size**
3Dconv Layer 1	1000	1500	(1, 1, 4)
3Dconv Layer 2	1500	1000	(6, 1, 1)

The encoder uses two 3D convolutional layers to compress the input data into a form (batchsize, 1000, 7). To guarantee effective training, this autoencoder uses a custom activation mechanism. The baseline models then process the compressed data with the shape (batchsize, 1000, 7) to execute subsequent tasks.

3.5 Training & Evaluation Metrics

In our experimentation, we ensured a fair comparison between the 2D structure and our proposed structures by maintaining consistency in the training process. For conformer-related models, the number of training epochs was fixed at 150, while for cnn-related models, it was set to 100. Additionally, we kept the hyperparameters such as learning rate and training techniques like data augmentation the same as those used in the baseline models. By doing this, we aimed to isolate the impact of the structural modifications on model performance. The focus was on evaluating the potential of our proposed structures objectively and comprehensively.

To assess the models, we employed several evaluation metrics. The average highest accuracy across all subjects allowed us to evaluate overall performance. Moreover, analyzing the standard deviations among all subjects helped us understand the consistency and reliability of the models' predictions. Furthermore, we conducted a detailed examination of how specific models performed on certain subjects. This analysis provided valuable insights into potential variations.

Chapter 4

Results & Evaluation

4.1 Comparison between Baselines and Performance of Proposed Structures

The results presented in the upper plot of Figure 4.1 showcase the average highest accuracy achieved by different models on the Graz dataset. The vertical error bars, representing the standard deviation, provide insights into the consistency of each model's performance.

The baseline models, 'conformer_2D' and 'cnn_2D,' were used as reference points. Interestingly, the models trained on the proposed data structures demonstrated a performance that closely matched the baselines. This suggests that the new data structures, particularly the 'semi_conformer_3D' and '3DConv_cnn' models, have the potential to be effective alternatives to the traditional 2D representations.

However, a notable exception was the 'semi_conformer_4D' model, which displayed significantly lower accuracy compared to all other models. The reason behind this lower performance is discussed later in this chapter.

Surprisingly, the '3DConv_conformer' model emerged as the best performer in this evaluation. With an average accuracy of 0.802 and a standard deviation of 0.129, it outperformed both baseline models. This finding indicates that the proposed 3D data structure has an effect, leading to improved predictive capabilities and enhanced generalization.

The lower plot in Figure 4.1 presents the average highest accuracy achieved by different models on the Motor dataset, with the error bars representing the standard deviation. The results displayed in this figure are highly promising and indicate significant

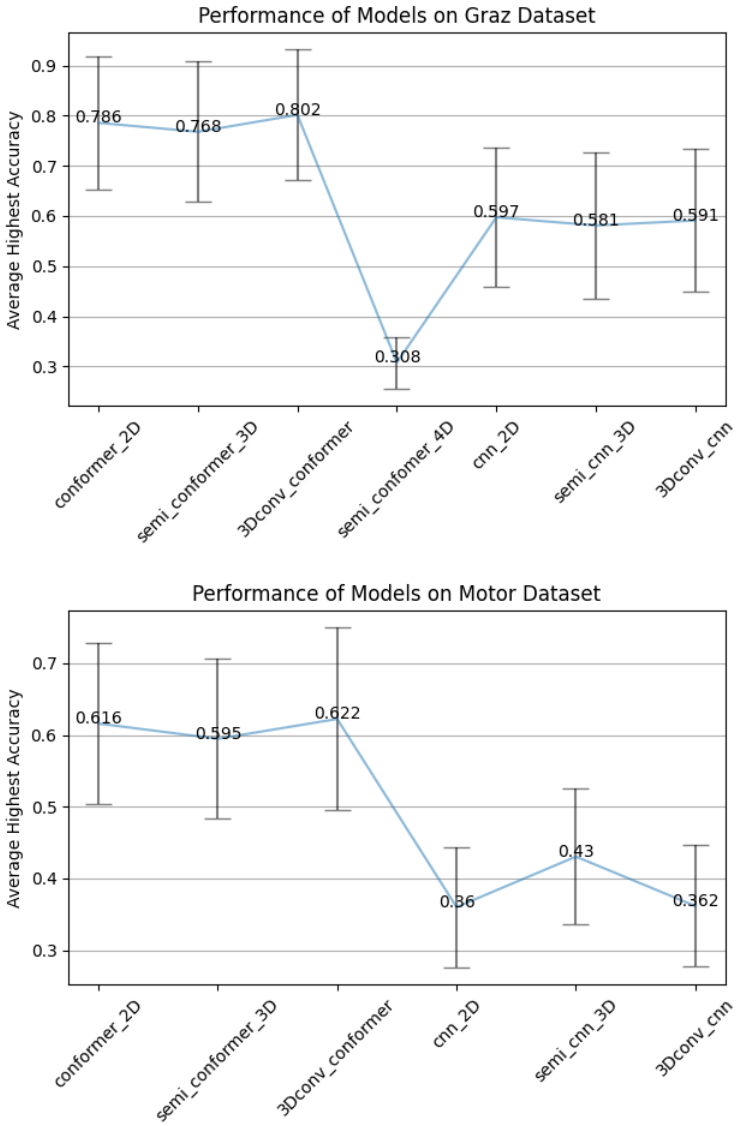
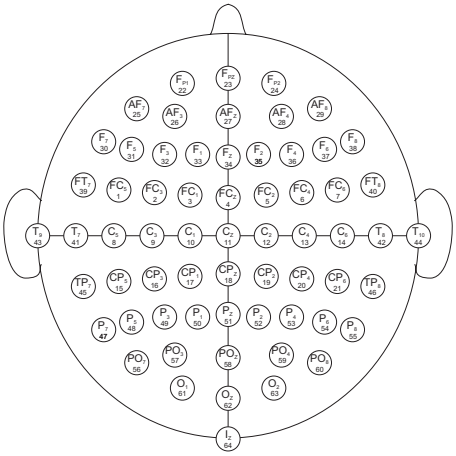


Figure 4.1: Performance of Models

improvements over the previous evaluation.

Among the models, the '3Dconv_conformer', 'semi_cnn_3D', and '3Dconv_cnn' models surpassed their associated baselines. Notably, the '3Dconv_conformer' model emerged as the top performer, achieving an accuracy of 0.622 across 50 subjects. This result is particularly positive as it surpassed its corresponding conformer baseline by 0.01 on the average. Moreover, the 'semi_cnn_3D' model also surpassed its corresponding CNN baseline by 0.16 on the average. Thus, they all showcased the potential of incorporating spatial information in EEG classification.



	0	0	0	F22	0	F23	0	F24	0	0	0
	0	AF25	0	AF26	0	AF27	0	AF28	0	AF29	0
	0	F30	F31	F32	F33	F34	F35	F36	F37	F38	0
	0	FT39	FC1	FC2	FC3	FC4	FC5	FC6	FC7	FT40	0
	T43	T41	C8	C9	C10	C11	C12	C13	C14	T42	T44
	0	TP45	CP15	CP16	CP17	CP18	CP19	CP20	CP21	TP46	0
	0	P47	P48	P49	P50	P51	P52	P53	P54	P55	0
	0	PO56	0	PO57	0	PO58	0	PO59	0	PO60	0
	0	0	0	O61	0	O62	0	O63	0	0	0
	0	0	0	0	0	I64	0	0	0	0	0

Figure 4.2: The left one is the illustration from the Motor dataset, and the right one is the illustration for our transformed data structure.

Moreover, by comparing the general performance between the two datasets, it is obvious that the benefit of 3D data in the Motor dataset was better presented than in the Graz dataset. Notably, the Graz dataset's samples have 1000 time points, whereas the Motor dataset's samples only have 646 time points in the experiments. This information and the different performances suggest that the spatial information in the new data structure was helping the EGG classification; and it is possible that the impact is more obvious when the length of time points in a sample is relatively short. This feature is aligned with the practical application in BCI, where it needs the classification models to respond instantly. Additionally, the Motor dataset included data from 50 subjects, compared to only 9 subjects in the Graz dataset. This performance further emphasizes the general benefits and robustness of the proposed 3D data structure across diverse datasets and subject populations.

Furthermore, the 3D structure of the data has demonstrated a certain level of tolerance to inaccuracies in the relative electrode locations. Despite not strictly adhering to the 2D illustration for the Motor dataset as shown in Figure 4.2, the models trained with the 3D structure data have shown superior performance compared to the baseline models. This finding suggests that the 3D structure is highly applicable and generalizable

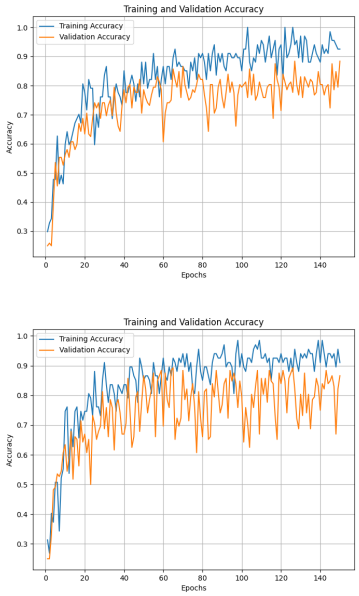


Figure 4.3: The upper plot and the lower plot are the training process of the baseline conformer model and the '3Dcon_conformer' model on subject 1 from the Graz dataset.

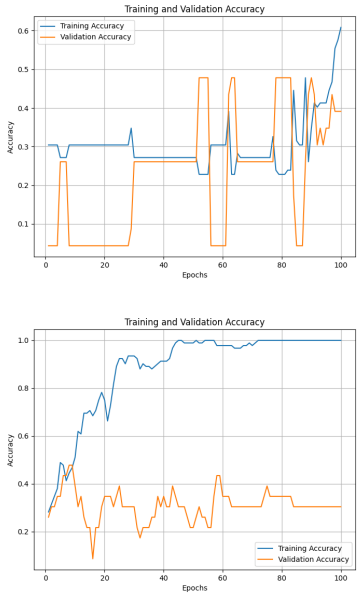


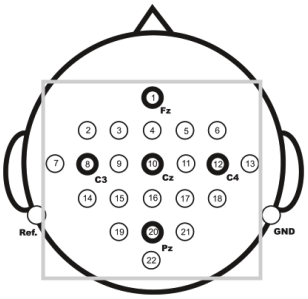
Figure 4.4: The upper plot and the lower plot are the training process of the 'semi_cnn_3D' and the baseline CNN model on the same subject from the Motor dataset.

across various situations. However, it also raises the question of whether increasing the accuracy of electrode locations in 3D data may lead to even a better model performance.

Despite all the advantages mentioned above, the performance of our 3D structure is not optimal, particularly during the training process of the '3Dconv_conformer' model, which appeared to be less stable compared to the 'conformer_2D' baseline model, as illustrated in Figure 4.3. We suspect three potential reasons for this behavior.

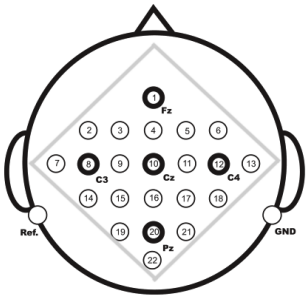
Firstly, inaccuracies in the 3D structure might have affected the model's performance. The electrode placements on the subject's scalp, conducted by practitioners, might not have strictly followed the pattern shown in Figure 3.1. Consequently, when constructing the 3D data, we adhered to the relative electrode locations in the figure, which could have introduced inaccuracies into the model.

Another possible reason lies in the data augmentation process used in the conformer model. The author of the model employed a data augmentation method known as 'the segmentation and reconstruction approach' [26]. This technique involves segmenting samples into N parts with the same time domain length and randomly concatenating them with parts from other samples while preserving the original order in time. However,



0	0	0	1	0	0	0	
0	2	3	4	5	6	0	
7	8	9	10	11	12	13	
0	14	15	16	17	18	0	
0	0	19	20	21	0	0	
0	0	0	22	0	0	0	

Figure 4.5: The 2D matrix at the same time point in the proposed 3D structure.



			0				
		0	1	0			
	2	3	4	5	6		
7	8	9	10	11	12	13	
	14	15	16	17	18		
		19	20	21			
			22				

Figure 4.6: The new 2D matrix at the same point to reduce the number of added zeros.

when adapting the conformer models for our purposes, we ensured that the segmentation occurred within the time domain. Moreover, the training process of non-augmentation CNN models with 3D data still exhibited relative instability, as seen in Figure 4.4, where both models lacked data augmentation. Thus, this reason is less likely.

The third reason relates to the addition of zeros during the construction of 3D data from the 2D data. In our proposed method to construct the 3D data, as shown in Figure 4.5 (where the gray square represents the 2D matrix at the same time point), values from the same time point were used to create a (6, 7) matrix, resulting in 20 added zeros. This means that the number of added zeros is nearly equal to the number of non-zero values in the 3D data, potentially contributing to the training process instability. Alternatively, we could have adopted the matrix construction approach depicted in Figure 4.6, where the resulting matrix would have been much more compact, with only 3 added zeros.

These suspected reasons warrant further investigation and potential adjustments to improve the stability and performance of our 3D structure.

4.2 P-values of the Performance Results

To examine the robustness of the improvement shown by the 3D data structure, p-values were calculated between the performance of the baseline models and their corresponding models using 3D EEG data. The null hypothesis for this evaluation is that the 3D data improved the model performance by chance, and the alternative hypothesis is that

Table 4.1: P-values

Dataset	**2D Baseline**	**3D Model**	**p-value**
Graz	conformer_2D	3Dconv_conformer	0.09
Graz	cnn_2D	3Dconv_cnn	0.58
Motor	conformer_2D	3Dconv_conformer	0.07
Motor	cnn_2D	semi_cnn_3D	9e-05

the improvement is not occurred by chance. Since we wanted to evaluate the groups of accuracy from the 3D data and 2D data, this satisfied the conditions of a t-test, which are comparing means, involving two independent groups and a small sample size. Therefore, a two-tailed test was conducted to further calculate the p-values. The table 4.1 presents the calculated p-values for each dataset and model comparison.

The outcomes presented in Table 4.1 offer intriguing insights into the significance of the advancements realized by the 3D models. When examining the Graz dataset, both the conformer_2D and 3Dconv_conformer models yield a p-value of 0.09, while the cnn_2D and 3Dconv_cnn models exhibit a p-value of 0.58. This means that we have 9% chance to reject the null hypothesis in the setting of the first comparison and 58% chance in the other comparison setting. Notably, both of these p-values are relatively high in comparison to the conventional significance threshold of 0.05. Therefore, these values are not enough to reject the null hypothesis.

Shifting our attention to the Motor dataset, the baseline models, conformer_2D and cnn_2D, result in p-values (0.79 and 9e-05, respectively) when contrasted with their corresponding 3D counterparts, 3Dconv_conformer and semi_cnn_3D. Remarkably, the p-value associated with the cnn_2D vs. semi_cnn_3D comparison falls below 0.01. This lower p-value indicates that the observed improvement is not an outcome of chance, but rather a consequence of the novel data structure.

Considering the fact that we conducted the four p-values, there was the probability of obtaining at least one false positive result. To control this effect, Bonferroni Correction was applied, which reduced the conventional significance threshold to 0.125. It is obvious that only one value is lower than this significance threshold and successfully reject the null hypothesis.

Therefore, we can draw the preliminary conclusion that the integration of a 3D structure is a promising avenue for EEG classification but lack of a solid proof. It needs more experiments and data to verify its general benefit to EEG classification. In

general, the results does establish the potential value of explicitly preserving the spatial information of electrodes within the EEG data structure.

4.3 Comparison between 3DSemi and 3Dconv models

As depicted in Figure 4.1, the models equipped with 3D convolutional layers was more likely to exhibit superior performance compared to the models with autoencoders, particularly when their associated baseline models achieved satisfactory results. The lower performance of the semi models could be the result of the information loss during the encoder training and the flattening step.

Additionally, our initial concerns, as outlined in Section 3.4.2.2, regarding the 3D convolutional layer's impact did not happen. We attribute this to the fact that the added values in the 3D structure were zeros, thereby mitigating any adverse effects stemming from increased complexity in the data structure. This finding suggests an encouraging trend that integrating 3D convolutional layers maybe align better with our proposed 3D structure. However, further experiments is needed to confirm it.

4.4 Explanation about the Performance of 4D data

The 'semi_conformer_4D' model exhibits an average highest accuracy of 0.308. It is quite low since this slightly surpasses the accuracy achieved by random guessing (0.25). A closer examination of the training accuracy reveals that the increase of training accuracy is limited from 0.2 to 0.4. This observation suggests a concerning trend: the model appears to struggle in effectively learning from the training data.

Two possibilities could account for this. Firstly, the structure of the model might hinder its capacity to obtain meaningful information from the data. However, it is important to note that other models employing the same conformer architecture have demonstrated an aptitude for learning. Thus, this explanation becomes less likely.

The second plausible explanation revolves around the integrity of the data during its transformation into a 4D format. Therefore, we examined our transformation process and found where the potential destruction happened. In our proposed 4D structure, the top electrode was placed at the center of a 2D matrix, and its four surrounding electrodes on the scalp were located in a lower matrix surrounding the projection of the top electrode onto this lower matrix. This arrangement treated the horizontal and vertical

distances between adjacent electrodes as equal. However, in reality, the horizontal and vertical distances between two adjacent electrodes on a 3D scalp are different.

This discrepancy might explain why the performance of the 4D data structure contradicts the results seen with the 3D data structures, which showed promising improvements over the baselines. To evaluate this reason, a potential solution could be to introduce zero paddings on the horizontal levels. By doing so, the difference between horizontal and vertical distances can be respected.

Applying zero paddings alone, however, is insufficient to fully use the capabilities of the 4D data structure because the datasets only show the electrode positions in a 2D representation. In order to generalise to a 3D visualisation of the positions of the electrodes, a new EEG dataset must be created that precisely documents the locations of electrodes for each participant in 3D.

Researchers would have a more complete and accurate representation of 4D EEG data if they created a new EEG dataset that incorporates 3D information about electrode positions. With the use of this new dataset, future research might examine the possible advantages of the 4D data structure and learn more about how spatial information affects EEG categorization tasks.

Chapter 5

Conclusion

This study aimed to explore the potential benefits of explicitly adding spatial information of electrodes in EEG data through the introduction of two novel data structures: a 3D structure and a 4D structure. The 3D structure preserved the locations of the projected electrodes on an imaginary 2D plane parallel to the floor, while the 4D structure maintained the spatial information of the electrodes on the 3D scalp. The hypothesis was that both the 3D and 4D structures could outperform the baseline models, with the 4D structure expected to perform even better due to its inclusion of more comprehensive spatial information.

To test this hypothesis, compatible models were created and trained using the proposed 3D and 4D data structures, and their performances were compared with two baseline models on EEG classification tasks using the Graz and Motor datasets. The results were highly encouraging, as the proposed 3D data structures surpassed the baseline 2D models, demonstrating the potential of incorporating spatial information to enhance predictive capabilities and generalization in EEG classification.

5.1 Limitation & Future Work

There are areas of improvement for the 3D structure. Firstly, constructing 3D data could have been more precise by closely adhering to the 2D dataset illustrations. Secondly, optimizing the 3D data's efficiency by reducing unnecessary zero values would have enhanced its compactness. Lastly, while baseline model performance aligned with original papers [23, 26], substituting the CNN model with a contemporary one could have bolstered the persuasiveness of our conclusions, highlighting the relevance of our findings in light of recent advancements.

Regarding to the 4D structure, while the 4D structure was expected to outperform the 3D structure because it included more spatial information, the 'semi_conformer_4D' model displayed lower accuracy compared to all other models. This unexpected outcome could be attributed to inaccuracies in the representation of the interrelationship between electrode locations in the 4D structure. To address this issue and potentially improve the 4D structure's performance, the study suggested exploring solutions such as introducing zero paddings to preserve accurate spatial information and creating a new dataset is essential, which has a 3D illustration of the locations of the electrodes.

The comparison between the 'semi' models with autoencoders and the '3Dconv' models with 3D convolutional layers favored the latter, suggesting that it is possible that integrating 3D convolutional layers aligns better with the proposed 3D data structure, leading to an area of further study.

5.2 Conclusion

Based on the experimental findings, the 3D structure demonstrated improvements across diverse datasets and subject populations. This is evident from the promising results on the Graz and Motor datasets, despite variations in electrode count and data length.

The outcomes were further evaluated using p-values. The resulting p-values indicate that only one setting can successfully prove the improvement is not happened by chance. Consequently, our experiment only partially proves the advantage of explicitly preserving electrode spatial information in EEG data structure over the 2D structure. However, these results still showed the potential of explicitly preserving the locations of the electrodes in the structure of EEG data. Further experiments are warranted to thoroughly explore its benefits.

Bibliography

[1] U Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, and Hojjat Adeli. Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals. *Computers in biology and medicine*, 100:270–278, 2018.

[2] Amirmasoud Ahmadi, Saeideh Davoudi, and Mohammad Reza Daliri. Computer aided diagnosis system for multiple sclerosis disease based on phase to amplitude coupling in covert visual attention. *Computer methods and programs in biomedicine*, 169:9–18, 2019.

[3] Amjed S Al-Fahoum and Ausilah A Al-Fraihat. Methods of eeg signal features extraction using linear analysis in frequency and time-frequency domains. *International Scholarly Research Notices*, 2014, 2014.

[4] Kai Keng Ang, Zheng Yang Chin, Chuanchu Wang, Cuntai Guan, and Haihong Zhang. Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Frontiers in neuroscience*, 6:39, 2012.

[5] Zülfikar Aslan and Mehmet Akin. Automatic detection of schizophrenia by applying deep learning over spectrogram images of eeg signals. *Traitement du Signal*, 37(2), 2020.

[6] Biniyam A Ayele, Heera Tesfaye, Mehila Z Wuhib, and Guta Zenebe. Factors associated with eeg slowing in individuals with parkinson's disease. *Ethiopian Journal of Health Sciences*, 32(1), 2022.

[7] Varun Bajaj and Ram Bilas Pachori. Automatic classification of sleep stages based on the time-frequency image of eeg signals. *Computer methods and programs in biomedicine*, 112(3):320–328, 2013.

[8] Pouya Bashivan, Irina Rish, Mohammed Yeasin, and Noel Codella. Learning representations from eeg with deep recurrent-convolutional neural networks. *arXiv preprint arXiv:1511.06448*, 2015.

[9] Clemens Brunner, Robert Leeb, Gernot Müller-Putz, Alois Schlögl, and Gert Pfurtscheller. Bci competition 2008–graz data set a. *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*, 16:1–6, 2008.

[10] Fafa Chen, Baoping Tang, Tao Song, and Li Li. Multi-fault diagnosis study on roller bearing based on multi-kernel support vector machine with chaotic particle swarm optimization. *Measurement*, 47:576–590, 2014.

[11] Denis A Engemann, Federico Raimondo, Jean-Rémi King, Benjamin Rohaut, Gilles Louppe, Frédéric Faugeras, Jitka Annen, Helena Cassol, Olivia Gosseries, Diego Fernandez-Slezak, et al. Robust eeg-based cross-site and cross-protocol classification of states of consciousness. *Brain*, 141(11):3179–3192, 2018.

[12] Kai Fu, Jianfeng Qu, Yi Chai, and Yong Dong. Classification of seizure based on the time-frequency image of eeg signals using hht and svm. *Biomedical Signal Processing and Control*, 13:15–22, 2014.

[13] Gregory A Light, Lisa E Williams, Falk Minow, Joyce Sprock, Anthony Rissling, Richard Sharp, Neal R Swerdlow, and David L Braff. Electroencephalography (eeg) and event-related potentials (erps) with human participants. *Current protocols in neuroscience*, 52(1):6–25, 2010.

[14] Otavio G Lins, Terence W Picton, Patrick Berg, and Michael Scherg. Ocular artifacts in eeg and event-related potentials i: Scalp topography. *Brain topography*, 6:51–63, 1993.

[15] Fabien Lotte, Marco Congedo, Anatole Lécuyer, Fabrice Lamarche, and Bruno Arnaldi. A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of neural engineering*, 4(2):R1, 2007.

[16] Ernst Niedermeyer and FH Lopes da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[17] S Noachtar, C Binnie, J Ebersole, F Mauguiere, A Sakamoto, and B Westmoreland. A glossary of terms most commonly used by clinical electroencephalographers and proposal for the report form for the eeg findings. the international federation of clinical neurophysiology. *Electroencephalography and clinical neurophysiology. Supplement*, 52:21–41, 1999.

[18] Paul L Nunez and Ramesh Srinivasan. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.

[19] Nobutaka Ono, Kenichi Miyamoto, Jonathan Le Roux, Hirokazu Kameoka, and Shigeki Sagayama. Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram. In *2008 16th European Signal Processing Conference*, pages 1–4. IEEE, 2008.

[20] Ronald J Racine. Modification of seizure activity by electrical stimulation: Ii. motor seizure. *Electroencephalography and clinical neurophysiology*, 32(3):281–294, 1972.

[21] Mamunur Rashid, Norizam Sulaiman, Anwar PP Abdul Majeed, Rabiu Muazu Musa, Ahmad Fakhri Ab Nasir, Bifta Sama Bari, and Sabira Khatun. Current status, challenges, and possible solutions of eeg-based brain-computer interface: a comprehensive review. *Frontiers in neurorobotics*, page 25, 2020.

[22] Gerwin Schalk, Dennis J McFarland, Thilo Hinterberger, Niels Birbaumer, and Jonathan R Wolpaw. Bci2000: a general-purpose brain-computer interface (bci) system. *IEEE Transactions on biomedical engineering*, 51(6):1034–1043, 2004.

[23] Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. Deep learning with convolutional neural networks for eeg decoding and visualization. *Human brain mapping*, 38(11):5391–5420, 2017.

[24] J Ignacio Serrano, María Dolores Del Castillo, Verónica Cortés, Nuno Mendes, Aida Arroyo, Jorge Andreo, Eduardo Rocon, María Del Valle, Jaime Herreros, and Juan Pablo Romero. Eeg microstates change in response to increase in dopaminergic stimulation in typical parkinson's disease patients. *Frontiers in neuroscience*, 12:714, 2018.

[25] H Shibasaki, G Barrett, Elise Halliday, and AM Halliday. Components of the movement-related cortical potential and their scalp topography. *Electroencephalography and clinical neurophysiology*, 49(3-4):213–226, 1980.

[26] Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. Eeg conformer: Convolutional transformer for eeg decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2022.

[27] Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. Eeg conformer: Convolutional transformer for eeg decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:710–719, 2023.

[28] Abdulhamit Subasi. Eeg signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, 32(4):1084–1093, 2007.

[29] Robert W Thatcher, G Lyon, J Rumsey, and N Krasnegor. *Developmental neuroimaging: Mapping the development of brain and behavior.* Academic Press, 1996.

[30] Vanessa Valladares, Célio Pasquini, Silvana Carvalho Thiengo, and Clélia Christina Mello-Silva. Feasibility of near-infrared spectroscopy for species identification and parasitological diagnosis of freshwater snails of the genus biomphalaria (planorbidae). *Plos one*, 16(11):e0259832, 2021.

[31] Filipa Campos Viola, Jeremy Thorne, Barrie Edmonds, Till Schneider, Tom Eichele, and Stefan Debener. Semi-automatic identification of independent components representing eeg artifact. *Clinical Neurophysiology*, 120(5):868–877, 2009.

[32] Lonce Wyse. Audio spectrogram representations for processing with convolutional neural networks. *arXiv preprint arXiv:1706.09559*, 2017.

[33] Xinyi Yong, Rabab K Ward, and Gary E Birch. Sparse spatial filter optimization for eeg channel reduction in brain-computer interface. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 417–420. IEEE, 2008.

[34] Yuni Zeng, Hua Mao, Dezhong Peng, and Zhang Yi. Spectrogram based multi-task audio classification. *Multimedia Tools and Applications*, 78:3705–3722, 2019.

[35] Wei-Long Zheng, Jia-Yi Zhu, Yong Peng, and Bao-Liang Lu. Eeg-based emotion classification using deep belief networks. In *2014 IEEE international conference on multimedia and expo (ICME)*, pages 1–6. IEEE, 2014.