This dissertation is presented for the degree of Honours in Computer Science at The University of Western Australia

Multi-Class Seizure Classification with WaveNet and Bidirectional LSTM on EEG Signals

Stefan Nio

Supervised by:
Dr. Ghulam Mubashar Hassan
Professor Amitava Datta (Co-Supervisor)

Department of Computer Science and Software Engineering
School of Physics, Mathematics, and Computing

October 2024

To my parents for their endless love and encouragement,
to my caring siblings,
to my dear friends for their guidance and support,
to my devoted partner,
to my late grandmother, who could not see this thesis completed,
and to all who believed in me, thank you for your unwavering support throughout this journey.

Abstract

Epilepsy is a neurological disorder affecting millions of people worldwide, characterized by recurrent seizures which causes brief episodes of involuntary movement that may involve a part of the body or the entire body. Accurate detection and classification of seizures and their types are crucial for effective treatment and management of epilepsy. The goal of this research was to classify seven types of seizures from one another, including focal non-specific, generalized non-specific, simple partial, complex partial, absence, tonic, clonic, tonic-clonic, and atonic seizure just from electroencephalogram (EEG) signal data. However, the manual interpretation of EEG signals by neurologists is labor-intensive, time-consuming, and prone to human error, highlighting the need for automated and reliable diagnostic tools. In this paper, we present WaveNet-BiLSTM, an end-to-end dual-path seizure type classification deep learning network designed to classify seven types of seizures using raw EEG signal data as input, without the use of manual feature selection or extraction. The proposed model combines the strengths of WaveNet, which excels at capturing local patterns through dilated convolutions, and bidirectional long short-term memory (BiLSTM) networks with channel-wise attention mechanism, which effectively model long-term temporal dependencies and focus on the most informative EEG channels. We evaluated the model using the most updated Temple University Hospital EEG Seizure Corpus, TUSZ 2.0.3, to ensure a comprehensive and diverse dataset. The model's performance was assessed using both seizure-wise and patient-wise cross-validation methods. The proposed dual-path model achieved a weighted F1-score of 96% in seizure-wise validation and 63% in patient-wise validation. These results demonstrate the effectiveness of the proposed model in accurately classifying multiple seizure types and its potential to generalize across different patients. The integration of advanced deep learning architectures and the utilization of raw EEG data contribute to the development of more robust and generalizable diagnostic tools, ultimately improving patient care and treatment outcomes in epilepsy.
You can find this work on GitHub.

Alternatively, here is the full link: `https://github.com/stefanniouwa/SeizureNet`

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my saviour and my God. Without His strength, wisdom, and guidance, this work would not have been possible. Through countless challenges and moments of doubt, it was His grace and peace that gave me the perseverance to push through, guiding me towards clarity and helping me see the light at the end of every difficult road. I am truly grateful for His blessings that have led me to this point in my journey.

A special note of gratitude goes to my supervisor, Dr. Ghulam Mubashar Hassan, whose unwavering support and patience have been the cornerstone of this research. Through our many meetings and discussions, Dr. Hassan consistently provided invaluable insights, helping me refine my ideas and push the boundaries of what I thought was possible. His dedication to my progress and his constructive feedback have shaped not only this research but also my growth as a researcher. His commitment to excellence has been an inspiration, and I owe much of the success of this project to his mentorship.

I would also like to extend my thanks to Professor Amitava Datta for his contributions and guidance throughout this journey. His expertise and thoughtful input helped shape the direction of this research, and I am grateful for the time and effort he invested in providing feedback and encouragement along the way.

To my parents, words cannot express the depth of my gratitude. You have always instilled in me the value of striving for the best, encouraging me to chase my dreams no matter how distant they seemed. Your unwavering belief in my abilities and the sacrifices you have made are the foundation of who I am today. It is because of your endless love and guidance that I have had the strength and courage to pursue this path. I am forever indebted to you both for everything you have done and continue to do for me.

To my younger brothers, thank you for always pushing me to be the best version of myself. Your constant encouragement, whether through words or simply leading by example, has motivated me to keep improving and striving for greatness. The bond we share has been a source of strength, and I am grateful for the role you have played in my personal and academic journey.

To my late grandmother, Bobo Lan. Since I moved to Australia over seven years ago, I've had the privilege of living with you, and your presence has been a guiding light in my life. You nurtured me with love, wisdom, and care, and though you are no longer here to witness the completion of this work, I know you are watching over me. Your memory inspires me every day, and I dedicate a part of this thesis to you, in honor of the values and resilience you instilled in me.

I would also like to extend my heartfelt appreciation to my esteemed colleagues, friends, and my club members who stood by me throughout this journey. Your support, whether through academic discussions, shared experiences, or simple hang outs, made the hard days easier and the good days even better. The collaboration and friendships we built during this time have left a lasting impact, and I am grateful for every moment we shared, both in and out of campus.

To my partner, your unwavering support, love, and understanding have been my anchor throughout this journey. You have been by my side through every high and low, encouraging me to keep going even when the path seemed unclear. Your belief in me, especially during the most challenging times, has meant more to me than words can express. Thank you for your patience, your endless encouragement, and for being my constant source of strength.

I would also like to acknowledge The University of Western Australia for providing the academic and research environment that made this work possible. The resources, facilities, and support from the university were instrumental in allowing me to conduct my research at the highest level. The UWA community fostered a space of growth, curiosity, and innovation, and I am proud to have been a part of it.

Lastly, to everyone who has believed in me and supported me in any way, I extend my deepest thanks. Your encouragement and belief in me have made this journey worthwhile, and I am truly grateful to have had you all by my side.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Epilepsy is a neurological disorder characterized by recurrent seizures, affecting numerous individuals globally. Proper detection and classification of seizures are crucial for the effective treatment of this condition. Accurate diagnosis requires a thorough understanding of seizure types, typically achieved through meticulous analysis of electroencephalogram (EEG) data by qualified neurologists. However, limited access to neurologists, particularly in low- and middle-income countries, poses significant challenges in meeting the urgent need for accurate seizure classification [1]. The detailed examination of electrical brain activity to diagnose epilepsy is a time-consuming and often labor-intensive process that can be influenced by personal bias. Experts must carefully study EEG signals and correlate them with patient history, which may necessitate prolonged observations to identify distinct abnormalities. Given the costly and demanding nature of current diagnostic techniques, there is a pressing need for automated approaches to facilitate the early identification of epilepsy.

Severe seizures can have debilitating effects on a person's ability to function in daily life, leading to social isolation and diminished educational and occupational achievements [1]. This underscores the importance of accurately classifying seizure types to effectively manage and treat epilepsy. Precise classification is essential for guiding treatment decisions and delivering individualized care to those with specific seizure types. Healthcare professionals can improve patient outcomes and quality of life by customizing treatment approaches to target particular epilepsy subtypes. Additionally, diagnosing potential comorbidities linked to various seizure categories and forecasting prognosis depend on a comprehensive understanding of seizure types.

Electroencephalography (EEG) plays a vital role in neuroscience and clinical neurology, serving as the gold standard method in research for detecting and classifying seizures. This non-invasive technique involves placing scalp electrodes on the patient's head to capture the brain's electrical activity. The electrodes detect minute electrical potentials generated by neuronal activity, providing real-time insights into brain function [4]. Over the years, EEG technology has evolved from simple recording devices to complex digital systems that can be integrated with other neuroimaging modalities for in-depth analysis. The principles of EEG involve recording voltage fluctuations resulting from ionic current flows within neurons [4]. Noise and artifacts are minimized through filtering and amplification of the recorded signals. The resultant waveforms are displayed as continuous voltage-versus-time traces, typically divided into frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (>30 Hz) [12]. Specific frequency bands are associated with different states of brain activity, such as sleep, relaxation, or cognitive processes.

Seizures are mainly categorized into focal (or partial) and generalized seizures. Focal seizures originate in a particular region of the brain and can be further classified as simple focal seizures, which do not cause loss of awareness, or complex focal seizures, which impair awareness [15]. Generalized seizures include absence seizures, tonic-clonic seizures, myoclonic seizures, atonic seizures, and clonic seizures. Each type of seizure has distinct EEG patterns; for instance, absence seizures are associated with 3 Hz spike-and-wave discharges, while tonic-clonic seizures are characterized by slow waves followed by high-frequency activity [26]. Seizure types are classified by the International League Against Epilepsy based on symptom manifestation. Recognizing these distinct EEG patterns enables clinicians to accurately classify seizure types, which is essential for determining appropriate treatment strategies.

Different seizure types may require specific therapeutic interventions, emphasizing the importance of precise classification for effective patient care [13]. For example, medications effective for focal seizures might not be suitable for generalized seizures and vice versa. Accurate classification informs medical decisions and ensures that patients receive the most appropriate treatment, ultimately enhancing patient safety and improving the overall quality of life for those with epilepsy. Despite its effectiveness, manual interpretation of EEG data is time-consuming, labor-intensive, and subject to human error. The process demands extensive expertise and can be influenced by subjective biases, leading to potential inconsistencies in diagnosis [36]. Factors such as patient movement, electrode placement, and external artifacts can complicate the analysis, making accurate identification and classification of seizures challenging [36].

This paper focuses on the innovative use of deep learning technologies to restructure the classification of seizure types. Deep learning, a subset of machine learning, has demonstrated significant success in handling large datasets and extracting complex patterns, making it particularly suitable for EEG signal analysis. Moreover, end-to-end deep learning models typically require less human intervention, as they can manage the entire learning process from raw data preprocessing to model training in a more automated manner, eliminating the need for feature extraction and selection. While traditional machine learning has been widely and effectively utilized in seizure detection and classification, it typically requires detailed preprocessing of data before analysis [3]. This preprocessing often involves substantial expert input to extract relevant features from raw EEG data, which can be time-consuming and labor-intensive. In contrast, deep learning is capable of processing raw EEG data directly, thus potentially transforming the landscape of seizure detection and classification [33]. This capability not only streamlines the workflow but also enhances the efficiency of neurologists by allowing models to autonomously learn from raw data.

The field of EEG signal analysis for seizure classification highlights significant differences between traditional machine learning and modern deep learning approaches. Traditionally, EEG analysis has utilized feature-based machine learning, which requires thorough signal preprocessing to extract relevant features like frequency bands and time-domain characteristics [33]. This method depends on domain expertise and is well-suited to structured datasets but struggles with scalability and adaptability to new data [27]. In contrast, deep learning methods offer a more holistic approach by utilizing frameworks capable of learning representations directly from raw data, thereby removing the need for manual feature selection and extraction [27]. This shift is evident through the use of our proposed network utilizing a deep learning approach to process raw EEG signals and automatically learn the complex patterns associated with different seizure types. The advantage of this approach lies in its ability to adapt to the data's intrinsic properties without prior biases about which features are most relevant, potentially leading to the discovery of novel insights into seizure mechanisms.

The work of researchers like Albaqami [6] has been instrumental in the application of deep learning models, specifically his chapters on Multi-Path SeizureNet [9]for classification, as well as the use of WaveNet and LSTM for abnormal detection, to EEG-based classification problems [8]. Albaqami's research has demonstrated groundbreaking advancements in applying machine learning and deep learning techniques for seizure classification, achieving high classification performance using the Temple University Hospital Seizure Corpus (TUSZ v1.5.2) [28]. Notably, regarding patient-independent classification, he achieved relatively high scores through five-type classification, obtaining an 87.6% F1-score [9]. Patient-independent techniques involve training and evaluating models on data from different patients, ensuring that the models generalize well to new, unseen individuals. This work exemplifies cutting-edge research that leverages complex neural network architectures to enhance diagnostic accuracies across new patients.

The aim of this study is to create a deep learning model capable of accurately and robustly classifying multiple seizure types from raw EEG data, focusing on seven-class patient-independent classification to enhance generalizability. We hypothesize that employing advanced deep learning models like WaveNet and bidirectional long-short term memory (Bi-LSTM) on the latest TUSZ Corpus will significantly improve the accuracy and generalizability of multi-class seizure classification compared to traditional approaches. Building upon existing architectures, we explore how different combinations of deep learning models can be integrated for classifying seizure types using EEG signals. A dual-path seizure type classification deep learning network was trained on raw EEG data for epileptic seizure type classification. The proposed dual-path model comprises WaveNet and bidirectional LSTM components, each extracting complex patterns from raw EEG signals. The outputs of both paths are fused for the final classification of seizure types. We utilize the latest updated TUSZ version 2.0.3 (March 2024) for our seizure type classification, and to our knowledge, this is the first application of this dataset in this area of research.

By integrating these advanced deep learning techniques, we aim to enhance the performance of seizure classification models, particularly in patient-independent scenarios. Our approach seeks to address the limitations of manual EEG interpretation and traditional machine learning methods, offering a more efficient and potentially more accurate solution for seizure type classification. This could have significant implications for improving diagnostic precision and patient outcomes in epilepsy care.

CHAPTER 2

Literature Review

2.1 Review of Machine Learning and Deep Learning Applications in Seizure Classification

The implementation of machine learning and deep learning techniques has greatly improved seizure classification by enabling automation of EEG signal data analysis and decreasing the need for manual feature extraction. Traditionally, machine learning approaches for EEG analysis required domain experts to manually extract relevant features from the EEG signals, which is labor-intensive and time-consuming process. By reducing the need for personal involvement, automated machine and deep learning techniques seek to optimise this procedure, increasing seizure classification accuracy and efficiency.

Acquisition of data, preprocessing, feature extraction, classification, and performance evaluation are common steps in the systematic process of integrating machine learning and deep learning in EEG analysis, as previous research has demonstrated [6, 9, 8, 21, 7, 10, 34, 37, 18, 14, 31, 5, 32, 11, 38, 23]. This can be illustrated in Figure 2.1. Deep learning techniques enable automatic feature extraction from raw EEG signal data, whereas machine learning approaches typically require manual feature extraction and feature selection. Both methods classify the EEG recordings into various seizure types, after which the performance of the models is evaluated to determine their accuracy and precision.

Significant EEG signal features are manually retrieved and fed into classification algorithms in feature-based machine learning techniques. These techniques have a long history of being used effectively in a variety of EEG analyses, including brain–computer interface (BCI) technologies and the diagnostic evaluation of neurological illnesses [24]. For example, machine learning models have used feature-based methods for binary classification tasks to detect abnormalities in EEG signals. López et al.[19] achieved 78.8% accuracy in detecting abnormal EEG activity by using convolutional neural networks (CNNs) and multilayer

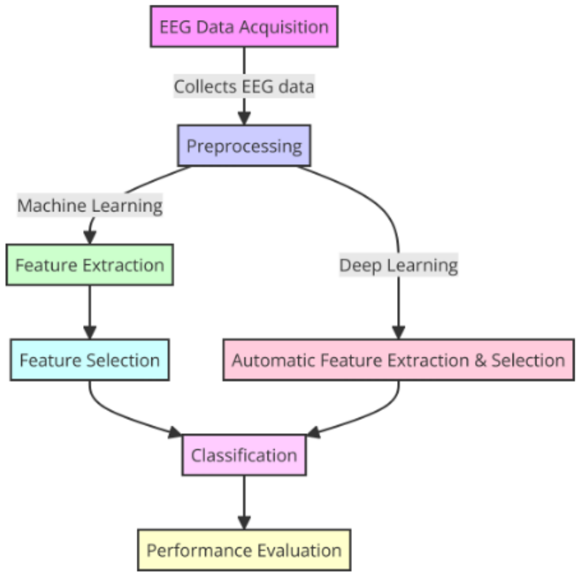


Figure 2.1: A typical machine learning and deep learning seizure type classification flow diagram

perceptrons (MLPs) in combination with band power-based features.

Numerous feature extraction methods, such as time-domain, frequency-domain, and time-frequency-domain approaches, have been used in seizure classification [29] [21] [30]. Wavelet transform (WT)-based feature extraction has become increasingly common among time-frequency methodologies since it can record both frequency and time information. Wavelet packet decomposition (WPD), continuous wavelet transform (CWT), and discrete wavelet transform (DWT) are often used approaches [7]. For instance, Alhussein et al.[10] used AlexNet and MLP to analyse fast fourier transform (FFT) band-limited signals and obtained an accuracy of 89.13%. Using an MLP with CWT for feature extraction, Saric et al. [34] achieved a 95.14% accuracy rate in a three-class seizure classification task. Similar to this, Wijayanto et al. [37] achieved a 95% classification accuracy by using support vector machines (SVMs) with empirical mode decomposition (EMD) and statistical features to categorise five-class seizure classification.

Feature-based techniques have several of disadvantages despite their effectiveness. The manual feature extraction procedure makes them labor-intensive and necessitates extensive subject knowledge. Moreover, these methods might be difficult to scale or modify for new, unseen data. When faced with complex or nuanced patterns in EEG signals, an over-reliance on predefined features might lead to inferior performance, which restricts the models' capacity to generalise.

By automating feature extraction and learning directly from raw EEG data, deep learning provides an innovative approach that uncovers intricate patterns in the signals. In particular, recurrent neural networks (RNNs) and convolutional neural networks (CNNs) have shown great promise in seizure classification [18]. Without the need for manual feature extraction, these models can analyse raw EEG data, which reduces the workload for experts and potentially increase classification accuracy.

Convolutional Neural Networks (CNNs) have been widely employed in seizure classification due to their proficiency in analysing spatial and temporal characteristics in EEG data. Liu et al.[25] utilised a bilinear model integrating a CNN and long short-term memory (LSTM) network to classify spectrogram pictures obtained from EEGs, attaining an F1 score of 97.4%. Shankar et al. [35] employed a CNN model utilising Gramian Angular Field Transformation (GAFT) to classify EEG signals into five-class seizure classifications, achieving accuracies between 84.19% and 96.01% across various classification scenarios.

Recurrent neural networks (RNNs) and LSTM networks are adept at sequence modelling and have shown effectiveness in identifying temporal patterns in EEG signals. Baghdadi et al. [14] developed a deep LSTM model including channel-wise attention to classify eight seizure types, attaining an F1 score of 98.41%. Priyasad et al. [31] employed an attention-based data fusion method to categorise seizures from raw EEG data, achieving an F1 score of 96.7%.

Hybrid models that integrate CNNs and LSTMs capitalise on the advantages of both architectures, improving the precision and dependability of seizure categorisation. Aristizabal et al. [5] devised a neural memory network (NMN) to classify seven seizure types, attaining an F1 score of 94.5%. Asif et al. [11] introduced SeizureNet, a deep learning framework that integrates a deep CNN and LSTM network, attaining a 95% F1 score at the seizure level.

Numerous studies have employed the Temple University Hospital Seizure (TUSZ) Corpus version 1.5.2 to train and evaluate machine learning and deep learning models, highlighting its significance in enhancing techniques for categorising seizure types. To our knowledge, the most recent version utilised in this research domain is TUSZ version 2.0.0 by Zhu et al. [38], mostly for seizure detection rather than classification. Numerous studies on seizure classification have predominantly utilised version 1.5.2. However, this dataset exhibited shortcomings, including patient overlap between the training and testing sets, as well as inconsistencies in seizure annotations [16]. These features may generate biases and diminish the generalizability of the models to novel cases.

Our research mitigates these deficiencies by utilising the enhanced TUSZ version 2.0.3, with the objective of improving model generalizability and performance via patient-independent assessments and uniform annotations. We aim to enhance prior research and address the constraints found in past studies by employing the most recent version of the dataset. Our methodology emphasises the creation of deep learning models that enhance generalisation for new patients and deliver more accurate seizure type classification, hence advancing diagnostic tools in epilepsy.

2.2 Motivations and Contributions

This research is motivated by the inherent constraints and challenges associated with manually interpreting EEG readings. This technique is both labor-intensive and time-consuming, as well as prone to human error and subjective prejudice. It is imperative to develop an automated system that is capable of accurately and reliably classify seizure types, while also being generalizable across various patient groups. In response to this requirement, our research seeks to create a system utilising advanced deep learning architectures—namely WaveNet and bidirectional LSTM networks—on the recently released Temple University Hospital Seizure (TUSZ) Corpus version 2.0.3.

The TUSZ v2.0.3 dataset, launched in March 2024, signifies a notable improvement compared to earlier iterations[28]. This work is, to our knowledge, the first application of this particular version for seizure classification tasks. The revised dataset [28] provides a more varied and extensive compilation of EEG recordings, encompassing a greater number of patients, as well as a higher rate of seizure events. The increased diversity enhances the training data accessible to the models, allowing them to learn more intricate and diverse patterns present in the EEG signals. As a result, the models developed are expected to be more robust and generalizable, and are proficient in appropriately classifying seizures across various patients.

Prior studies often show higher performance metrics in seizure-wise classification, whereby seizure occurrences are derived from patients, and the same patient's data may be present in both training and validation sets. For example, Albaqami[7] utilised LightGBM in conjunction with wavelet packet decomposition for seven-class seizure classification, attaining an F1-score of 89.6% in seizure-wise evaluation, though only 64% in patient-wise evaluation. Likewise, Asif et al.[11] created SeizureNet utilising characteristics derived from fast fourier transform (FFT), achieving a notable seizure-wise F1-score of 95%, which diminished to 62% in patient-wise evaluation. Jia et al. [23] employed a $ResNet18\_v2$ model alongside a variable weight CNN method and correlation matrix IBMFS, attaining a seizure-wise F1-score of 94% while achieving a patient-wise accuracy of merely 54.2%. These examples demonstrate that although these methods effectively localise seizure occurrences within EEG data for specific patients, they insufficiently account for the models' capacity to generalise to new, unseen patients. The inclusion of the same patient data in both training and validation phases may result in too optimistic performance metrics that fail to represent real-world clinical situations.

Consequently, our research emphasises optimising patient-wise (patient-independent) classification metrics, which more accurately evaluate a model's generalizability to novel patient groups post-training. This methodology reflects actual clinical environments, where models must precisely diagnose seizures in patients whose data were excluded from the training set. Prior research demonstrate that patient-independent classification is more problematic due to variations in EEG patterns among individuals and a lack of patient data. This intricacy may lead to significantly decreased performance measures in contrast to seizure-wise classification. Nonetheless, it offers a more accurate and significant assessment of the model's relevance in clinical practice, thereby enhancing patient care.

We focus on patient-wise classification to develop a robust and generalizable classifier for a seven-class seizure classification problem. In contrast to prior research that mostly depended on feature extraction from EEG signals for machine learning models, our methodology utilises raw EEG signals directly. By allowing the models to examine raw EEG data without manual feature extraction, we utilise the deep learning models' ability to independently detect and recognise complex patterns within the data. This reduces the workload on experts and removes possible biases from manual feature engineering, potentially revealing new insights on seizure causes.

The key contributions of this research are summarized as follows:

1. Utilization of the Latest TUSZ v2.0.3 Dataset: We are among the first to employ the newly released TUSZ v2.0.3 dataset for multi-class seizure classification. The dataset's increased diversity—with more patients, seizures, and seizure types—provides a richer training set, enabling the development of more robust and generalizable models.

2. Analysis of Raw EEG Signals Without Feature Extraction: Our models operate directly on raw EEG signals, eliminating the need for manual feature extraction. This approach allows the models to learn intricate patterns and representations inherent in the data autonomously, potentially capturing subtle features that manual methods might overlook.

3. Focus on Patient-Independent Results: By emphasizing patient-wise classification, we address the critical issue of model generalizability to new

patients. This focus aligns with real-world clinical scenarios, where models must perform accurately on unseen patient data, thereby providing a more realistic assessment of the model's effectiveness and applicability.

CHAPTER 3

Methodology

3.1 Dataset

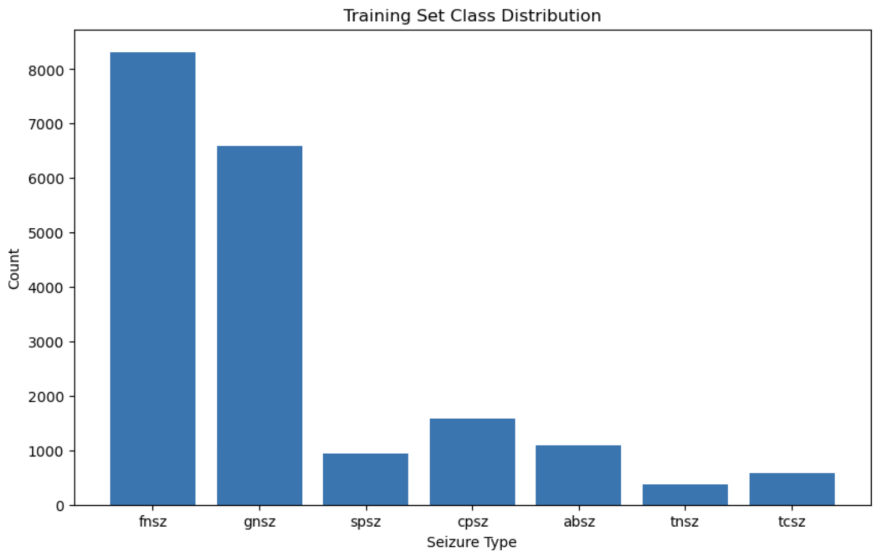


Figure 3.1: TUSZ version 2.0.3 Class Distribution

The Temple University Hospital EEG Data Corpus [28] is recognised as one of the most extensive publicly available datasets for research on seizure detection and classification using EEG data. This dataset, created and maintained by Temple University, comprises over 25,000 EEG recordings from more than 14,000 patients. This resource includes several seizure types, aiding researchers in the development and evaluation of novel algorithms for seizure detection and classification. The comprehensive annotations and metadata accompanying each recording facilitate analysis and efficient model training.

The TUH EEG Corpus is systematically categorised into subsets, each concentrating on specific elements of seizure study. The primary subset for this study,

Seizure Type	Definition	No. of Seizure Events		Duration (s)		No. of Patients	
		ver 1.5.2	ver 2.0.3	ver 1.5.2	ver 2.0.3	ver 1.5.2	ver 2.0.3
FNSZ	Focal seizures which cannot be specified with its type	1836	18,958	121,139	1,078,760	150	186
GNSZ	Generalized seizures which cannot be further classified into one of the groups below	583	23,717	59,717	1,822,190	81	120
CPSZ	Partial Seizures during unconsciousness; Type specified by clinical signs only	367	3597	36,321	281,774	41	39
ABSZ	Absence Discharges observed on EEG; patient loses consciousness for few seconds	99	2508	852	19,702	12	11
TNSZ	Stiffening of body during seizure	62	410	1204	11,485	3	4
TCSZ	At first stiffening and then jerking of body (Grand Mal)	48	857	5548	63,739	12	15
SPSZ	Partial seizures during consciousness; Type specified by clinical signs only	52	942	2146	31,985	3	2
MYSZ	Myoclonous jerks of limbs	3	44	1312	28,443	2	1

Table 3.1: Data distribution for different types of seizures in TUSZ version 1.5.2 and TUSZ version 2.0.3 (March 2024).

referred to as the TUH EEG Seizure Corpus (TUSZ), offers recordings classified in accordance with the standards established by the International League Against Epilepsy (ILAE) [20]. This subset encompasses various seizure types, including focal non-specific seizures (FNSZ), generalised non-specific seizures (GNSZ), simple partial seizures (SPSZ), complex partial seizures (CPSZ), absence seizures (ABSZ), tonic seizures (TNSZ), clonic seizures (CNSZ), tonic-clonic seizures (TCSZ), atonic seizures (ATSZ), and myoclonic seizures (MYSZ). The data points representing the distribution of each seizure class are illustrated in Figure 3.1.

Every recording is subjected to meticulous annotation by neurologists to guarantee high-quality and reliable data for research and clinical applications. The most recent iteration of TUSZ, version 2.0.3, launched in March 2024, includes significant enhancements. This dataset comprises approximately 51,000 seizure events from more than 670 patients, offering a substantial array of data for seizure analysis [28]. The dataset includes data from hundreds of thousands of sessions, with each session having one EDF file that includes the wave signal from each channel, alongside a corresponding CSV file that details the identified seizure type, including start and stop times for each channel. The specifics of the comparison between TUSZ version 1.5.2 and 2.0.3, together with the description of the chosen seizure types, are presented in Table 3.1

TUSZ v2.0.3 presents a substantial improvement over earlier versions by delivering a more generalizable, patient-independent evaluation, attributed to critical improvements implemented in this update. Earlier iterations, such as v1.5.2[16], had challenges with patient overlap across the training, development, and assessment datasets, potentially resulting in exaggerated performance metrics as the model was exposed to same patient data throughout both training and testing phases. [16] noted that one subject was shared between the development and evaluation sets, five subjects were shared between the development and training sets, and thirteen subjects were shared between the evaluation and training sets. While this overlap may not have significantly impacted performance for the technology of that era, it presented a possible concern for future developments[16]. To resolve this issue, [16] reconstructed the data partitions to eliminate the overlap, requiring the augmentation of the evaluation and development datasets with fresh subjects that had not been previously annotated, while preserving the subset sizes. Thus, TUSZ v2.0.3 guarantees that patients are mutually exclusive within the training, development, and evaluation sets, which is essential for assessing patient-wise classification through patient-independent validation methods. This clear distinction avoids the model's performance from being biassed towards particular patients, thereby facilitating a more precise evaluation of the model's generalizability to novel, unseen patient data, which is crucial for practical clinical applications.

TUSZ v2.0.3 adjusts prior anomalies in seizure annotations concerning frequency content. Previous iterations displayed inconsistencies, with certain files comprising solely low-frequency epileptiform events (2.5 Hz to 3 Hz)[16], but others included events with substantial frequency content above 3 Hz. This discrepancy required a comprehensive examination and reannotation of the entire corpus to guarantee consistency in seizure event annotations. The incorporation of additional epileptiform events with higher frequency content significantly enhanced the recorded seizure events, increasing from 673 in v1.5.2 to 1,239 in v1.5.3[16], thereby providing a more comprehensive dataset for the training and assessment of machine and deep learning models.

3.2 Preprocessing

The Temple University Hospital EEG Seizure Corpus (TUSZ) v2.0.3 displays diversity in montage configurations and sample rates, reflecting of its origin from real clinical environments. The recordings in this dataset [28] exhibit sampling rates between 250 Hz and 1024 Hz and utilise different montages. To maintain consistency and enable successful analysis, we executed several preprocessing procedures to standardise the input data before model training.

Initially, we concentrated on extracting EEG segments that specifically corresponded to seizure events. This was accomplished with the annotation files accompanying the dataset, which contain exact start and finish times for each seizure episode across various channels. Utilising these annotations, we isolated relevant segments from the continuous EEG recordings, ensuring our models are trained on data directly linked to seizure activity.

To highlight spike activity characteristic of seizures, we utilised the Transverse Central Parietal (TCP) montage [32]. An EEG montage is a differential representation of data, generated by calculating the difference between signals obtained from two specific electrodes (e.g., Fp1-F7, F7-T3), as outlined by the international 10-20 electrode placement scheme [22]. This differential method improves the identification of localised cerebral activity and assists in minimising noise. Neurologists are very cautious regarding the montage utilised in EEG interpretation, as it profoundly affects the visibility of neurological patterns [35].

The TCP montage has proven to be especially effective for seizure detection in numerous investigations [32]. It highlights spike and sharp wave activities, which are essential for recognising epileptiform patterns. Utilising the TCP montage, we computed the differential signals for the relevant electrode pairs, hence enhancing the characteristics relevant to seizure classification.

Due to the differing sampling rates in the TUSZ v2.0.3 dataset, standardisation was required for consistent data processing. All EEG recordings were standardised to a uniform sampling rate of 250 Hz. This sample rate is frequently employed in EEG analysis, offering a compromise between sufficient temporal resolution and computing efficiency [7, 8, 6, 9]. Resampling was performed utilising signal processing techniques that maintain the integrity of the original signals, guaranteeing that no substantial information was lost throughout the procedure.

Following to the application of the TCP montage and resampling, we partitioned each EEG recording into fixed-duration, non-overlapping intervals of two seconds. This segmentation produces segments comprising 500 data points each at a sampling rate of 250 Hz. The selection of a two-second window length is based on previous research that has determined it to be the ideal timeframe for recording seizure dynamics while ensuring computational effectiveness [32]. Segmenting the data into uniform chunks standardises the input for our algorithms and facilitates the identification of patterns within a clinically relevant timeframe.

We normalised each EEG segment for input into our neural network architecture. We computed the mean and standard deviation for each channel within a segment. The mean was deducted from the signal to centre it at zero, followed by division by the standard deviation to normalise amplitude variations. This standardisation is essential for neural networks, since it prevents any individual channel or segment from disproportionately affecting the model due to variations in scale.

Although the majority of segments typically comprised 500 data points following segmentation, there were occasions where segments were marginally shorter or longer due to the exact start and stop times of seizure events. We modified these segments to maintain uniformity across all input data.

• Signals Shorter than 500 Data Points: Zero-padding was applied to extend the signal to the required length.

• Signals Longer than 500 Data Points: The signal was truncated to fit the 500 data point requirement.

This adjustment ensures that all segments fed into the model have the same shape, which is essential for batch processing in deep learning frameworks.

A key aspect of our methodology is the direct use of raw EEG signals as inputs to our deep learning models. Unlike traditional approaches that rely on extensive feature extraction and handcrafted signal transformations, our method harnesses the power of deep learning architectures—specifically WaveNet and Bidirectional Long Short-Term Memory (Bi-LSTM) networks—to learn complex patterns directly from the normalized EEG data. This approach not only preserves the integrity of the original signals but also reduces preprocessing complexity and potential biases introduced by manual feature selection.

3.3 Proposed Model

3.3.1 WaveNet

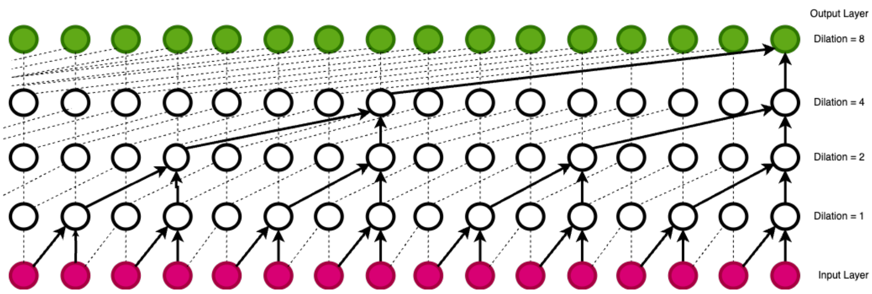


Figure 3.2: WaveNet Dilated Causal Convolution Layers

WaveNet, introduced by Google's DeepMind [2], is a deep generative model first designed for the generation of raw audio waveforms. It utilises dilated causal convolutions to effectively capture both short-term and long-term temporal dependencies in sequential data. WaveNet has attained significant success in speech synthesis and audio generation tasks due to its ability to simulate intricate temporal patterns without relying on recurrent connections [2].

In EEG signal processing, temporal dynamics are essential for precise seizure classification. EEG signals contain complex temporal patterns and long-range dependencies that are difficult to model using conventional frameworks. Recurrent neural networks (RNNs), including Long Short-Term Memory (LSTM) networks, are extensively utilised for sequential data but frequently encounter limitations in short-term memory while processing lengthy sequences [17]. We used the WaveNet architecture to analyse raw EEG signals, facilitating efficient feature extraction and enhancing performance in seizure classification.

The fundamental element of WaveNet consists of a series of one-dimensional convolutional layers characterised by exponentially increasing dilation rates [2]. A dilated convolution introduces gaps between kernel elements, effectively expanding the receptive field of the convolutional layer without increasing the parameter count. This design enables the network to effectively capture long-range temporal dependencies. Figure 3.2 illustrates the mechanism of dilated convolutional layers within the WaveNet architecture. The dilation rate d at each layer l is conventionally established to double with each subsequent layer, adhering to the sequence $d = 2^l$. The receptive field R of the network grows exponentially with the number of layers L:

$$R = (K-1) \times \left(\sum_{l=0}^{L-1} 2^l \right) + 1$$

where K is the kernel size.

WaveNet utilises gated activation units to regulate and shape the information flow inside the network. In each convolutional layer, the gated activation unit integrates the outputs of two convolution operations—one processed through a tanh activation function and the other through a sigmoid function:

$$z = \tanh(W_f * x) \odot \sigma(W_g * x)$$

where:

- W_f and W_g are the filter weights for the tanh and sigmoid convolutions, respectively;
- $*$ denotes the convolution operation;
- $\odot$ represents element-wise multiplication;
- x is the input to the layer.

Residual connections are utilised to enhance the training of deep networks by mitigating challenges including the vanishing gradient problem, speeding convergence, and enhancing overall performance.

In adapting WaveNet for EEG signal processing, we modified the original architecture to suit the characteristics of EEG data and the specific requirements of multi-class seizure classification. We have designed WaveNet layers with specific dilation rate sequences to capture both extensive and localised temporal patterns in the EEG signals. This method enables the model to acquire both overarching characteristics and intricate details crucial for distinguishing between different seizure types. In contrast to the original WaveNet, our implementation excludes skip connections, aligning with [8] to enhance control over the receptive field size and to facilitate a more effective capture of temporal information for our purpose.

3.3.2 Bidirectional LSTM

To further enhance the model's ability to capture temporal dependencies, we integrated a Bidirectional Long Short-Term Memory (BiLSTM) network pathway with a channel-wise attention mechanism. LSTMs are engineered to preserve long-term dependencies in sequential data using a memory cell architecture, incorporating gates that regulate the information flow. Every LSTM cell comprises three gates: the input gate, the forget gate, and the output gate. These gates govern the entry, storage, and exit of information within the cell state, thereby determining what to retain and what to discard [17].

The computations within an LSTM cell at time step t are as follows:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [x_t, h_{t-1}] + b_f), \\ i_t &= \sigma(W_i \cdot [x_t, h_{t-1}] + b_i), \\ o_t &= \sigma(W_o \cdot [x_t, h_{t-1}] + b_o), \\ g_t &= \tanh(W_g \cdot [x_t, h_{t-1}] + b_g), \\ C_t &= f_t \odot C_{t-1} + i_t \odot g_t, \\ h_t &= o_t \odot \tanh(C_t), \end{aligned}$$

where:

- f_t, i_t, and o_t are the forget gate, input gate, and output gate activations, respectively;
- g_t is the candidate cell state;
- C_t is the cell state at time t;
- h_t is the hidden state (also the output) at time t;
- W_* and b_* are the weight matrices and bias vectors for each gate;
- σ denotes the sigmoid activation function;
- $\odot$ represents element-wise multiplication;
- x_t is the input at time t;
- h_{t-1} is the hidden state from the previous time step.

Standard LSTMs process sequences in one direction, relying on prior information, however Bidirectional LSTMs enhance this by analysing input data in both the forward and backwards directions. This is accomplished by employing two distinct LSTM layers: one that analyses the sequence from past to future and another from future to past. The outputs of the forward and backwards LSTMs at each time step are combined, usually through concatenation:

$$h_t = \overrightarrow{h_t} \parallel \overleftarrow{h_t},$$

where $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ are the hidden states from the forward and backward LSTMs, respectively, and $\parallel$ denotes concatenation. This methodology enables the model to incorporate information from both past and future contexts, hence enhancing its capacity to discern temporal patterns within the EEG data.

In our model, the BiLSTM directly processes the raw EEG data, with inputs formatted as (time steps $\times$ EEG channels). We employ a BiLSTM layer with 64 units to capture temporal dependencies in both directions. To enhance the model's emphasis on the most informative features, we utilise a channel-wise attention mechanism subsequent to the BiLSTM layer. The attention mechanism computes attention scores for each channel, enabling the model to dynamically assess the significance of various channels.

Channel Wise Attention

The attention mechanism is applied to the output of the BiLSTM layer in our implementation. The BiLSTM analyses the input data and produces a sequence of hidden states $H \in \mathbb{R}^{T\times 2D}$, where T is the number of time steps, and $2D$ is the dimensionality of the hidden state (due to the concatenation of forward and backward passes in the BiLSTM).

However, in our context, since we are dealing with EEG data where the channels can be considered analogous to features at each time step, the BiLSTM output H is reshaped or interpreted as $H \in \mathbb{R}^{C\times D}$, where C represents the number of channels, and D is the hidden dimension.

The attention mechanism calculates attention scores for each channel to assess its relative importance. The calculation proceeds as follows:

1. **Compute Attention Scores**
   For each channel c, we compute an attention score e_c using a fully connected layer with a tanh activation function:

   $$e_c = \tanh(W_a h_c + b_a)$$

   where:

   - $h \in \mathbb{R}^D$ is the hidden state output from the BiLSTM for channel c.

- $W_a \in \mathbb{R}^{D\times 1}$ and $b_a \in \mathbb{R}^1$ are the trainable weights and bias of the attention layer.
- The tanh activation ensures non-linearity in the attention computation.

2. **Compute Attention Weights**
   The attention scores are then normalized across all channels using the softmax function to obtain the attention weights α_c:

$$\alpha_c = \frac{\exp(e_c)}{\sum_{k=1}^{C} \exp(e_k)}$$

   where α_c represents the normalized importance of channel c.

3. **Compute Context Vector**
   The context vector s is computed as the weighted sum of the BiLSTM hidden states, using the attention weights:

$$s = \sum_{c=1}^{C} \alpha_c h_c$$

   This context vector aggregates the most relevant information from all channels, emphasizing those with higher attention weights.

By integrating the channel-wise attention mechanism as executed in your code, our model efficiently utilises the spatial information across EEG channels. This method enables the model to identify the channels most essential for distinguishing different seizure types, potentially enhancing classification performance.

3.3.3 Combined Dual-Path Approach

The suggested combined dual-path technique is illustrated in Figure 3.3. Our hybrid model effectively captures the intricate temporal dynamics and spatial information inherent in EEG data by merging the outputs from both the modified WaveNet path and the BiLSTM with an attention mechanism. The combined features are further processed through dropout and batch normalisation layers to mitigate overfitting and enhance the stability of the learning process. The final classification is attained through fully connected layers that conclude with a softmax activation function, producing the probability distribution across the seizure classifications.

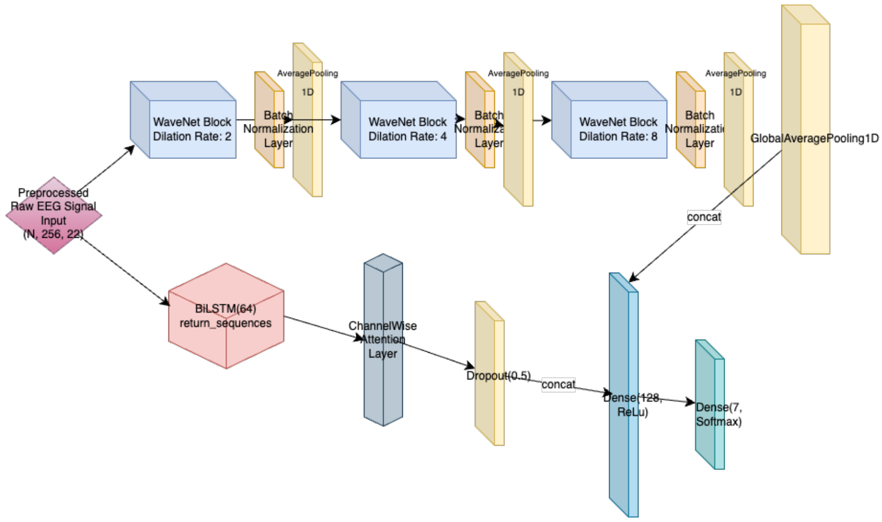


Figure 3.3: Combined Dual-Path Architecture: WaveNet and Bi-LSTM with Channel Wise Attention

Utilising the advantages of both the modified WaveNet and the BiLSTM with an attention mechanism enables our model to effectively analyse raw EEG signals without requiring significant feature extraction or manual signal manipulations. This method maintains the authenticity of the original signals and improves the model's capacity to discern complex patterns directly from the data, hence facilitating more precise and generalizable seizure classification.

CHAPTER 4

Results

This study utilised the Temple University Hospital Seizure Corpus (TUSZ) EEG dataset version 2.0.3 for seven-class seizure type classification. Table 3.1 outlines the statistics of the TUSZ dataset for various seizure types and the number of patients. We omitted Myoclonic (MYSZ) seizures from our research due to an insufficient amount of events for statistically significant analysis, with just three seizures recorded, as indicated in Table 3.1. This decision aligns with prior studies [32, 8, 9, 7]. We conducted preliminary preprocessing (3.2 Preprocessing) to the EEG data to obtain raw EEG signals, which served as inputs for our proposed deep learning models. We utilised three distinct architectures: a WaveNet model, a bidirectional long short-term memory (BiLSTM) model with channel-wise attention, and a hybrid dual-path model for multi-class classification.

4.1 Training and Performance Metrics

Evaluation Metrics

The TUSZ dataset demonstrated class imbalance and a skewed class distribution, with seizure types such as focal non-specific seizures (FNSZ) and generalised non-specific seizures (GNSZ) exhibiting a higher number of instances relative to the other classes. Figure 4.1 illustrates the distribution of seizure types throughout the dataset. Consequently, relying solely on accuracy would inadequately reflect the models' performance. Therefore, we utilised the weighted F1-score as our primary evaluation metric, in addition to accuracy, precision, recall, specificity, and sensitivity for each class. The weighted F1-score is computed as follows:

$$\text{Weighted F1-score} = \sum_{i=1}^{N} w_i \times \text{F1-score}_i$$

Where:

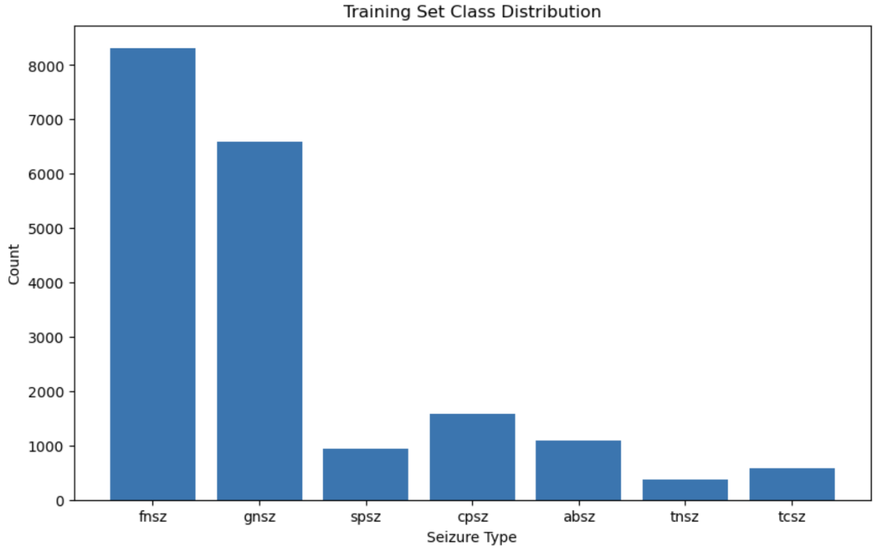


Figure 4.1: TUSZ v2.0.3 Seizure Type Distribution

- N is the number of classes,
- w_i is the proportion of instances belonging to class i,
- F1-score$_i$ is the F1-score for class i

We applied our technique to three models—the WaveNet model, the BiLSTM model with channel-wise attention, and the combined dual-path structure—and compared their performances.

Cross-Validation Strategies

To evaluate the generalizability and robustness of our models, we tested our technique using both seizure-wise and patient-wise cross-validation.

Seizure-Wise Cross-Validation

Drawing upon the techniques of [6] and [32], we utilised stratified 5-fold cross-validation for classification on a seizure-wise approach. The dataset was divided into five folds, ensuring each fold had a proportional representation of each class relative to the whole dataset. The model was trained on four folds and was evaluated on the remaining fold, iterating this procedure until each fold served as the evaluation set. This technique enables comprehensive evaluation across several data subsets and assists neurologists in accurately identifying seizure events within EEG signals across channels for an individual patient. The final performance metric was derived by calculating the average weighted F1-score across all five folds.

Patient-Wise Cross-Validation

We aimed to evaluate the model's generalizability to new patients following training on data from an earlier group. This technique, despite being constrained by a lack of patient data, guarantees the validity of the proposed automated system in real-world scenarios, where new data regularly originates from previously unseen patients. We utilised the dataset splits specified in the revised TUSZ v2.0.3 dataset, whereby the training, development, and evaluation sets are clearly indicated to have mutually exclusive patients in each directory[28]. This setup ensures that the data utilised for testing comes exclusively from different patients whose information has not been employed during the training phase. The comprehensive performance was assessed via the weighted F1-score across all folds.

Training Procedure

All three models—the WaveNet, the BiLSTM with channel-wise attention, and the combined dual-path model—were trained for 100 epochs with a batch size of 32. To combat overfitting, we implemented several regularization techniques:

- Early Stopping: Implemented with a patience of 10 epochs to prevent overfitting by halting training when the validation loss ceased to improve.
- L2 Regularization: Applied within the WaveNet blocks and the final dense layer to penalize large weights and promote generalization.
- Batch Normalization: Added after each WaveNet block to stabilize and accelerate training by normalizing layer inputs.
- Dropout Layers: Included in the BiLSTM path and after the concatenation in the combined model to prevent overfitting by randomly dropping units during training.

Categorical cross-entropy was used as the loss function, and the Adam optimizer was employed with an initial learning rate set to 0.001. We utilized a learning rate scheduler callback, where the learning rate was reduced during training based on a step decay function defined as:

$$L(e) = L_0 \times (d)^{\left\lfloor \frac{e+1}{N} \right\rfloor}$$

Where:

- $L(e)$ is the learning rate at epoch e

- L_0 is the initial learning rate
- d is the drop factor
- N is the number of epochs after which the learning rate drops
- $\lfloor \cdot \rfloor$ denotes the floor function, which returns the largest integer less than or equal to the input.

4.2 Experimental Results

To evaluate the effectiveness of our proposed dual-path method approach, we tested the performance of each submodel—the WaveNet and the BiLTSM with channel-wise attention—against the combined dual path model. Table 4.1 displays the seizure-wise performance metrics across the WaveNet, BiLSTM, and Proposed Dual-Path Method, while Table 4.2 evaluates these metrics in the patient-level.

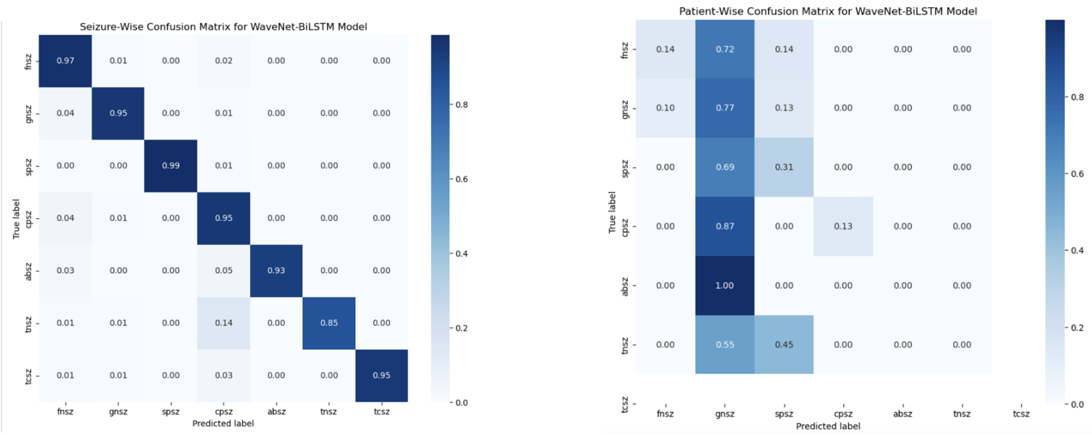
Table 4.1: Seizure Wise Performance Metrics

Model	**Avg F1-Score (%)**	**Avg Accuracy (%)**	**Avg Precision (%)**
WaveNet Path	86.6	86.6	89.7
BiLSTM Path	98.9	98.8	98.9
BiLSTM Path with Channel-Wise Attention	98.9	98.8	98.9
Proposed Dual-Path Method	96	96	96
Model	**Avg Recall (%)**	**Avg Sensitivity (%)**	**Avg Specificity (%)**
WaveNet Path	87.6	84.1	97.5
BiLSTM Path	98.9	99	99
BiLSTM Path with Channel-Wise Attention	98.9	99	99
Proposed Dual-Path Method	96	94	99

Table 4.2: Patient Wise Performance Metrics

Model	**Avg F1-Score (%)**	**Avg Accuracy (%)**	**Avg Precision (%)**
WaveNet Path	61.2	58.2	68.9
BiLSTM Path	37.8	37.7	69.3
BiLSTM Path with Channel-Wise Attention	46.7	45.2	71
Proposed Dual-Path Method	63	60.6	64.6
Model	**Avg Recall (%)**	**Avg Sensitivity (%)**	**Avg Specificity (%)**
WaveNet Path	58.2	20.7	88.6
BiLSTM Path	37.7	32.5	85.2
BiLSTM Path with Channel-Wise Attention	45.2	29.1	86.2
Proposed Dual-Path Method	60.6	22.4	84

Seizure Wise Performance Results



(a) Seizure-wise Confusion Matrix (b) Patient-wise Confusion Matrix

Figure 4.2: Confusion matrices of 7-class seizure types for seizure-wise and patient-wise validation techniques using the Proposed Dual-Path approach

Both WaveNet and BiLSTM with Channel-Wise Attention submodels performed generally well. The WaveNet model attained an F1-score of 86.6%, but the BiLSTM model earned an F1-score of 98.9%. The suggested dual-path network achieved an F1-score of 96%, just falling short of the BiLSTM model by 2%. The performance of the BiLSTM model is similar from that of the BiLSTM model incorporating a Channel-Wise Attention Mechanism at the seizure level. The superior performance of a model utilising BiLSTM demonstrates its effectiveness in classifying various seizure types.

The BiLSTM models, both with and without Channel-Wise Attention, demonstrated superior performance in average precision, recall, sensitivity, and specificity. Both models had average precision and recall rates of 98.9%, signifying their robust ability to accurately detect seizure events while reducing misclassification. The implementation of the Channel-Wise Attention Mechanism did not significantly affect the performance metrics of the model. This indicates that the BiLSTM model individually excels in capturing temporal dependencies within the data.

In comparison to our dual-path design, the proposed model attained an average precision and recall of 96%, which, although inferior to the two BiLSTM models, nevertheless demonstrates commendable performance overall. Its average sensitivity of 94% and specificity of 99% demonstrate its robust capacity to differentiate between seizure and non-seizure events. Conversely, the WaveNet path trailed behind the other models, achieving an average precision of 98.7% and a recall of 87.6%. It attained an average sensitivity of 84.1% and specificity of

97.5%, indicating less effectiveness in identifying actual seizure events compared to BiLSTM-based models.

The confusion matrix for seizure-wise classification displayed in Figure 4.2a indicates that the Proposed Dual-Path model achieved high accuracy for across seizure types, displaying low misclassification rates. Focal non-specific seizures (FNSZ) were classified with 97% accuracy, generalised non-specific seizures (GNSZ) with 95% accuracy, and simple partial seizures (SPSZ) with 99% accuracy. Absence seizures (ABSZ) showed a slight reduction in performance, achieving 93% accuracy, whereas tonic seizures (TNSZ) attained an accuracy of 85%, suggesting challenges in differentiating between these seizure types.

Patient Wise Performance Results

In patient-wise classification, all examined models exhibit a decline in performance, as illustrated in Table 4.2. This may be anticipated due to the difficulty of accurately classifying seizure types in novel patient data. The proposed dual-path model attained a weighted F1-score of 63%, surpassing the individual WaveNet model's score of 61.2% by around 2%. Our BiLSTM with Channel-Wise attained a 46.7% F1-score at the patient level. The channel-wise attention layer significantly enhances the performance of seizure type classification, yielding an 8.9% improvement over the BiLSTM model that does not emphasise individual channels. This enhancement demonstrates that combining the local feature extraction strength of WaveNet with the temporal modelling capabiltieis of BiLSTM with channel-wise attention, improves the model's ability to generalise to unseen patient data.

The proposed dual-path approach and WaveNet path outperformed their BiLSTM counterparts in average recall, achieving 60.6% and 58.2%, respectively, compared to 37.7% and 45.2% for the BiLSTM Path and BiLSTM Path with Channel-Wise Attention. This indicates that the Dual-Path and WaveNet models were more effective in detecting actual seizure events in novel patients.

Nonetheless, the BiLSTM models demonstrated superior average precision rates in comparison to alternative models. Among these models, the BiLSTM Path with Channel-Wise Attention attained an average precision score of 71%. This suggests that while they were more conservative in predicting seizures (resulting in fewer false positives), they also missed more true seizure events (lower recall). The Proposed Dual-Path Method achieved a balance between precision and recall scores, attaining 64.6% and 60.6%, respectively, indicating a more consistent performance across both evaluation measures.

The tendencies are further emphasised by the measurements of sensitivity and specificity. Regarding sensitivity, which measures the true positive rate, the models demonstrate relatively low performance in patient-wise classification, attaining scores of 22.4% for the Proposed Dual-path Method and 20.7% for the WaveNet Path model. This indicates challenges in accurately classifying seizures in novel, unseen patient data. On the other hand, the models show higher average specificity values, with 88.6% for the WaveNet Path model and 84% for the Proposed-Dual Path Method. This signifies the ability to accurately differentiate non-seizure events.

The confusion matrix for patient-wise classification (Figure 4.2b) underscores the significant challenges that all models face in generalizing to unseen patients. Specifically, the WaveNet model exhibited a high misclassification rate for focal non-specific seizures (FNSZ), with only 14% of these seizures being correctly classified, while 72% were incorrectly classified as generalized non-specific seizures (GNSZ). Notably, the model achieved its highest performance with GNSZ, correctly classifying 77% of these cases. However, the other six seizure types were predominantly misclassified as GNSZ. For instance, 87% of complex partial seizures (CPSZ) were incorrectly identified as GNSZ. Additionally, absence seizures (ABSZ), tonic seizures (TNSZ), and tonic-clonic seizures (TCSZ) were not correctly classified at all. This pattern indicates a tendency of the model to overgeneralize and default to predicting GNSZ when confronted with unfamiliar seizure patterns in new patients. These results highlight the difficulty in capturing the diverse characteristics of different seizure types across various patients.

CHAPTER 5

Discussion

This study aimed to evaluate the effectiveness of our proposed dual-path technique, which combines a WaveNet submodel and a BiLSTM submodel with channel-wise attention, for classifying seizure types using EEG data. We evaluated the performance of each submodel separately and contrasted them with the integrated dual-path model. The outcomes were evaluated utilising both seizure-wise and patient-wise metrics, and we compared our findings with prior studies to contextualise our contributions.

Our WaveNet submodel attained an average F1-score of 86.6% for seizure-wise classification and 61.2% for patient-wise classification. Prior research has investigated the application of WaveNet topologies for seizure detection based on EEG data. [8] utilised WaveNet for binary categorisation of EEG signals derived from the TUAB (Temple University Hospital Abnormal Corpus), attaining an accuracy of around 88.7%. [8] concentrated on differentiating between normal and abnormal EEG signals by analysing the initial 30 seconds of each recording.

In contrast to previous efforts, our WaveNet submodel extends the use of WaveNet to a more complex, multi-class seizure type classification problem. The seizure-wise performance illustrates the model's proficiency in effectively capturing local temporal aspects. The decline in accuracy in patient-specific classification implies difficulties in generalising to new patients, a prevalent issue due to inter-patient variability in EEG patterns.

The BiLSTM submodel incorporating channel-wise attention attained an average F1-score of 98.9% for seizure-wise classification and 46.7% for patient-wise classification. [14] previously developed a channel-wise attention-based LSTM model for seizure classification, attaining an F1-score of 96.87% in seizure-wise evaluation. A separate study, [9], attained a 97.01% F1-score in five-class seizure classification utilising a hybrid CNN-BiLSTM model on wavelet-processed EEG signals. Our BiLSTM model not only exceeds this performance, attaining an F1-score of 98.9% in 7-class seizure-wise classification, but also accomplishes this with a less complex network architecture and by utilising solely raw EEG inputs.

The high seizure-wise performance of our BiLSTM submodel demonstrates its effectiveness in capturing temporal relationships and prioritising the most valuable channels within the EEG data. In contrast to other studies that employed attention mechanisms separately for each channel, our method integrates the attention mechanism to identify dominant features across all channels, hence augmenting the model's discriminative skills.

In patient-wise classification, the BiLSTM submodel's performance diminishes to an F1-score of 37.8%, reflecting trends observed in previous studies where models demonstrate reduced accuracy on unfamiliar patient data. Nonetheless, our BiLSTM, with a channel-wise attention layer, dramatically enhanced the performance, increasing it by 8.9% to 46.7%. This boost illustrates the effectiveness of attention mechanisms in prioritising the most informative EEG channels, thus enhancing the model's ability to generalise across patients. The attention mechanism allows the model to weigh the contribution of each channel dynamically, emphasizing channels that are more indicative of seizure activity for different patients. This adaptability is essential for addressing the variability of EEG patterns among the patient group.

Our proposed dual-path methodology integrates the advantages of the WaveNet and BiLSTM submodels to enhance generalisation in seizure classification tasks. The dual-path model attained an average F1-score of 96% in seizure-wise classification, closely matching the performance of the BiLSTM submodel. Furthermore, it attained an average F1-score of 63% in patient-wise classification, surpassing both individual submodels and numerous models from prior research. For example, [11] attained a patient-specific F1-score of 62% utilising SeizureNet with Fast Fourier Transform (FFT) features, whilst our model reached 63%. Other models, such the K-Nearest Neighbours (KNN) and XGBoost classifiers utilising FFT features, displayed low patient-wise F1-scores of 40.1% and 54.2%, respectively [32]. However, [7] reported a superior patient-level F1-score of 64% for a seven-class seizure classification test utilising wavelet-based feature extraction techniques. They attained F1-scores of 64%, 63.9%, and 63.3% utilising Wavelet Packet Decomposition (WPD), Dual-Tree Complex Wavelet Transform (DT-CWT), and Discrete Wavelet Transform (DWT), respectively. Wavelet-based feature extraction can enhance model performance by accurately capturing both temporal and spectral features of EEG signals, hence facilitating improved representation of seizure patterns. This method can improve the model's capacity to generalise to new patients by emphasising variables that are more resilient to inter-patient variability than raw EEG signals. Variations in datasets, preprocessing techniques, and evaluation protocols may influence these comparisons; still, the findings suggest that our dual-path approach is comparable to existing state-of-the-art models.

The enhancement in individual patient performance indicates that the combination of local feature extraction (WaveNet) with temporal modelling and attention mechanisms (BiLSTM with channel-wise attention) improves the model's capacity to generalise to unfamiliar patients. The dual-path technique proficiently captures both local patterns and long-term dependencies in EEG signals, essential for precise seizure type categorisation.

Despite the advancements, the suggested dual-path methodology continues to encounter obstacles. In patient-wise classification, the sensitivity is significantly lower at 22.4%, signifying challenges in accurately identifying all true seizure events in new patients. This limitation highlights the persistent issue in EEG analysis about inter-patient variability and indicates that additional refining is required to improve sensitivity and overall performance.

The computational complexity of this dual-path architecture may increase training duration and resource demands relative to single-path models. The BiLSTM submodel with channel-wise attention demonstrates superior performance; nonetheless, it is susceptible to overfitting, as evidenced by its diminished patient-wise performance, necessitating larger datasets or further regularisation methods to enhance generalisation.

Although improvements in machine learning present considerable promise for enhancing seizure classification, it is essential to maintain collaborative integration between AI seizure classification tools and experienced neurologists to guarantee effectiveness in real-world usage. Interpreting machine learning models, particularly those utilised in decision-making within healthcare, is crucial for ensuring alignment with established medical principles and practices. This alignment not only supports the models' effectiveness in relation to physicians' clinical expertise but also promotes acceptance and trust among the medical community. Moreover, ethical considerations including responsibility, data privacy, and prediction biases are crucial when implementing AI models in healthcare. Utilising formal verification methods to evaluate the reliability and resilience of these systems can improve their safety and trustworthiness, especially in vital domains such as healthcare, where errors may yield substantial repercussions. Future study should to concentrate on incorporating these methodologies into the advancement of machine learning applications to enhance their dependability and promote their implementation in healthcare environments.

Table 5.1: Performance Comparison of Various Methods on 7-class Seizure Classification

Method	**Features**	**Seizure Wise Performance (%)**	**Patient Wise Performance (%)**
SeizureNet [11]	Fast Fourier Transform	95 F1-score	62 F1-score
KNN [32]	Fast Fourier Transform	90.1 F1-score	40.1 F1-score
XGBoost [32]	Fast Fourier Transform	85.1 F1-score	54.2 F1-score
SGD [32]	Fast Fourier Transform	80.7 F1-score	46.9 F1-score
Plastic Neural Memory Network (NMN) [5]	Correlation Matrix IBMFS	94.5 F1-score	-
LightGBM [7]	Wavelet Packet Decomposition	89.6 F1-score	64 F1-score
ResNet18_v2 with variable weight CNN algorithm [23]	Correlation Matrix IBMFS	94 F1-score	54.2 Accuracy
LSTM [14]	- (no features)	96.7 F1-score	-
Proposed Dual-Path	- (no features)	96 F1-score	63 F1-score

CHAPTER 6

Conclusion

This research evaluated a dual-path neural network that combines a WaveNet submodel and a BiLSTM submodel with channel-wise attention for the classification of seven seizure types using EEG data. We conducted classification directly on raw EEG signals by employing a new dataset and innovative data preparation approaches, bypassing conventional feature extraction methods. Our dual-path methodology attained an average F1-score of 96% in seizure-level classification and 63% in patient-level classification. The results indicate successful classification of seizure types among patients and encouraging generalisation to novel patients, notwithstanding difficulties arising from inter-patient variability. Although not entirely nove, our results highlight the possibility of utilising raw EEG data without feature extraction, hence streamlining the data processing pipeline. The integration of WaveNet and BiLSTM architectures, supplemented by channel-wise attention, improved generalisation relative to individual submodels. Our research contributions to the field by demonstrating the usefulness of the dual-path architecture, highlighting the promise of innovative data preprocessing, and providing insights into model generalization challenges. Future work may concentrate on improving patient-specific sensitivity through the integration of larger and more heterogeneous datasets, applying data augmentation techniques, and the optimisation of the model to more effectively manage variability among patients.

Bibliography

[1] Diagnosis of epilepsy a neurological disorder using electroencephalogram (eeg). *International Journal of Modern Trends in Engineering & Research 4* (06 2017), 144–149.

[2] , A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., and Kavukcuoglu, K. Wavenet: A generative model for raw audio, 2024.

[3] Abbasi, B., and Goldenholz, D. M. Machine learning applications in epilepsy. *Epilepsia 60* (09 2019), 2037–2047.

[4] Action, E. About seizures, 2022.

[5] Ahmedt-Aristizabal, D., Fernando, T., Denman, S., Petersson, L., Aburn, M. J., and Fookes, C. Neural memory networks for seizure type classification. *QUT ePrints (Queensland University of Technology)* (07 2020).

[6] Albaqami, H. Leveraging machine learning for enhanced detection and classification of brain pathologies using eeg. *the UWA Profiles and Research Repository* (2023).

[7] Albaqami, H., Hassan, G. M., and Datta, A. Wavelet-based multi-class seizure type classification system. *Applied Sciences 12* (06 2022), 5702–5702.

[8] Albaqami, H., Hassan, G. M., and Datta, A. Automatic detection of abnormal eeg signals using wavenet and lstm. *Sensors 23* (06 2023), 5960–5960.

[9] Albaqami, H., Hassan, G. M., and Datta, A. Mp-seiznet: A multi-path cnn bi-lstm network for seizure-type classification using eeg. *Biomedical Signal Processing and Control 84* (03 2023), 104780–104780.

[10] Alhussein, M., Muhammad, G., and Hossain, M. S. Eeg pathology detection based on deep learning. *IEEE Access 7* (2019), 27781–27788.

[11] Asif, U., Roy, S., Tang, J., and Harrer, S. Seizurenet: Multi-spectral deep feature learning for seizure type classification, 2024.

[12] Attar, E. T. Review of electroencephalography signals approaches for mental stress assessment. *Neurosciences 27* (10 2022), 209–215.

[13] Aylward, R. L. Epilepsy: a review of reports, guidelines, recommendations and models for the provision of care for patients with epilepsy. *Clinical Medicine 8* (08 2008), 433–438.

[14] Baghdadi, A., Fourati, R., Aribi, Y., Daoud, S., Dammak, M., Mhiri, C., Chabchoub, H., Siarry, P., and Alimi, A. A channel-wise attention-based representation learning method for epileptic seizure detection and type classification. *Journal of Ambient Intelligence and Humanized Computing 14* (05 2023), 9403–9418.

[15] Bereda, G. Seizures are the main sign of epilepsy: Stages of seizures. *Pathology and Laboratory Medicine* (02 2023).

[16] Buckwalter, G., Chhin, S., Rahman, S., Obeid, I., and Picone, J. Recent advances in the tuh eeg corpus: Improving the interrater agreement for artifacts and epileptiform events. In *2021 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)* (2021), pp. 1–3.

[17] Chair, D. Long short-term memory — bibsonomy, 2016.

[18] Craik, A., He, Y., and Contreras-Vidal, J. L. Deep learning for electroencephalogram (eeg) classification tasks: a review. *Journal of Neural Engineering 16* (04 2019), 031001.

[19] Diego, L. d., and Isabel, S. Automated interpretation of abnormal adult electroencephalograms.

[20] Fisher, R. S., Cross, J. H., French, J. A., Higurashi, N., Hirsch, E., Jansen, F. E., Lagae, L., Moshé, S. L., Peltola, J., Roulet Perez, E., Scheffer, I. E., and Zuberi, S. M. Operational classification of seizure types by the international league against epilepsy: Position paper of the ilae commission for classification and terminology. *Epilepsia 58* (03 2017), 522–530.

[21] He, J., Cui, J., Zhang, G., Xue, M., Chu, D., and Zhao, Y. Spatial–temporal seizure detection with graph attention network and bidirectional lstm architecture. *Biomedical Signal Processing and Control 78* (06 2022), 103908–103908.

[22] Jasper, H. H. The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology 10* (01 1958), 371–375.

[23] Jia, G., Lam, H.-K., and Althoefer, K. Variable weight algorithm for convolutional neural networks and its applications to classification of seizure phases and types. *Pattern Recognition 121* (08 2021), 108226–108226.

[24] Kevric, J., and Subasi, A. Comparison of signal decomposition methods in classification of eeg signals for motor-imagery bci system. *Biomedical Signal Processing and Control 31* (01 2017), 398–406.

[25] Liu, T., Truong, N. D., Nikpour, A., Zhou, L., and Kavehei, O. Epileptic seizure classification with symmetric and hybrid bilinear models. *IEEE Journal of Biomedical and Health Informatics 24* (10 2020), 2844–2851.

[26] Moeller, F., LeVan, P., Muhle, H., Stephani, U., Dubeau, F., Siniatchkin, M., and Gotman, J. Absence seizures: Individual patterns revealed by eeg-fmri. *Epilepsia 51* (08 2010), 2000–2010.

[27] Nafea, M. S., and Ismail, Z. H. Supervised machine learning and deep learning techniques for epileptic seizure recognition using eeg signals—a systematic literature review. *Bioengineering 9* (12 2022), 781.

[28] Obeid, I., and Picone, J. The temple university hospital eeg data corpus. *Frontiers in neuroscience 10* (05 2016).

[29] O.K., F., and R., R. Time-domain exponential energy for epileptic eeg signal classification. *Neuroscience Letters 694* (02 2019), 1–8.

[30] Polat, K., and Güneş, S. Classification of epileptiform eeg using a hybrid system based on decision tree classifier and fast fourier transform. *Applied Mathematics and Computation 187* (10 2006), 1017–1026.

[31] Priyasad, D., Fernando, T., Denman, S., Sridharan, S., and Fookes, C. Interpretable seizure classification using unprocessed eeg with multi-channel attentive feature fusion. *IEEE Sensors Journal 21* (09 2021), 19186–19197.

[32] Roy, S., Asif, U., Tang, J., and Harrer, S. Seizure type classification using eeg signals and machine learning: Setting a benchmark. *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)* (12 2020).

[33] Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., and Faubert, J. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering 16* (05 2019).

[34] Sarić, R., Jokić, D., Beganović, N., Pokvić, L. G., and Badnjević, A. Fpga-based real-time epileptic seizure classification using artificial neural network. *Biomedical Signal Processing and Control 62* (09 2020), 102106.

[35] Shankar, A., Dandapat, S., and Barma, S. Seizure type classification using eeg based on gramian angular field transformation and deep learning. *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (11 2021).

[36] Wang, Y., Zibrandtsen, I. C., Lazeron, R. H., Dijk, v., Long, X., Aarts, R. M., Wang, L., and Arends, J. B. Pitfalls in eeg analysis in patients with non-convulsive status epilepticus. *medRxiv (Cold Spring Harbor Laboratory)* (10 2020).

[37] Wijayanto, I., Hartanto, R., Nugroho, H. A., and Winduratna, B. Seizure type detection in epileptic eeg signal using empirical mode decomposition and support vector machine. In *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)* (2019), pp. 314–319.

[38] Zhu, Y., and Wang, M. D. Automated seizure detection using transformer models on multi-channel eegs. In *2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)* (2023), pp. 1–6.