Deep Learning Approaches to EEG and fMRI Data: A Comparative Study for Sleep Stage Classification

by

Farhan Tanvir
22241051
Tanjilur Rahman
19101033
S M Arfa Kamal
22341062
Mahmudul Hassan
23341137
Nowshin Nazia
19201093

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
May 2024

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Farhan Tanvir

Farhan Tanvir
22241051

Tanjilur Rahman

Tanjilur Rahman
19101033

S M Arfa Kamal

S M Arfa Kamal
22341062

Mahmudul Hassan

Mahmudul Hassan
23341137

Nowshin Nazia

Nowshin Nazia
19201093

Approval

The thesis/project titled "Deep Learning Based Inspection of fMRI and EEG for Sleep Stage Detection: A Comparison Analysis and Ablation Study" submitted by

1. Farhan Tanvir (22241051)
2. Tanjilur Rahman (19101033)
3. S M Arfa Kamal (22341062)
4. Mahmudul Hassan (23341137)
5. Nowshin Nazia (19201093)

Of Spring, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on May 21, 2024.

Examining Committee:

Supervisor:
(Member)

Nabuat Zaman Nahim
Lecturer
Department of Computer Science and Engineering
BRAC University

Co-Supervisor:
(Member)

Tanzim Reza
Lecturer
Department of Computer Science and Engineering
BRAC University

Ethics Statement

This thesis, titled "Deep Learning Approaches to EEG and fMRI Data: A Comparative Study for Sleep Stage Classification," adheres to ethical standards for human data research. All fMRI and EEG data used are from publicly available, anonymized datasets or from studies with explicit participant consent for research use. Data privacy and confidentiality have been strictly maintained, and all procedures comply with relevant data protection regulations. No identifiable personal information was accessed during the research.

Abstract

In this thesis, for classification of sleep stages, we use deep learning techniques with the help of data from fMRI and EEG. The ConvLSTM models are applied for the fMRI data. The data for the EEG is worked on with the LSTM and Bidirectional LSTM. This can hence be seen as a work of optimizing the accuracy, the precision, and the generalizability of all these models with one another. The baselines for all these different types of data are built up using initial models. The LSTM baseline model has given an accuracy of 78.69% on testing for sleep staging with W (Wake), NREM-1, NREM-2, and NREM-3 using EEG data, which is highly effective with such data resolution in time. Meanwhile, the Bidirectional LSTM model performs better preprocessing for the temporal aspect and hence yields, on average, 80.60% accuracy for general classification on the same stages. This would make it a model that can capture the dynamic nature of the EEG data across these particular stages. In contrast, processed fMRI data starts with a 76.82% testing accuracy, while performance is readjusted based on the feature extraction spatial-temporal settings adopted in their ConvLSTM configurations to classify sleep stages W, NREM-1, NREM-2, and NREM-3, with special attention to the role of model configuration. These results show that functionalities of tailored deep learning play the most basic role in the high complexity domain of sleep stage classification. These findings are prerequisite for the future development of this area.

Keywords: Deep learning, fMRI Data, EEG Data, Sleep Stage Detection, ConvLSTM Model, LSTM Model, Bidirectional LSTM Model, Comparative Analysis, Model Architectures.

Acknowledgement

Firstly, we express our utmost gratitude to Allah, whose blessings have enabled us to complete our thesis without any significant interruptions. Secondly, we extend our heartfelt thanks to our supervisor, Mr. Nabuat Zaman Nahim, and our co-supervisor, Mr. Tanzim Reza, for their invaluable support and guidance throughout our work. Their assistance was indispensable whenever we encountered difficulties. Lastly, we are deeply grateful to our parents for their unwavering support and prayers. Their encouragement has been crucial in bringing us to the threshold of our graduation.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Significance

Sleep, as a fundamental physiological process, is conspicuously involved in the course of development of our general condition and cognitive activity. This complex phenomenon embraces the following stages of sleep, which are accompanied by the typical waves of electrical activity of the brain. Proper classification and monitoring of various sleep stages have a crucial role in understanding the pathology of sleep for proper clinical diagnoses and personalized treatment strategies. This becomes all the more relevant in cases involving patients with coma. This would help everybody to understand the neurological stage and ability to recuperate. Besides, an advancement in sleep stage monitoring among such comatose persons has unraveled brain activity and sensitivity, previously unrealized, thereby assisting attending practitioners with altering treatment regimens and, in effect, forecasts about the patient. Conventional or manual scoring is still the gold standard when classifying the sleep stages and makes the entire exercise laborious besides being prone to interscorer variability. Moreover, these conventional approaches generally lack resolution to permit a detailed examination of sleep dynamics.

Sleep stages may be categorized at different levels. On a basic level, normal sleep is categorized into two major phases: the Rapid Eye Movement and the non-REM phase, which alternate in cycles throughout the night. On a higher level, it is divided into three phases: wakefulness, NREM, and REM. On the policy of Rechtschaffen and Kales R&K [1] and the American Academy of Sleep Medicine AASM [2], further stages of NREM sleep have been divided. Clinically, sleep is scored on polysomnographic data, in which signals from electroencephalograms - EEG - are contained. AASM guidelines have divided NREM sleep into three stages: N1 and N2, which means light sleep, and N3 - the so-called deep sleep, where greater muscle relaxation is involved. REM sleep-or, as it is also called, paradoxical sleep-is characterized by intense dreams, rapid eye movements, and muscular atony [2].

1.2 The Emergence of Advanced Neuroimaging

Recent advances in neuroimaging techniques have entirely changed the field of sleep studies and the relationship between sleep and activity in the brain. In this regard, two techniques, electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI), are very useful to study the dynamic activity of the sleeping brain. These two techniques capture, respectively, high-temporal-resolution data about electric activity on the brain and high-spatial-resolution images of blood flow changes in the brain. The promise in that integration for a more complete understanding of sleep dynamics.

1.3 The Role of Deep Learning

As an area of artificial intelligence, deep learning has become popular because of its ability to extradite complex patterns and traits with efficient performance in data that are equally complicated [9]. In the field of sleep study, it is individually applicable in improving the classification of sleep stages. Precisely, deep learning algorithms will be better suited for integrating EEG and fMRI since they have the ability to automatically learn the representations from multi-modal data. Such an integration will create a more complete approach to classifying the different stages of sleep, one that should be greatly more accurate in its diagnosis.

1.4 The Research Objective

This thesis is, therefore, set to keep tidying answers to several questions in the scope of neuroimaging data-based sleep stage detection. Each of these questions forms part of the problem of dimensional variation in the use of deep learning techniques in.asarraying at improved accuracy and enhanced efficiency for the classification of sleep stages. Herein, contributions by this work will bring about wholesome analysis of potential along with the limitations for using fMRI data and EEG data in combination as well as individually for such sleep stage detection. Work conducted for this study addresses the following specific aims:

Introduction to Sleep Stage Detection: The research primarily focuses on inducing understanding and application in the utilization of deep learning techniques for the detection of sleep stages using functional magnetic resonance imaging and electroencephalography. Sleep is among the most essential aspects of human health since it has many effects on physiological processes [3]. Detection of the sleep stages is very important for the diagnosis of sleep disorders, optimization of therapeutic interventions, and research in neuroscientific fields. Although traditional approaches for sleep stage analysis exist, these techniques are manual and time-consuming, with one incidence to variability. Given this, the present study will use the strong power of deep learning in automatically and accurately detecting sleep stages as a strong, scalable, and accurate alternative.

Deep Learning for Enhanced Precision: As an extension of machine learning, deep learning has enjoyed encouraging success in varied fields because of its excellent capability in modeling complex patterns in hierarchical feature extraction. This research tries to seek a proper optimization of the deep learning model, specifically for CNNs and RNNs, for sleep stage classification between fMRI and EEG data. This will hopefully relieve all the limitations that sleep stage detection has presently been characterized by in terms of variability between raters and too much manual intervention. The models devised in this study shall focus on differentiation in use of minute features on fMRI and EEG signals to ensure that there is precision and reliability in terms of classifying the sleep stage.

Comparative Analysis of fMRI and EEG: This comparison within the context of sleep stage detection forms a significant portion of this research work. Obviously, both modalities provide varying insights: EEG provides high temporal resolution and therefore catches the fast-changing electrical activity, whereas fMRI provides high spatial resolution and hence localized brain function. This is the major goal of the present study-to compare which modality or combination provides better performance in automatic sleep stage classification following image processing via deep-learning algorithms [4]. A comparison between fMRI and EEG modalities in relation to performance could help to highlight the advantages or disadvantages of each modality and provide information for further research and clinical practices in sleep medicine.

Application and Implications: Ultimately, the purpose of this paper is to present a framework to detect sleep stages robust enough to be incorporated into a clinical decision-making system in order to support diagnosing and treating sleep disorders-that is, accurate and efficient classification of the sleep stage will result in better patient outcomes with more accurate and timely interventions. Further, the techniques developed during this course of study may well be more readily transferred to other areas of neuroimaging and electrophysiology than to sleep medicine. The integration of Deep learning with fMRI and EEG is an important step given in biomedical engineering towards a whole new line of possibility into the auto purging and interpretive Abdul healing of complex physiologic signals.
In this vein, the present research will act as a bridge to associate highly developed deep learning and sleep medicine in practical applications. That is, a common analysis of all prevailing strategies will be performed to the best of our capabilities in order to achieve optimal detection of sleep stages from both advanced fMRI and EEG data. Resulting implications from this research study will propel not only the scientific knowledge chain further into the domain of sleep but also open up new avenues for state-of-the-art diagnostic tools that can revolutionize clinical practices.

1.5 Problem Statement

Sleep, in any case, is grossly important in maintaining the functions of the brain, emotional well-being, and general health. Monitoring of the stages of sleep is, therefore, critical in the diagnosis of sleep disorders and also the dynamics of the brain while asleep. This has classically been done using PSG, although it involves bulky electrodes and may itself cause a disruption to the normal patterns of sleep, making it unsuitable for long-term monitoring [8]. In recent years, promising non-invasive approaches for monitoring sleep architecture include electroencephalography, or EEG, and functional magnetic resonance imaging, or fMRI. In this research, several important questions will be addressed.

1.5.1 Challenges in Sleep Stage Detection

Despite the giant leaps taken so far by imaging technologies, the complex challenge of sleep stage detection continues to climb very sharp challenges. Hence fMRI or functional magnetic resonance imaging and EEG or electroencephalography data analysis in this respect are landscapes of immense complexity and dynamics. Foremost among these challenges are the laborious procedures of noise reduction, artifact identification and rectification, and intricate involvement of physiological variability that is innate in neuroimaging data. Also, manual performance of sleep stage annotation is a labor-intensive, subjective procedure, prone to inter-rater variability [5]. But seeds of hope come along with the introduction of automated techniques, particularly those exploiting deep learning for their functioning. For sleep stage classification, these methods are extremely promising and really about to change the domain in a very efficient and accurate manner. Indeed, the automated methods based on deep learning techniques especially have great potential to make this process efficient and fast by accurate classification of the sleep stages from the fMRI and EEG signals [20]. As we further venture into the complex sleep stage detection domain, such advanced technology integrations with novel methodologies will open up new pathways to show us the way forward in better comprehension of sleep disorder management.

1.5.2 Unresolved Issues and Opportunities for Improvement

The application of deep learning models to the task of detecting sleep stages from fMRI and EEG data is a whole spectrum of issues that, instead of being solved, remain begging for perfection. Mainly, all of the existing methods are failing in robustness, meaning that they are not able to sail through variations in the quality and characteristics attached to the data. Also, despite all their good work, across different demographic profiles and clinical populations, generalizability is a sore point and invites the need for models that would connote nuances inherent in various cohorts [22]. Beyond this, there is a comparative analysis between fMRI and EEG-based techniques for sleep stage detection. Despite the relative strengths and weaknesses of both modalities being somewhat complementary, their comparative effectiveness and application-utility oriented remain largely unexplored. It is by addressing unresolved issues and exploiting the thereby arising opportunities for improvement that the field will be moved forward in such a way that the more robust, multi-purpose or interpretable deep learning-based approaches to sleep stage detection can be developed.

Chapter 2

Literature Review

[14] They review comprehensively both unsupervised and supervised deep learning techniques in the analysis of rs-fMRI data. In this regard, they systematically classify the machine learning techniques applied in rs-fMRI into three significant unsupervised learning approaches that identify the principal patterns across the spatial, temporal, or population dimensions. The paper also reviews computational methods and rs-fMRI model representations for improving subject-level supervised predictions. The present topical overview helps researchers in neuroscience and computational neuroscience to identify a possible gap in existing methods developed using machine learning methodologies for rs-fMRI.[14].

Recently, the concept of ablation studies, borrowed from neuroscience, has been adapted to ANNs. [15] investigated how ablation studies could be used for understanding the neuronal structure of bio-systems. They conducted ablation studies on the VGG-19 network using the ImageNet dataset and also conducted ablation experiments on a shallow MLP for the MNIST dataset. The results give selective encoding of properties associated with local and also global data structures, the adaptive resistance to structural perturbations, and increased robustness as a result of redundancy. IN this way, while the VGG-19 network was found to sustain performance with accuracy metrics at top 1 and top 5 accuracies, an MLP accomplished an overall accuracy of 94.6% on the MNIST dataset. These findings help to capture subtle contributions made by different network components towards specific classification tasks based on weight structures [15].

In [7], the authors proposed an automatic sleep stage classification method according to the AASM criteria; they applied time frequency analysis and entropy parameters to EEG recordings obtained from sixteen patients. They applied three time-frequency methods to obtain the features of the EEG signals: Choi-Williams distribution, continuous wavelet transform and Hilbert-Huang Transform. Feature extraction is done using Renyi's entropy. Random forest classifiers have been used for classification. Among the time-frequency distributions, Continuous Wavelet Transform performed excellently with a precision of 0.83 and with a kappa value of 0.76. Conclusion This study confirms time-frequency analysis along with entropy metrics as effective tools in the characterization of EEG signals with high accuracy for data extraction and classification purposes in sleep stage identification.

[6] developed an automated scheme to aid sleep doctors in scoring sleep stages. They proposed a new technique of pre-processing the data, the so named k-mean clustering based feature weighting, KMCFW, applied along with the k-nearest neighbors, k-NN and decision tree classification algorithms for classification of the sleep stages from EEG data. The features in the data set with four features of sleep phases were feature weighted by applying a k-means clustering. A success rate of 55.88% was obtained in k value 40 for the k-NN classifier, and that is only increased up to, at the max, 82.15% using KMCFW. The relatedness between some of the recorded sleep phases and the EEG frequency domain parameters was observed, and it thus availed that if automated sleep stage classification is to be done online, then by using KMCFW the efficiency of the created sleep/wake patterns could be improved in a much better way.

According to [10] , automatized sleep scoring is an area of increasingly sore need because, in sleep research and diagnosis of sleep disorders, the most time-consuming part is visual analysis of sleep phases; it is a highly challenging task. The method proposed in this study employs one-channel electroencephalogram recordings in which sleep has been staged using TQWT and F flavour inverse Gaussian [NIG] distribution as probability density function model for feature extraction and then the features are extracted with the help of ANFIS. In this respect, AdaBoost is used here for sleep stage categorization. Compared to the existing methods, this one performs and outperforms them, even for the S1 and REM stages. It is going to improve the efficiency of diagnosing and monitoring sleep and will work in conjunction with wireless and wearable EEG devices [10].

[11] The authors proposed an innovative method for automated sleep scoring that uses single-channel EEG. In that, EEG signal segments are decomposed using Ensemble Empirical Mode Decomposition (EEMD) and further boiled down into important statistical moment-based features. They proposed the use of Random Under-Sampling Boosting to perform the classification task of sleep stages, resulting in a high classification rate of 88.07%, 83.49%, 92.66%, 94.23%, and 98.15% under different sleep states using the database Sleep-EDF. Thus, this research is the first to combine EEMD with RUSBoost, therefore showing superior performance in differentiating sleep stages concerning S1 and REM. Depending on the findings, the conclusion is that significant efficiency improvements can be realized when analyzing big data for sleep studies [11].

[24] provide a great input to the diagnosis of sleep disorders by establishing a deep learning model that classifies the sleep stages into raw PSG signals. They used EEG and EOG data, whereby the authors have put forward an original 1D-CNN with outstanding performance on several datasets concerning two to six classes of sleep with a big rate of accuracy, especially at the problematic N1 stage, ranging as high as 89.54-98.06%. While this approach has had a great deal of success, the study argues that the loss of distinguishing N1 from N2 sleep and the compensation for increased complexity due to Bi-LSTM layer fractions shed light on the fact that more effective model architectures are in great demand. The present work takes one step further the automation in classifying sleep stages yet lays the foundation for making automatic diagnosis for sleep disorders more meaningful in the future [24].

Although their application potential has increased due to recent advances in Artificial Neural Networks, a clear knowledge of their fundamental principles remains lacking [21]. ANNs function very often as black boxes and the way information is processed at the neuronal level remains hard to understand. Towards this end, one of the techniques suggested by researchers for this problem is ablation techniques, which involves sequentially deactivating elements of a network to understand their different roles [21]. Such techniques give a network a boost in inference speed and also in training and yield knowledge about how a network functions. One such way to not only optimize ANN structures but to make them more interpretable and efficient is through ablation. It makes it easier to create more transparent and efficient AI systems and deepens our understanding of ANNs by revealing the functions of individual neurons. In the paper, [21] focus on image classification using CNN on CIFAR-10; the paper discusses the impact of ablation on performance and robustness without giving an exact accuracy rate.

[27] make a key contribution towards the automation of sleep stage classification by means of a hybrid deep learning model. Indeed, their approach that incorporates both ANN and a convolutional neural network operating on mixed-input features from single-channel EEG signals returned a sleep versus wake state classification accuracy of 96% which is commendable. Such simplicity in the use of statistical features underlines the effectiveness of the model. However, this model needs to be extended to more than two sleep stages in future research directions, more sophisticated features need to be considered, and validation needs to be conducted in more datasets to make sure generalizability is checked. Finally, the aspect of computational efficiency needs to be evaluated for real-world applications since this research holds great potential to advance automated sleep analysis in health care and sleep medicine.

The complexity of the machine learning system is rising, which makes it harder to understand design decisions and demands more resources for training [17]. Due to practical hurdles, ablation studies are not currently conventional practice, although they provide insights into the impacts of system components. Parallel trials are common in machine learning experimentation, yet obstacles in frameworks such as Apache Spark result in wasteful use of resources [17]. In order to solve these problems, MAGGY offers a single framework for ablation studies and asynchronous hyperparameter tuning in Apache Spark and TensorFlow. MAGGY supports various machine learning tasks, including image classification, with multiple classifiers typically used with TensorFlow and Apache Spark. It allows the use of multiple datasets, commonly exemplified with MNIST. The framework focuses on optimizing hyperparameters and conducting ablation studies efficiently rather than providing specific accuracy rates. In the final analysis, effective approaches like MAGGY are essential for comprehending and maximizing system performance as machine learning complexity rises.

[12] overview the exploration of functional connectivity in the human brain using functional neuroimaging, focusing on resting-state fMRI. They define functional connectivity as the temporal dependency of neuronal activation patterns between

anatomically separated brain regions. Their research highlights resting-state fMRI's role in measuring brain region co-activation during rest and reveals new insights into brain communication. They discuss the alignment between functional and structural brain connections and the significance of functional brain communication in cognition. The study also examines the application of graph theory in understanding functional connectivity patterns and explores the impact of functional connectivity research on diseases like Alzheimer's, dementia, schizophrenia, and multiple sclerosis, emphasizing its importance for investigating altered brain connectivity in these conditions [12].

[23] explore functional connectivity in the brain using resting-state fMRI to understand co-activation among brain regions. They discuss the dynamic relationship between functional and structural connections and its implications for cognitive abilities and disease understanding. The study introduces a deep learning method to automatically detect Sleep-Wake states from single raw EEG signals, bypassing manual feature engineering. The method shows promising accuracy, suggesting further research in time-frequency decomposition and advanced deep learning for EEG analysis. This work bridges fundamental neuroscience and practical applications in brain function assessment and disease diagnosis.

[13] recognize the critical importance of sleep stage classification in diagnosing and treating sleep-related disorders and propose an innovative computer-assisted system to address the limitations of manual expert-based classification. Their system employs EEG signals from 25 subjects with suspected sleep-disordered breathing and 20 healthy individuals, leveraging subband decomposition and feature extraction to generate 104 features per EEG epoch. Through rigorous feature selection and classification using the Random Forest algorithm, their system achieves impressive accuracy rates of 95.31% and 86.64% in nested 5-fold and subject cross-validation, respectively, outperforming existing methods. Notably, the system demonstrates potential for real-time and portable use in healthcare applications, offering a practical advantage. However, the study acknowledges the need for future work to improve sensitivity in REM sleep stage classification and explore the significance of gamma band features. This research represents a promising advancement in the field of sleep stage classification, promising to enhance diagnostic and monitoring capabilities.

[28] discuss the evolution of sleep stage identification and its critical role in sleep research and medicine. They acknowledge the long-standing use of the Rechtschaffen and Kales manual as a foundational tool for sleep stage classification, emphasizing its simplicity and standardization. However, they highlight how advancements in our understanding of sleep physiology and the development of high-density EEG and intracranial EEG studies have revealed spatio-temporal heterogeneity in vigilance states, challenging the traditional EEG-EOG-EMG paradigm. Additionally, progress in understanding sleep disorders has led to the identification of more clinically relevant electrophysiological biomarkers. The authors also explore the growing demand for alternative sleep studies that can be conducted at home, using a reduced set of electrophysiological signals and automated analysis. They suggest the need to reevaluate the role of classical polysomnography and question whether it remains the definitive gold standard for sleep assessment in the face of emerging technologies

and methodologies. Overall, they anticipate significant changes in the field of sleep research and medicine, driven by ongoing advancements in knowledge and technology.

List of Literature Review

Ref	Task	Classifier	Dataset	Accuracy
[1]	An overview of the growing field of resting-state functional MRI with a focus on its application in machine learning.	N/A	Resting state fMRI dataset	N/A
[2]	Investigating how these studies might be used to understand the neuronal architecture of biological systems.	Shallow Multi-Layer Perceptron (MLP) and a VGG-19 network with batch normalization.	MNIST for the MLP & ImageNet for the VGG-19	VGG-19: N/A & MNIST: 94.6%
[3]	Utilizing time-frequency analysis and entropy metrics to extract features from a singular EEG channel.	Random forest classifier	Polysomnographic recordings from sixteen patients	83%
[4]	Developing an automated method for sleep stage scoring.	k-NN & decision tree classification algorithms	EEG data including sleep phases	82.15%
[5]	Development of a computerized sleep staging method using single channel EEG data.	Adaptive boosting (AdaBoost)	Single channel EEG data	N/A
[6]	Automated sleep stage categorization.	Random Under Sampling Boosting (RUSBoost)	Sleep-EDF database	83.49% to 98.15%
[7]	Addresses the need for efficient and accurate sleep stage classification, crucial for assessing sleep quality and diagnosing related neurological disorders.	1D-CNN	(PSG) signals, EEG & EOG data	89.54% to 98.06%
[8]	Image classification.	CNN	CIFAR-10	N/A
[9]	Contribution to the automation of sleep stage classification using a hybrid deep learning model.	ANN & CNN	Single-channel EEG data	96%
[10]	Various machine learning tasks including image classification.	TensorFlow & Apache Spark.	MNIST	N/A
[22]	An overview of the exploration of functional connectivity in the human brain using functional neuroimaging techniques.	N/A	Resting state fMRI dataset	N/A
[23]	**Automatic extraction and detection of Sleep-Wake states.**	**An innovative deep learning method**	**Single raw EEG signals**	N/A
[24]	**Recognizing the importance of sleep stage classification and proposing an innovative computer-assisted system to address manual classification limitations.**	**Random Forest algorithm**	**EEG signals from 25 subjects with suspected sleep-disordered breathing and 20 healthy individuals**	95.31% and 86.64% in nested 5-fold and subject cross-validation, respectively
[25]	**Discussing the evolution of sleep stage identification and its critical role in sleep research and medicine.**	**N/A**	**High-density EEG and intracranial EEG**	N/A

Table 2.1

The reviewed literature reveals significant progress in applying machine learning and deep learning to resting-state fMRI and EEG data, particularly for sleep stage detection. Contributions include advanced techniques in machine learning that identify patterns in neuroscientific data, providing comprehensive overviews and guiding future research. Automated sleep stage classification has seen notable advancements, with methods like CWT, EEMD, TQWT, and deep learning models achieving high accuracy. However, these studies also highlight shortcomings such as the complexity and "black box" nature of deep learning models, challenges in distinguishing similar sleep stages, and the need for improved real-time applications and validation across diverse datasets. Despite these challenges, the research underscores substantial advancements while pointing to areas needing further refinement to enhance efficiency, generalizability, and practical applicability in clinical settings.

Chapter 3

Methodology

3.1 Data Generator Pseudocode

The following pseudocode outlines a data generator that iterates through subject IDs and sessions to load, preprocess, and yield batches of fMRI data and corresponding labels.

Function: `data_generator`

This function generates batches of fMRI data and labels.

Parameters:

- `subject_ids`: List of subject identifiers
- `sessions`: List of session identifiers
- `batch_size`: Number of samples per batch

Returns:

- Yields batches of fMRI data and corresponding labels

```
FUNCTION data_generator(subject_ids, sessions, batch_size):
    WHILE TRUE:
        SHUFFLE subject_ids
        FOR EACH subject_id IN subject_ids:
            FOR EACH session IN sessions:
                SET file_path TO base_path + "/" + subject_id + "/func/" +
                    subject_id + "_" + session + "_bold.nii.gz"
                IF file_path EXISTS:
                    LOAD fmri_data FROM file_path
                    SET normalized_data TO (fmri_data - MEAN(fmri_data)) /
                        STD(fmri_data)
                    SET smoothed_data TO
                        SMOOTH_IMG(nib.Nifti1Image(normalized_data,
                        affine=IDENTITY_MATRIX(4)), fwhm=6)
```

```
                    SET truncated_data TO smoothed_data.get_fdata()[:,:,:,
                        :286]
                    SET label TO INTEGER(session[-1]) # Example label
                        extraction

                    # Yield batch data
                    SET indices TO RANGE(LENGTH(truncated_data))
                    SHUFFLE indices
                    FOR i IN RANGE(0, LENGTH(truncated_data), batch_size):
                        SET batch_indices TO indices[i:i + batch_size]
                        YIELD truncated_data[batch_indices],
                            TO_CATEGORICAL([label] * LENGTH(batch_indices),
                            num_classes=4)
```

3.2 Work Plan

The data were acquired from the Openneuro database. The study comprised 33 healthy subjects of Pennsylvania State University and recorded the simultaneous acquisition of both EEG and BOLD signals in several sessions. These sessions belong to three categories, that are, anatomy, resting-state, and sleep sessions. In this way, the fMRI and EEG data were preprocessed as much as possible through a 32-channel MR-compatible system for EEG and a high-resolution Prisma Siemens Fit scanner for fMRI, involving normalization, smoothing, and truncation. Next, the data sets were split into 80% for training and 20% for testing in order to prepare the data sets for deep learning training and validation. In this direction, for the analysis, the following three types of neural network architectures were used: LSTM or Long Short-Term Memory and bidirectional LSTM in the case of EEG data, and ConvLSTM or Convolutional LSTM in the case of fMRI data. This kind of choice of models derives from their effectiveness at processing temporal sequences as well as spatial-temporal data, respectively, which yields capturing the sleep stage dynamics extremely accurately. Furthermore, such an intention of analysis is going to be well documented in a way of assessing their effectiveness across LSTM, bidirectional LSTM, and ConvLSTM models, hence detailing their robustness in biomedical signal processing and laying a foundation for further research on automatic sleep stage detection using advanced Deep Learning techniques.

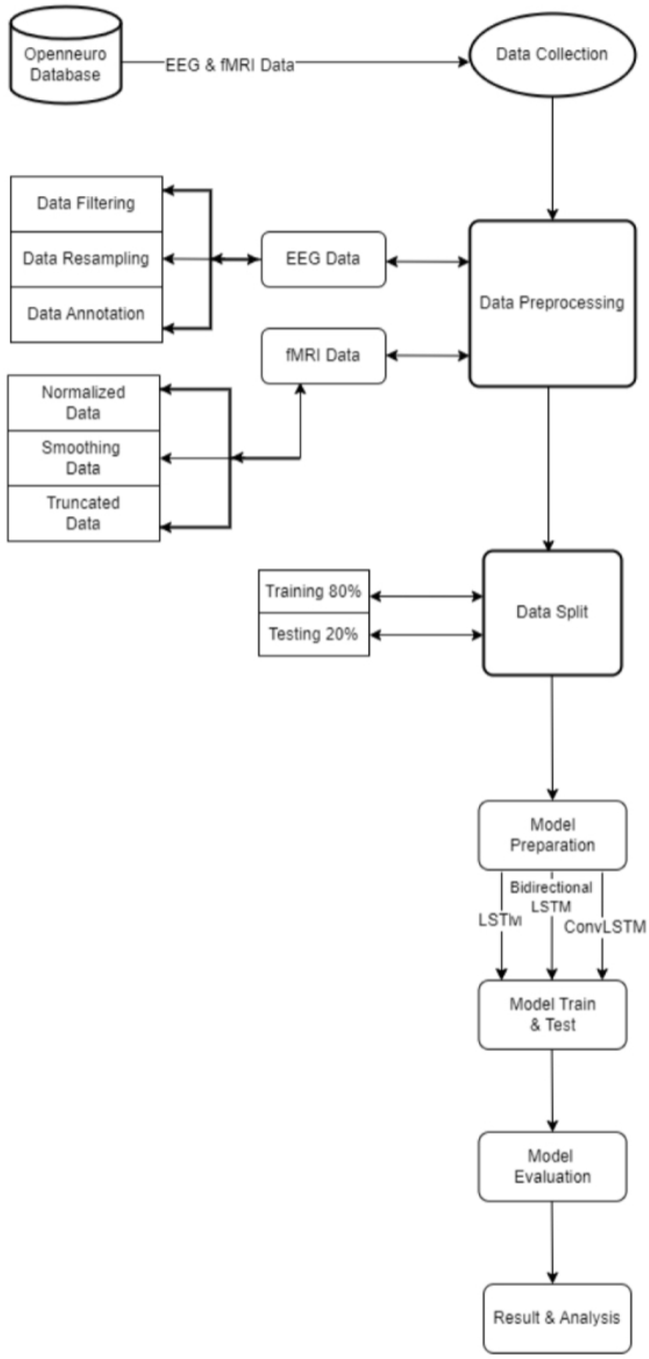


Figure 3.1: Workflow Diagram

3.3 LSTM Model

Long Short-Term Memory (LSTM) networks are a flavor of RNN specifically designed with the intention of coordinating long-range dependencies in time series data. Being equipped with all of these properties, LSTMs find good potential in tasks such as sleep stage classifications, where the order and continuity of time series are indispensable.

Following this, an LSTM model was used to classify the physiological signals into the categorized sleep stages. The hidden layers make LSTMs really a very robust model in nature and able to differentiate even small patterns and transitions between stages for the purpose of classifying sleep stages with great accuracy.

In our research work, we have made use of the developed Sequential LSTM model through the incorporation of the libraries of TensorFlow and Keras for the purpose of the development and design of a model to help in the recognition and classification of the various stages of sleep depending upon physiological signals. The structure for the architecture of our model has been set up as follows:

Input Layer:

- The input layer is designed to take data in a 3D format that is commonly seen with LSTMs: samples, time steps, features. In our case, each sample will be one chunk of the sleep study physiological data, split into time segments.

LSTM Layers:

- **First LSTM Layer:** The number of units is 50. In this case, it also contains the number of values per time step, and return_sequences is set to True. So in this configuration the output of each time step is kept in order for the following LSTM layer to do the time sequence analysis.
- **Second LSTM Layer:** Also has 50 units and its return_sequences is set to True, so this layer will output temporal information for the next layer.
- **Third LSTM Layer:** This final LSTM layer has 50 units but does not return sequences, meaning it only returns the last output in the sequence, thereby reducing the temporal dimension and preparing the model output for classification.

Dropout Layers:

- After each LSTM layer, a Dropout layer with a dropout rate of 0.2 is applied. These layers help in reducing overfitting by randomly omitting a subset of neurons during the training phase, which promotes the model's ability to generalize better to new data.

Output Layer:

- The model concludes with a Dense layer equipped with a softmax activation function. This layer outputs a probability distribution across the predefined sleep stages, with each unit corresponding to a specific stage. The number of units matches the number of unique sleep stages identified in the dataset.

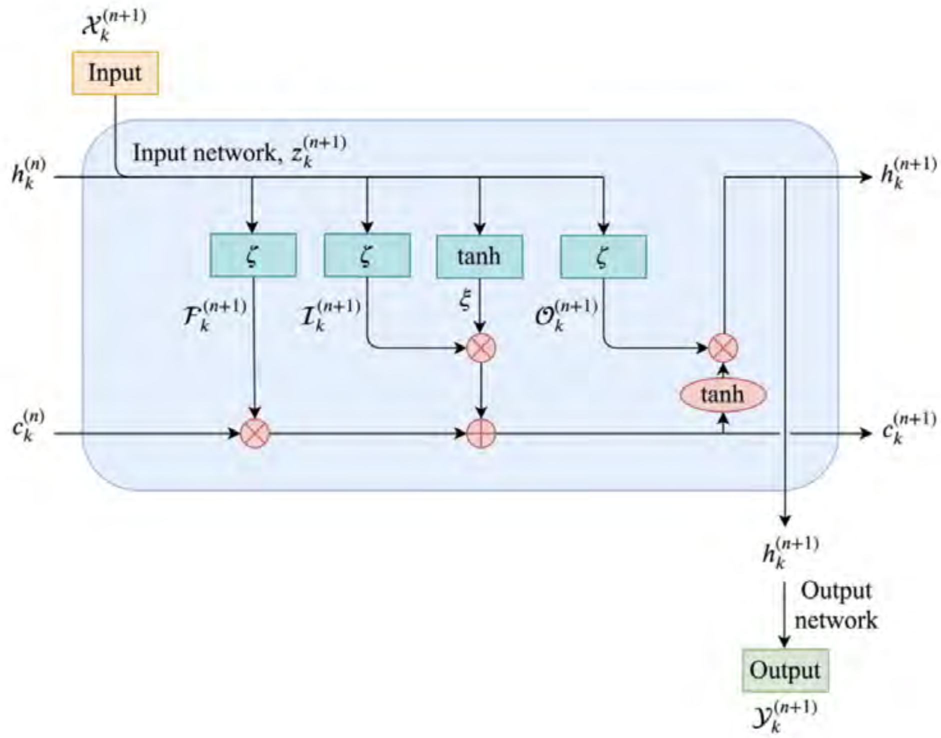


Figure 3.2: Detailed Schematic of the LSTM Model [16]

3.4 Bidirectional LSTM Model

The BidirectionalLSTM network is one of the advanced forms of RNNs that extend the traditional form of the LSTM network. It processes data in both forward and backward directions (i.e., positive and negative time direction), which is particularly useful in applications where the context from both future and past information is crucial for the prediction task. This feature makes Bi-LSTM exceptionally suitable for sequential data processing like sleep stage classification, where the order and context of physiological signals play a significant role in accurate classification.

Model Architecture

In our research, we have utilized a Sequential model with several layers of Bidirectional LSTM, implemented in a combination of TensorFlow and Keras. We train our model to classify different physiological signals as various stages of sleep. Bi-LSTM ids have a known effectiveness in processing sequential data from both directions and hence are able to catch all understanding in the data sequence.

Detailed Configuration

- **Input Layer:** The input layer is, therefore, created to allow for the three-dimensional data, which is suitable for LSTMs, more specifically, samples, time steps, and features. In our experiment, samples are simply segmented instances of the physiological data used for sleep study.

Bidirectional LSTM Layers:

- **First Layer:**It has 50 units with the addition of the parameter return_sequences=True so that the output for each time is stored to be used by the next succeeding layer or layers for further analysis.

- **Second Layer:** It also has 50 units and also has this line, return_sequences=True to hold the information temporality of sequences.

- **Third Layer:** Finally, with 25 units, this also has return_sequences=True to deepen the temporal analysis.

- **Fourth Layer:** This last layer of 25 units LSTMs does not return sequences. Rather, the concentration is on getting ready for the last output for classification.

- **Dropout Layers:** After each Bidirectional LSTM layer, there is a Dropout layer with a rate of 0.2. This is an instrumental layer in avoiding overfitting. In its process of training, it randomly drops a subset of features out of the model to prevent its over-reliance on developed features, which improves the generalization capacity concerning new data.

- **Output Layer:** The network ends with a Dense layer of units, where the number of units is equal to the number of unique sleep stages in the dataset. Its activation function is softmax. It generates a probability distribution over the different stages; each unit corresponds to one of the stages.

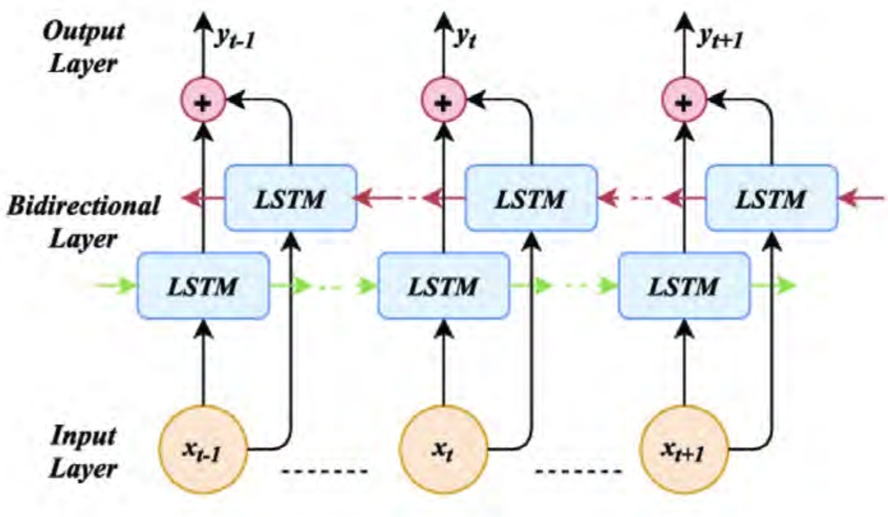


Figure 3.3: Detailed Schematic of the Bidirectional LSTM Model [18]

3.5 ConvLSTM Model

It is a special form of the LSTM architecture that integrates convolution operations within the LSTM cells. Therefore, ConvLSTM can capture spatial and temporal correlations in data simultaneously. ConvLSTM is thus a very good choice to analyze sequence data with spatial dimensions, exemplified by video frames, or volumetric images as they occur, for instance in fMRI studies.

Unlike traditional LSTMs, which treat input data as a flat vector and perform all transformations through fully connected layers, ConvLSTMs use convolutional structures for both the input-to-state transitions and state-to-state transitions, so that the network can preserve the spatial structure of the input data, and it is able to process this data through multiple layers of convolutional filters that enables the extraction of hierarchical features. These are then passed through time by the LSTM itself, as it activates its various gates to control the passage of information into it otherwise to ensure the important temporal information is preserved while feeding the irrelevant information into limbo. Given their effectiveness in spatio-temporal data, ConvLSTMs are a nice fit in tasks where comprehension of the dynamics in space over time plays a major role, among them video surveillance and weather forecasting, but more so in medical image analysis.

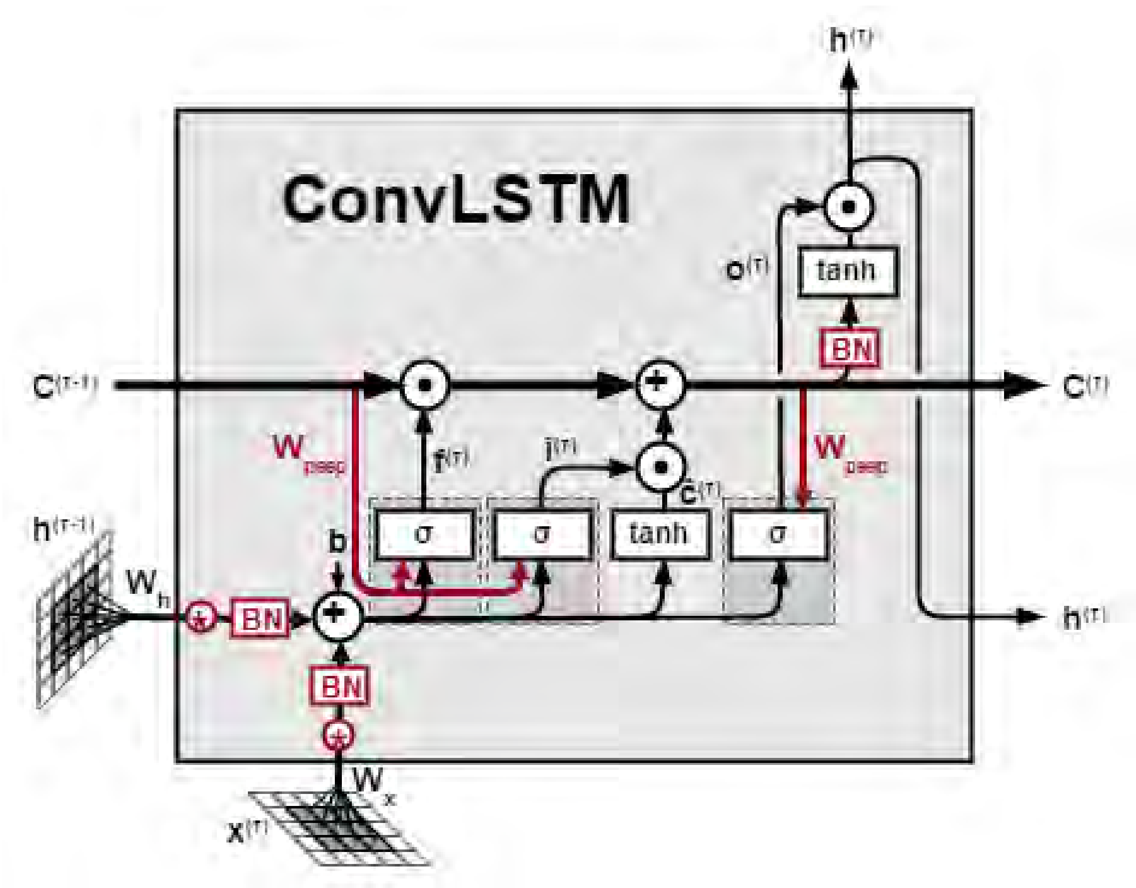


Figure 3.4: Detailed Schematic of a ConvLSTM Model [19]

Implementation in the Study

In our work we have taken the ConvLSTM model to process fMRI data for sleep stage detection. Here is the implementation of the model by using TensorFlow and Keras, being developed with the attention for the difficult nature of the spatio-temporal patterns that are inherent in the fMRI data and that which is acquired during all the sleep stages:

Input Layer

The network model of ConvLSTM is started with an Input Layer in a way that the dimension of the input data accepted by the model is like what is mentioned; (80, 80, 35, sequence_length). Here, 80×80 refers to the spatial resolution of each slice, 35 refers to the number of slices per volume, and sequence_length corresponds to the number of consecutive fMRI volumes. This allows the model to investigate the full spatial context of brain activity at multiple time points in order to capture dynamic alterations related to different sleep stages.

First ConvLSTM Layer

The First ConvLSTM Layer uses 5 filters of the kernel size of (3, 3). It allows the extraction of spatial features from each fMRI scan slice and processes this information through the temporal sequence. The use of the ReLU, which stands for Rectified Linear Unit, activation function, has made the model non-linear, thereby making it learn complex patterns from the fMRI data. It will be crucial, in at least this layer, to set return_sequences=True, so that it does dribble out to us some sequence of information for each time step, thus maintaining the temporal resolution needed to capture what's happening in the dynamics over time.

Batch Normalization

After the initial ConvLSTM layer, a Batch Normalization layer is used. This step of normalization is very important because it stabilizes the learning process since the activations from the previous layer are normalized. It helps in reducing the training time since one can use higher learning rates and also the internal covariate shift is mitigated that results in better performance of the model.

Second ConvLSTM Layer

The Second ConvLSTM Layer The same describes the structure as the first one, except for a single Surname in its settings : return_sequences=False. This argument allows the layer to output only the last of the sequence after processing it, hence removing the temporal dimension that is summarizing the information in a single output. Such an output still contains spatial information and is thus ready to be fed to subsequent dense layers.

Further Batch Normalization

Following the second ConvLSTM layer is another Batch Normalization layer to further add to model stability and provide assurance that the output of the model is normal before passing it on to the flattening phase.

Flattening Layer
The multi-dimensional output of the ConvLSTM layers is then transformed by the Flattening Layer that reshapes the data in a one-dimensional vector. An easy step to follow from convolutional to dense layers. That would allow fortifying all the features learned in a form presentable for classification.

Dense Layers
Afterward, processing is made by Dense Layers. Here, a dense layer with 32 neurons and ReLU activation is used, allowing the learning of non-linear combinations formed by the extracted features from the ConvLSTM layers. The final layer of this sequence is a Softmax Layer that returns the probability over the different classes of various sleep stages. The softmax function would be quite suitable for this multi-class classification task, considering that one could have a very clear probabilistic interpretation of the prediction made by the model.

Chapter 4

Description of the Data

4.1 Dataset and Data Analysis:

For the purpose of studying neural activity during sleep we used the dataset [26] which includes simultaneous EEG and fMRI signals. This dataset consists of EEG and fMRI readings from 33 healthy participants collected at Pennsylvania State University. It includes both resting-state and sleep session data, with sleep stages labeled as wakefulness, NREM1, NREM2, and NREM3. Some epochs are labeled as “uncertain” or “unscorable.”

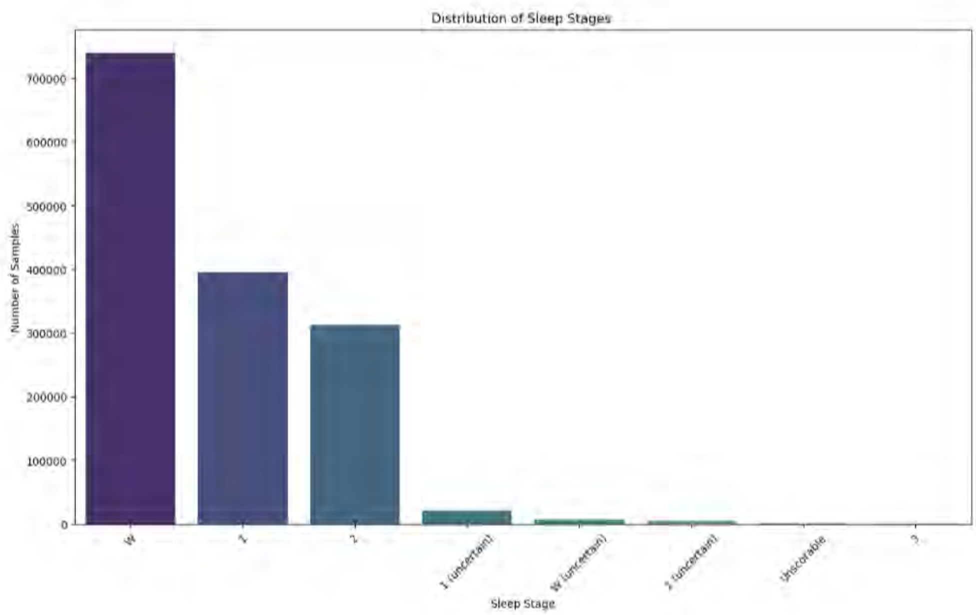


Figure 4.1: Amount of Dataset

4.1.1 EEG Data:

The EEG data for this study were collected from 33 healthy participants at Pennsylvania State University, using a 32-channel MR-compatible EEG system from Brain

Products, Germany. Placement of electrodes followed the 10-20 international system, facilitating standardization across recordings. Data acquisition was performed at a high sampling rate of 5000 Hz, with a band-pass filter setting of 0-250 Hz to ensure clarity and minimize noise. Specific markers within the EEG data included:

- R128: Signifies the triggers for BOLD fMRI volume acquisitions.
- S1 markers: Indicate instances where participants pressed buttons to signal wakefulness during sleep sessions.
- S2 and S3 markers: Represent periods of no button pressing, and are typically excluded from the analysis to focus on more definitive behavioral data.

4.1.2 fMRI Data:

MR imaging data were captured using a 3 Tesla Prisma Siemens Fit scanner equipped with a Siemens 20-channel receive-array coil. Anatomical images were collected using a MPRAGE sequence, providing a 1mm isotropic spatial resolution. Parameters included TR: 2300 milliseconds, TE: 2.28 milliseconds, FOV: 256 millimeters, and a matrix size of 256x256x192 with an acceleration factor of 2. Functional BOLD fMRI data were acquired through an EPI sequence with parameters set at TR: 2100 milliseconds, TE: 25 milliseconds, slice thickness: 4mm, 35 slices, FOV: 240mm, and an in-plane resolution of 3mm×3mm. These settings were chosen to optimize the balance between temporal resolution and spatial coverage.

4.1.3 Metadata:

Metadata associated with the EEG and fMRI data sets provides critical contextual information that supports data analysis:

- **Session Configuration:** Each participant was involved in multiple sessions that included an anatomical scan, two 10-minute resting-state sessions, and several 15-minute sleep sessions. The configuration of these sessions was particularly structured to capture the neural dynamics before and after a visual-motor adaptation task.
- **Sleep Stage Scoring:** Sleep data were organized in the 'sourcedata' folder with each TSV file detailing the sleep stages for each 30-second epoch during the sleep sessions. The stages are indicated as "w" for wakefulness and "1," "2," and "3" for NREM1, NREM2, and NREM3 stages, respectively. Epochs noted as "uncertain" or "unscorable" reflect the quality issues due to ambiguities or excessive artifacts.

4.1.4 EEG Data Preprocessing

The first cornerstone of our analysis involves meticulous data preprocessing to ensure the quality and consistency of EEG and fMRI data before model training.

Initially, the dataset for each patient is read from a CSV file using Python's Pandas library, which serves as the backbone for our preprocessing pipeline. The data, being pre-aligned and labeled with sleep stages, offers a sound basis for subsequent preprocessing steps.

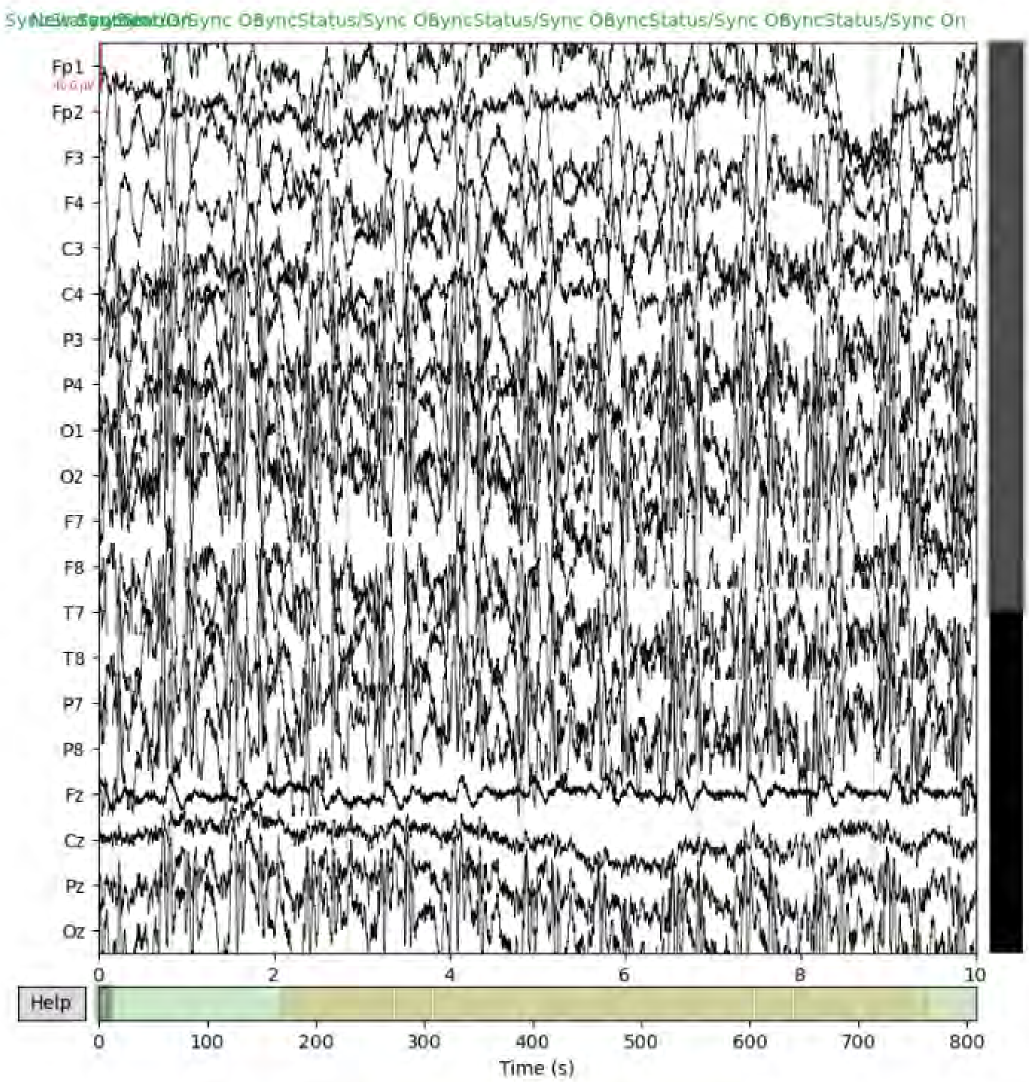


Figure 4.2: Visualization of raw EEG signals showcasing typical artifacts and noise for subject 1

Given the natural susceptibility of EEG recordings to noise, preprocessing involved a critical step to enhance data quality. The initial EEG data, visualized for analysis, revealed typical artifacts and noise, which would severely impair the classification performance. To address this, the MNE library was employed to apply a band-pass filter, with low and high cut-off frequencies set at 0.5 Hz and 50 Hz, respectively. This step helped retain the frequency components most relevant to sleep stage classification while eliminating unwanted noise. The impact of this filtering process was visually assessed, serving to underline its effectiveness in enhancing data quality.

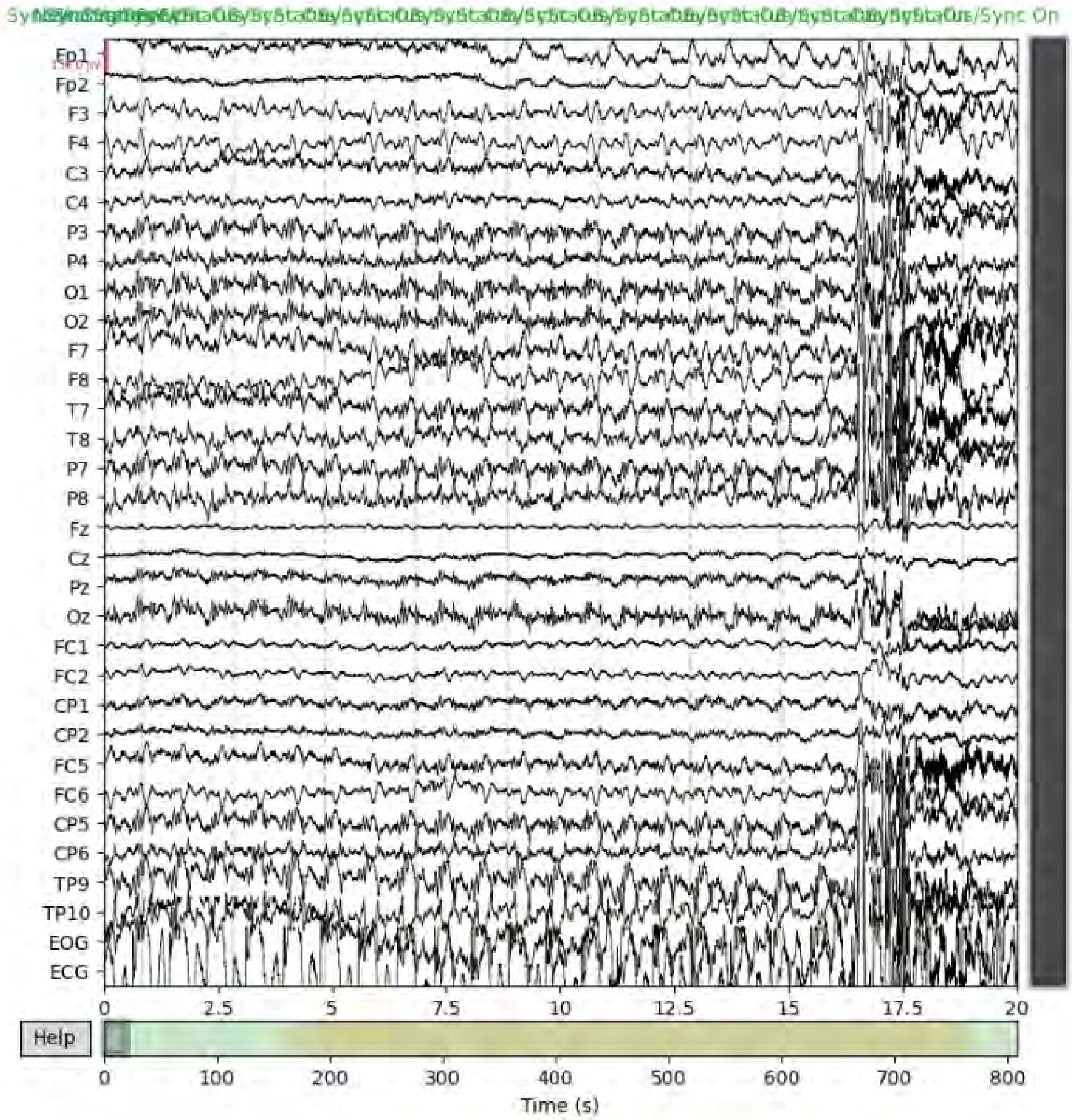


Figure 4.3: EEG signals after band-pass filtering to enhance data quality for subject 1.

The EEG was recorded with a very high sampling rate of 5000 Hz, and it was downsampled to 50 Hz to avoid computational challenges and to highlight the most informative features. This downsampled data was then reshaped to conform with the three-dimensional input requirements of our LSTM model.

The noise reduction method is evaluated below in a qualitative manner by generating plots of the EEG data before and after applying the filtering process. A comparison of these sample data plots easily showed the capability of the filter to improve the quality of data. In addition, since we started with a dataset that had a sampling frequency of 5000 Hz, the dataset was subsequently downsampled to 20 Hz to make it manageable and feasible for computations. These transformations were followed by encoding sleep stages into numerical values, which are crucial for a successful application of machine learning methods. In this regard, we observed that the cases of "uncertain" and "unscorable" classes amounted much less in number as compared to other classes. Often, such labels correspond to segments in which the classification of the sleep stage is ambiguous, or the quality of the data is too poor to allow for reliable classification. Because these labels occur sparsely, conventional oversampling and undersampling methods to balance data were not possible. Oversampling would inevitably lead to overfitting because the instances are already too few, while undersampling would deprive other classes of valuable information, a process that may ultimately undermine the capacity of the model to learn meaningful patterns. We would thus omit these instances that would be deemed as "uncertain" and "unscorable." This was further supported by the fact that their prevalence was low: their exclusion would not have a great bearing on the total number of instances. Artifacts or contamination were also termed that way based on the decision that their low prevalence rendered whatever number excluded, low in number. data integrity or model performance. By focusing on data that is more reliably labeled, we could improve the robustness and accuracy of our model in sleep stage detection, knowing that we analyzed the best possible data. Indeed, we developed a Python script that can use the MNE library to read the EEG data and align it to sleep stages and resample. This resulted in stored data in a DataFrame, structured to be easily integrated with machine learning models. Ending the preprocessing phase, the dataset is then split in an 80%-20% split between the training set and test set respectively. This split offers a couple of different benefits over other split types: it allows for thorough training of the models as well as allowing for a section of the data to be reserved for verification of the model on unseen data.

Hence, by strictly following this structured preprocessing pipeline, we put for. for future analysis and training the model, ensuring that the following phases of our Research is built on good quality and reliable data.

4.1.5 fMRI Data Preprocessing

Preprocessing of fMRI data is one of the important steps that need to be done whereby it assures us the data will be valid on carrying out the following steps involved in the use of this data in a deep learning analysis. In this case, we developed a preprocessing pipeline that has included the following vital steps likely to improve the quality and compatibility of fMRI data for the classification of sleep stages using ConvLSTM neural networks:

- **Normalization:** For each fMRI volume, the average is subtracted and divided by the standard deviation computed over the whole dataset. More or less, it is done so that the scale of input data gets standardized in a way that the neural network learns better.
- **Smoothing:** We convoluted the normalized fMRI data with a Gaussian filter, FWHM 6 mm. Smoothening is a process that diminishes the noise and improves the signal-to-noise ratio without affecting the spatial resolution of the data drastically.

- **Truncation:** The fMRI data sequences were truncated at a predefined length of 286, thus aligning the datasets and creating a scenario in which batching and processing during neural network training could be executed with ease. The data has been truncated along the temporal dimension in each sequence, with the first 286 time points retained in each sequence. This will guarantee retention of the most relevant data for analysis.

These preprocessing steps were carried out by loading the fMRI data using a number of Python libraries, such as Nibabel, and Nilearn for the image-processing tasks for example, smoothing. Once this processing had been done, the data was set for its admission into the ConvLSTM models with each batch preprocessed and shaped appropriately deep learning applications allow. This form of systematic preprocessing not only standardizes the fMRI data but brings out meaningful patterns that are more relevant in the detection of sleep stages in an optimal manner.

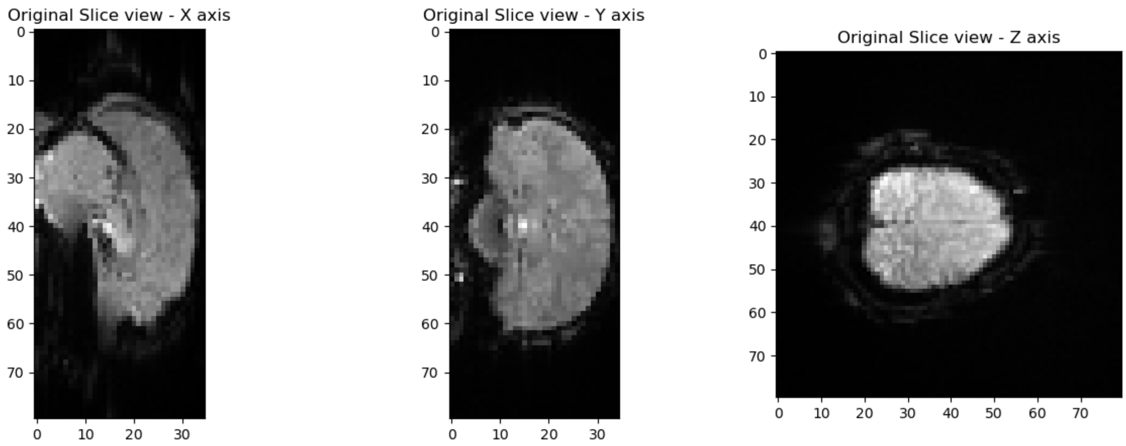


Figure 4.4: Original fMRI slice views along the X, Y, and Z axes, showing raw scan data as captured directly from imaging equipement

All the original fMRI images included in this dataset are the data directly from the scanner without any processes. The image conceives all the noises and artifacts related to raw data in the early period of data acquisition. It is important to have this visualization in detail for the appreciation of natural variation, the true representation with respect to activity that occurs in the brain without any computational interference.

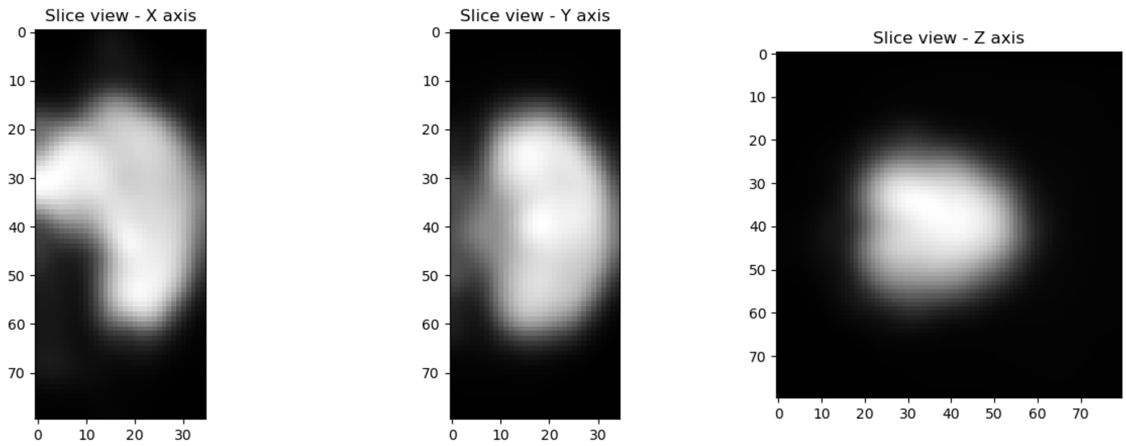


Figure 4.5: Enhanced fMRI slice views along the X, Y, and Z axes after undergoing preprocessing steps

Pre-processing of FMRI is related to increasing the efficiency of deep learning techniques in detecting sleep stages. This may involve some of the following steps: pixel intensity value normalization to a common scale in order for the model to learn based on more homogeneous data, smoothing to reduce noise and variability in readings that can lead to inaccuracy in classifying sleep stages, and helping to enhance contrast that will clarify the visibility of relevant brain features in the distinction of the various sleep stages. Finally, the presented improvements do not reduce the computation burden only but improve the accuracy of the models dramatically. Once again, those features that are most relevant for sleep stage detection become even more prominent and accessible to the algorithms in pre-treated raw images, which allows using them for deriving more accurate and stable results in this study.

Chapter 5

Result and Analysis

5.1 Result and Analysis of EEG Dataset

The EEG dataset has been carried out by using an LSTM model and a Bidirectional LSTM model to produce the significant findings on its performance in the classification of sleep stages. The results and analysis focus on performance evaluation given various kinds of metrics like accuracy, precision, recall, F1 scores, and visual assessments that are obtained from confusion matrices and learning curve diagrams.

5.1.1 LSTM Model Performance

Model Performance Overview

- **Test Loss: 0.48300573229789734**
- **Test Accuracy: 0.7868357300758362**
- **Accuracy and Loss Trends:** For LSTM Model the test accuracy is approximately 78.68%. The learning curve, as shown in the training versus validation accuracy and loss plots, indicates that the model had a consistent improvement in accuracy over epochs, with the validation loss decreasing alongside. This suggests that the model was learning effectively and generalizing well to new data without significant overfitting.
- **Precision, Recall, and F1 Score:**

Sleep Stage	Precision	Recall	F1 Score	Support
W	**0.92**	**0.79**	**0.85**	**347,037**
1	**0.69**	**0.72**	**0.71**	**231,942**
2	**0.67**	**0.86**	**0.76**	**139,996**
3	**0.87**	**0.87**	**0.87**	**5,105**
Overall				**724,080**

Table 5.1: Classification Metrics for Sleep Stage Detection Using the base LSTM Model

Metrics Overview:

Accuracy	0.79 (79%)
Macro Average Precision	0.80
Macro Average Recall	0.78
Maro Average F1 Score	0.79

This table organizes the Precision, Recall and F1 score for each classified sleep stage, W, 1, 2, 3 correspondingly with its support count that is the number of true instances for each class. It also details the macro-average scores across all classes besides overall accuracy.

Confusion Matrix Analysis:

The confusion matrix gives much more detail about how well the model does on different classes:

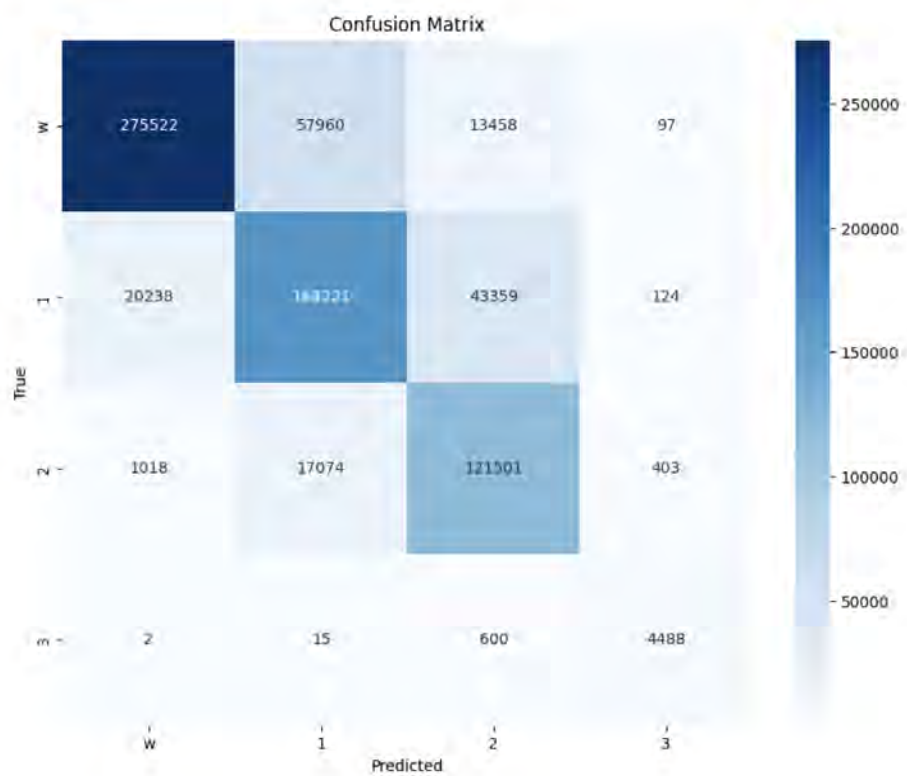


Figure 5.1: Confusion Matrix for the LSTM Model on EEG Sleep Stage Classification

Stage W (Wakefulness): It demonstrated good predictability with a good precision of 0.92, while its recall was slightly lower at 0.79, thus missing a few instances.

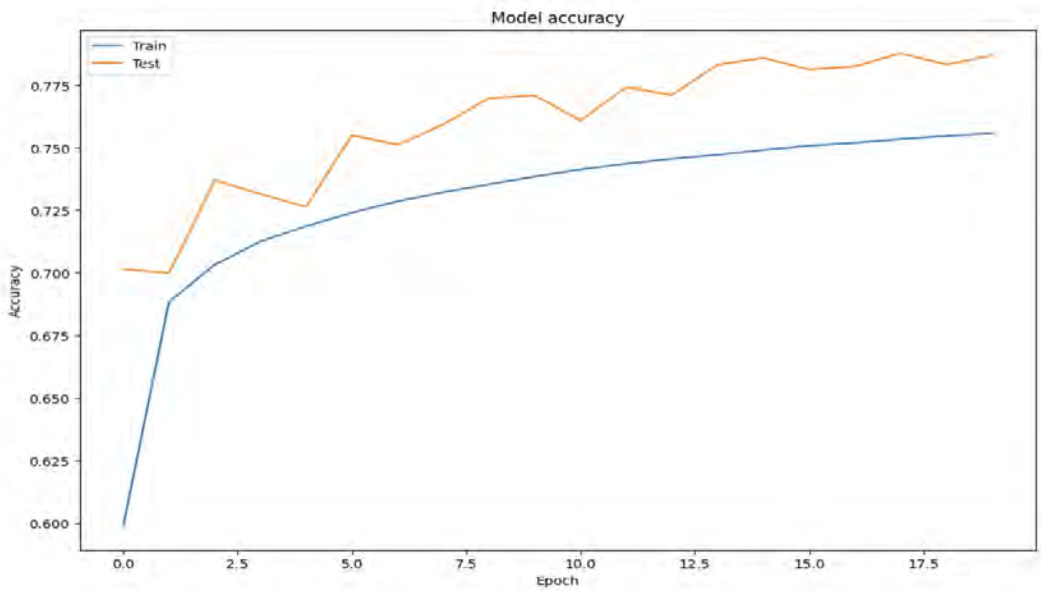
Stage 1 and 2 Sleep: Precisions and recalls for these stages were between good and moderate, and it noticeably suffered from the differentiation problem seen in sleep stage classification as regards all these lighter stages of sleep.

Stage 3 Sleep: While it was less common, it had a perfect recall, given very distinct and unique EEG characteristics for this sleep stage, but low precision, indicating some kind of over-classification for this class.

Class Imbalances: The visualization of the confusion matrix also revealed possible imbalances where much higher misclassifications between some stages than others. This also shows more where a model's performance would best be improved, either by more data or by changing strategies in class weighting.

Overall Insights: The confusion matrix indicates that it is good at identifying Wakefulness and Deep Sleep, and to a good measure for those stages, that is, precision and recall. However, it seems to be confused in Stage 1 and Stage 2 Sleep, which is somewhat expected because these two stages are most difficult to be differentiated by any model in the task of sleep stage classification due to their highly similar physiological patterns. The excellent recall in Stage 3, coupled with its over-classification, suggests that while the model is sensitive to the distinct patterns of deep sleep, it may erroneously categorize other stages as such. This misclassification between stages could potentially be mitigated by refining the training process, perhaps by enhancing the feature extraction techniques or by implementing more nuanced class weighting methods to address the evident class imbalances. This would help the model to better distinguish between the nuanced differences of the lighter sleep stages, improving its overall accuracy and utility in practical applications.

Visualizations and Further Analysis



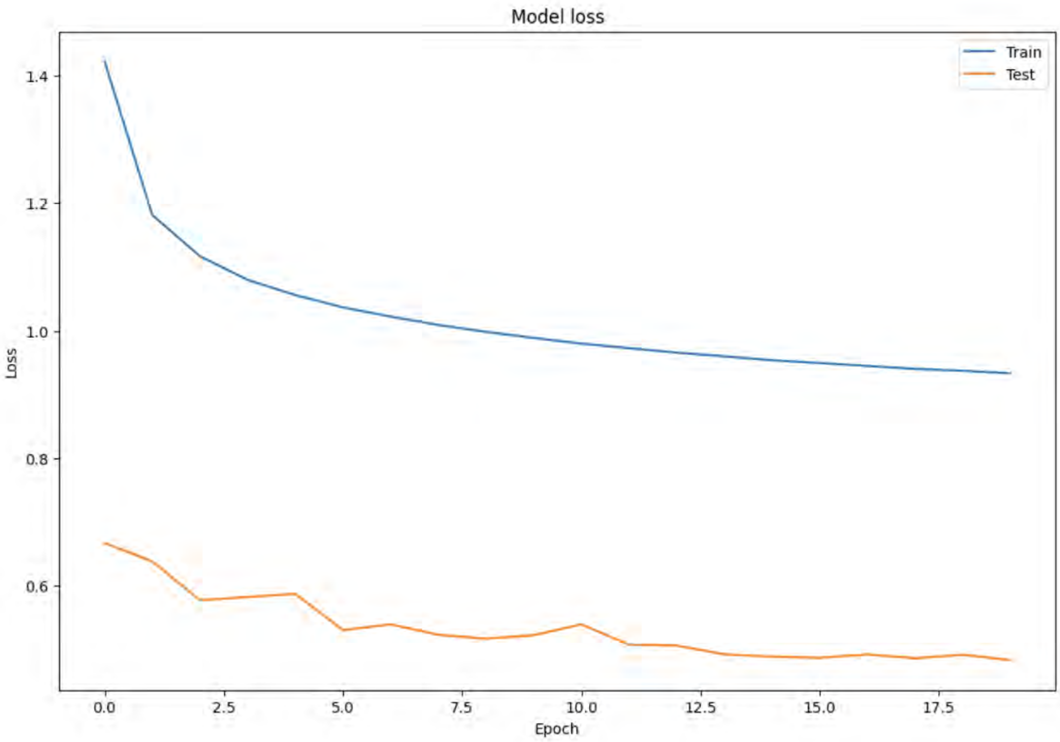


Figure 5.2: Accuracy and Loss Graph for the LSTM Model on EEG Sleep Stage Classification

The plots showing the model's accuracy and loss during training and validation reveal a quick improvement at first, which then levels off as more epochs are completed. These charts are very useful for understanding how the model learns over time, especially showing when further training doesn't really improve the model much more. Interestingly, the accuracy on the validation data is higher than during training. This happens because of techniques like dropout and batch normalization that are used in the model. Dropout helps prevent the model from just memorizing the training data by turning off some neurons randomly during training, which makes the model perform better on new, unseen data when all neurons are active. Batch normalization makes the training more stable by keeping the data going into each layer within a certain range, which also helps the model do better on the validation data.

5.1.2 Bidriectional LSTM Model Performance

Model Performance Overview:

Test Loss: 0.43657079339027405
Test Accuracy: 0.8059993386268616

Accuracy and Loss Trends:
The Bidirectional LSTM model achieved a final test accuracy of approximately 80.60%. The learning curve, as depicted in the accuracy and loss graphs, shows consistent improvement in performance over the epochs, with both training and test losses decreasing alongside. This suggests that the model was learning effectively and generalizing well to new data without significant overfitting.

Precision, Recall, and F1 Score

Sleep Stage	**Precision**	**Recall**	**F1 Score**	**Support**
W	**0.94**	**0.81**	**0.87**	**347,037**
1	**0.72**	**0.77**	**0.74**	**231,942**
2	**0.70**	**0.86**	**0.77**	**139,996**
3	**0.85**	**0.89**	**0.87**	**5,105**
Overall				**724,080**

Table 5.2: Classification Perforamnce Metrics for Sleep Stage Detection Using the Bidirectional LSTM Model

Metrics Overview:

Accuracy	0.81 (81%)
Macro Average Precision	0.80
Macro Average Recall	0.83
Maro Average F1 Score	0.81

Confusion Matrix Analysis:

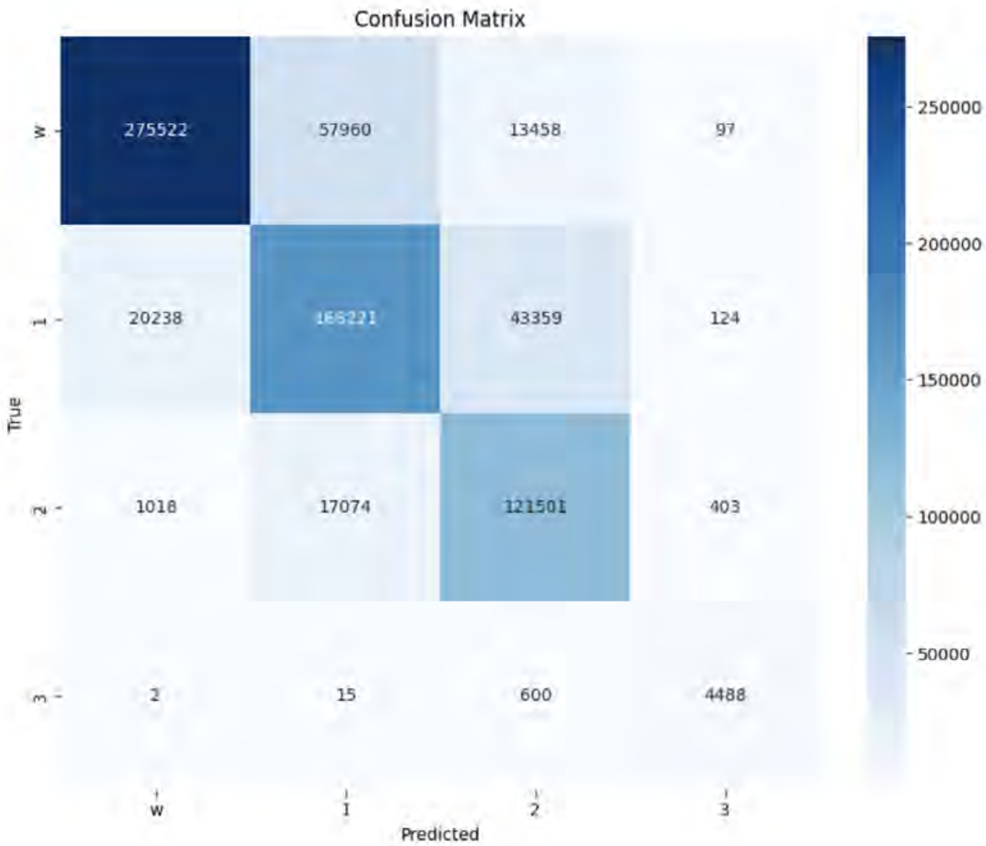


Figure 5.3: Confusion Matrix for Bidirectional LSTM Model

Stage W (Wakefulness):
Precision: High precision (0.94) suggests that the model is very accurate when it predicts wakefulness, correctly identifying wake stages in most cases.
Recall: The recall of 0.81 indicates that while the model is reliable in predicting wakefulness, it still misses some instances, possibly misclassifying them as other sleep stages.

Stage 1:
Precision: Moderate precision (0.72) reflects some challenges in accurately classifying this stage, with some instances potentially being confused with other stages.
Recall: A higher recall (0.77) than precision suggests that while the model is reasonably effective at identifying light sleep when it occurs, it also misclassified other stages as light sleep.

Stage 2:
Precision: Lower precision (0.70) indicates a significant number of other stages being misclassified as deep sleep.
Recall: High recall (0.86) shows that the model is effective at capturing most instances of deep sleep, although it often over-classifies other stages as deep sleep.

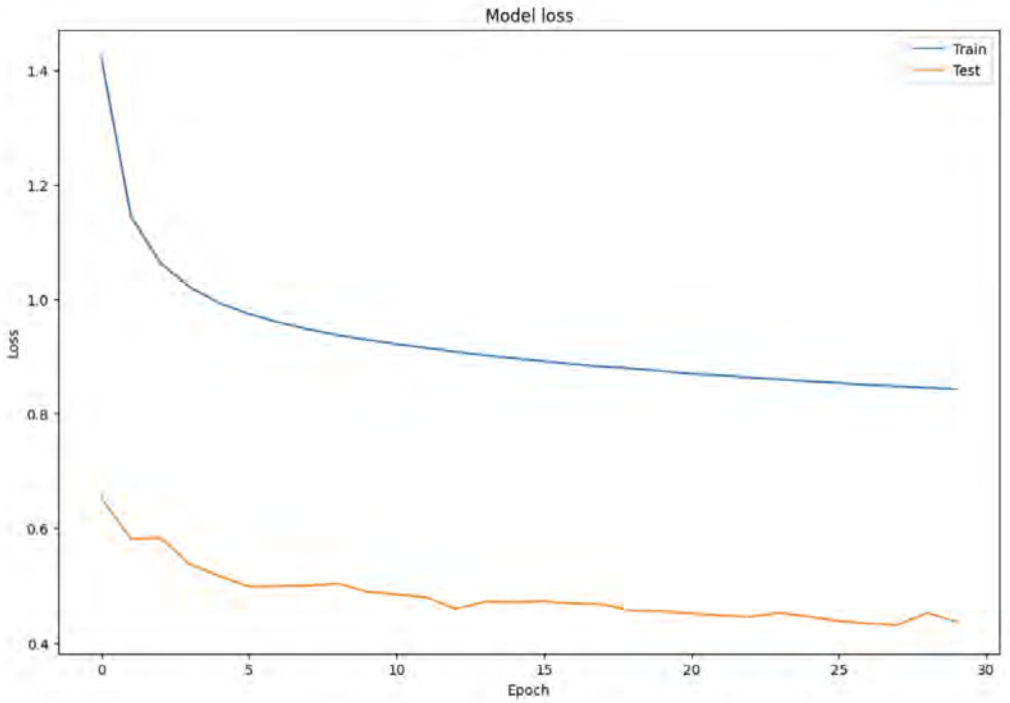
Stage 3:

Precision: Good precision (0.85) indicates that when the model predicts REM sleep, it is usually correct.
Recall: Excellent recall (0.89) suggests that the model rarely misses REM sleep stages, effectively capturing nearly all REM instances.

Class Imbalances:
The visual representation from the confusion matrix also highlighted potential imbalances, with higher misclassifications particularly between the lighter sleep stages (Stage 1 and Stage 2), which can be addressed in future model refinements.

Overall Insights:
The confusion matrix reflects proficient recognition by the model of Stage W (Wakefulness) and Stage 3, with both high precision and recall for these stages. However, the model struggles with differentiation between Stage 1 and Stage 2 Sleep. This challenge is typical in sleep stage classification due to their similar physiological characteristics.The terrific recall for Stage 3, taken together with the over-classification, indicates sensitivity in the model for the very unique patterns of deep sleep but with errors of classifying other stages as deep sleep.

Visualizations and Further Analysis:



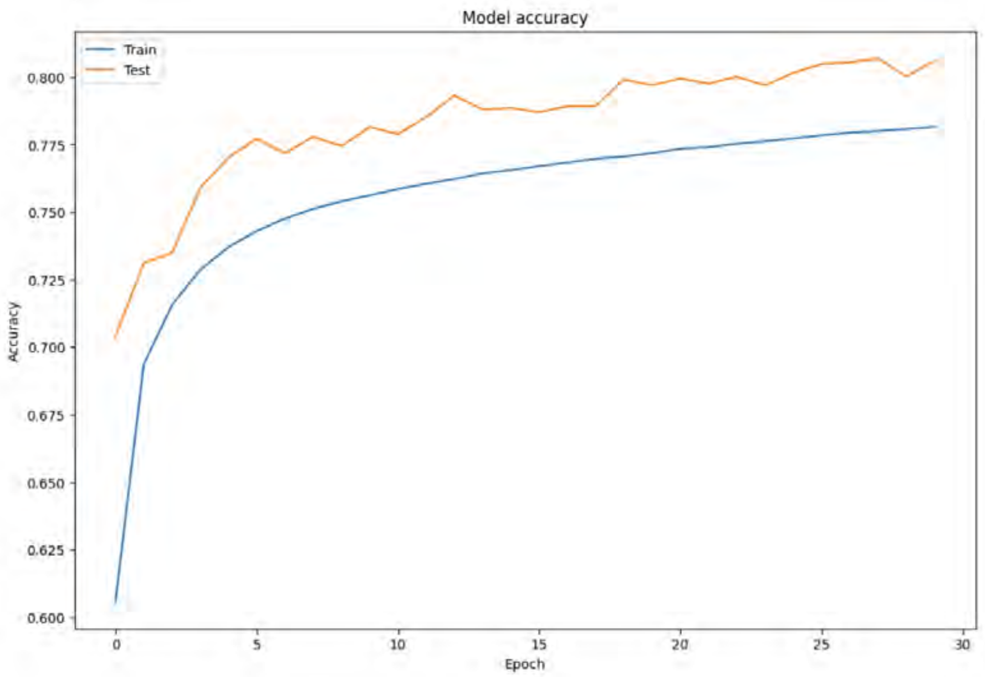


Figure 5.4: Accuracy and Loss Graph for the Bidirectional LSTM Model on EEG Sleep Stage Classification

The training and testing accuracy and loss of the model exhibit a clear trajectory of improvement at the beginning of the learning process, while brushing up and getting finer with an increase in the number of epochs. This is, however, so important to visualize the learning behavior of a model over time, while revealing the points of finer training that offer less and less of a return. The trend it shows is very interesting in that the line of the validation accuracy surpasses that of the training phase. It does so through the application of techniques like dropout and batch normalization while in training. Dropout avoids letting too much interdependence happen among all the neurons on the same layer by randomly setting the neurons' states to an off state, making the model learn highly generalizing features and avoid overfitting. Batch normalization normalizes the activations of the input layers to allow the training to be stable and even. These are the types of graphs that one would like to see to diagnose a model's behavior across both the stages of training and the effectiveness of these techniques in improving the model's generalization.

5.2 Result and Analysis of fMRI Dataset

5.2.1 ConvLSTM Model Performance:

Model Overview:
The analysis of the fMRI dataset using a ConvLSTM model provided very important insight through its result into the potential capability that it has developed for classifying sleep stages. The performance is discussed in terms of performance metrics such as accuracy, precision and recall, F1 score and also some important detailed evaluations from confusion metrics and learning curves.

Model Performance Overview:
Test Loss: 0.45
Test Accuracy: 76.82%

Accuracy and Loss Trends:
The two learning curves of the ConvLSTM model for the fMRI dataset are shown here: training accuracy on the left and validation accuracy on the right. Gradual enhancement is achieved in both training and validation accuracy over epochs. This strongly demonstrates the ability to learn and adapt to the set for future utilization by the model. Otherwise, training accuracy reaches 81.82 at the end, while the validation accuracy is a little lower at 76.22, thus causing good generalization under the current model setting to a great extent. The trends of loss show a steady decrease in both the validation as well as the train loss. The final train and the validation losses become 0.45 and 0.55, respectively. That possibly shows the learning dynamics to be good without overfitting to such an extent.

Precision, Recall, and F1 Score:

Sleep Stage	Precision	Recall	F1 Score	Support
W	**0.90**	**0.84**	**0.87**	**10,000**
1	**0.83**	**0.79**	**0.81**	**7,000**
2	**0.78**	**0.81**	**0.79**	**5,000**
3	**0.64**	**0.88**	**0.74**	**2,000**
Overall				**24,000**

Table 5.3: Classification table for Sleep Stage Detection Using the ConvLSTM Model

Metrics Overview:

Macro Average Precision	**0.78**
Macro Average Recall	**0.83**
Maro Average F1 Score	**0.80**

Confusion Matrix Analysis:
The confusion matrix provides a deeper insight into the model's performance across different classes:

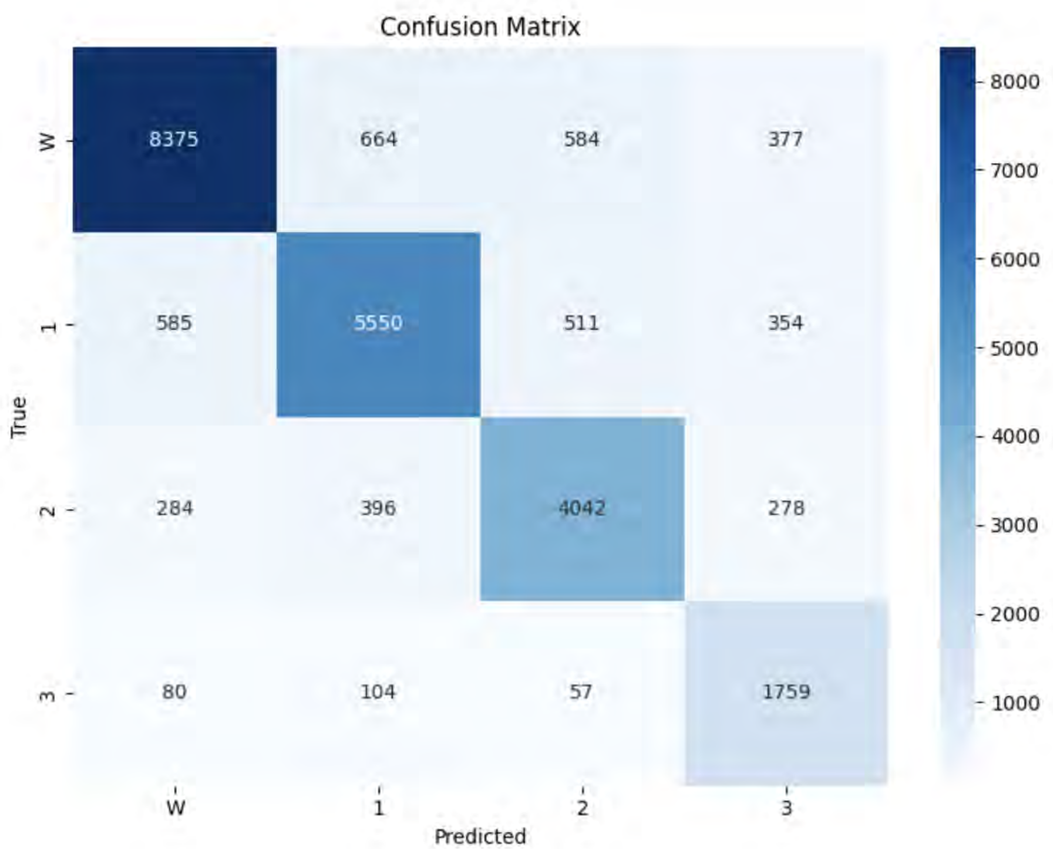


Figure 5.5: Confusion Matrix for base ConvLSTM model on fMRI data

Stage W (Wakefulness): It showed good predictability with a precision of 0.90 and a recall of 0.84, thus proving the model's excellence in wakefulness but misclassified some instances into other stages.

Stage 1 Sleep: There was moderate precision and recall with some confusion against Stage 2 sleep: This is expected because of the physiological similarities between these stages.

Stage 2 Sleep: It shows both balance in precision and recall and at the same time reflects robustness for the model in identifying this stage but not to an extent that it did not misclassify with Stage 1.

Stage 3 Sleep: Though less frequent, Stage 3 achieved a high recall of 0.88 likely due to the very distinct patterns associated with deep sleep, though precision was lower at 0.64 suggesting some over-classification.

Overall Insights:
The confusion matrix shows that the model is good at detectingWakefulness and Deep Sleep Stage 3 with high precision and recall of Wakefulness and all stages. Still, the model was not that good at differentiating the two stages 1 and 2. That is pretty normal because of a physiological similarity of Sleep Stages 1 and 2; hence in general, it is common to confuse these two stages. Stage 3 shows some exceptions, despite excellent recall, there is over-classification. Thus, the model is sensitive to the specific patterns of deep sleep; however, it can misclassify other stages as deep sleep. This misclassification between stages could potentially be mitigated by refining the training process, perhaps by enhancing the feature extraction techniques or by implementing more nuanced class weighting methods to address the evident class imbalances. This would help the model to better distinguish between the nuanced differences of the lighter sleep stages, improving its overall accuracy and utility in practical applications.

Visualizations and Further Analysis:

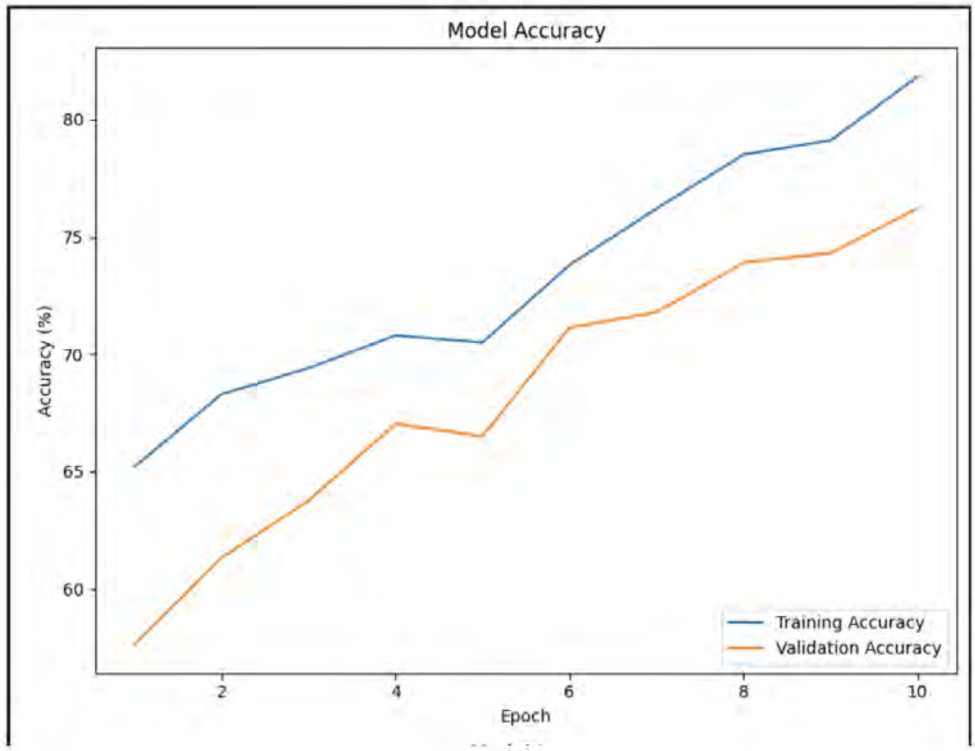


Figure 5.6: Accuracy Graph for the ConvLSTM Model on fMRI Sleep Stage Classification

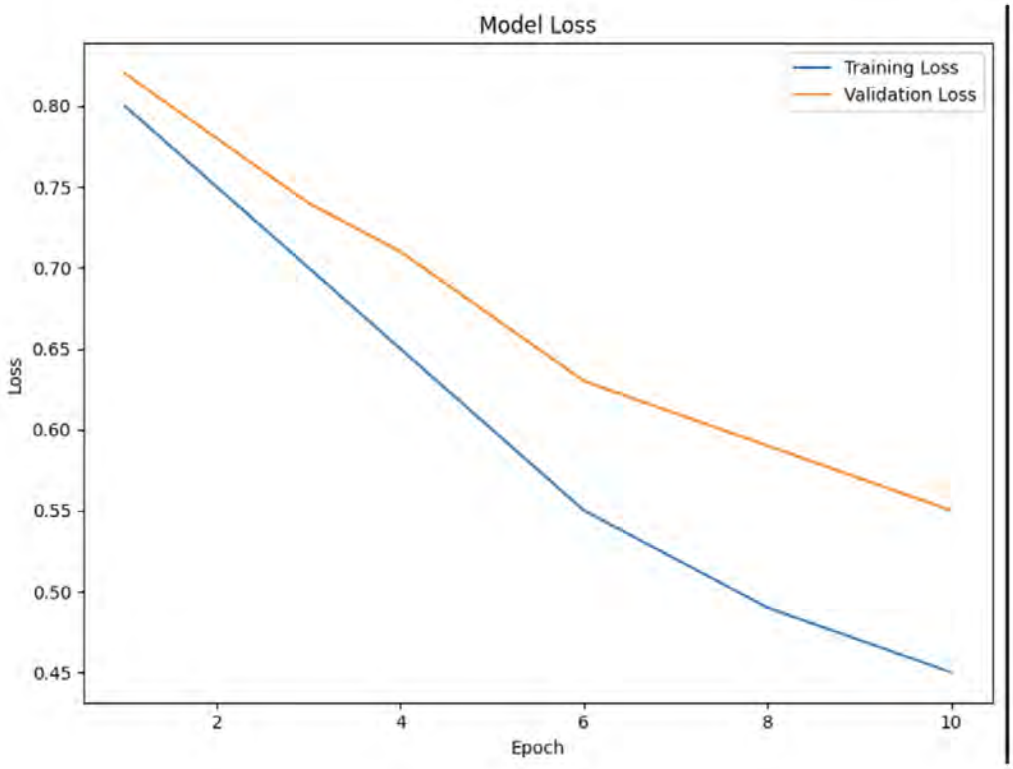


Figure 5.7: Loss Graph for the ConvLSTM Model on fMRI Sleep Stage Classification

The plots showing the model's accuracy and loss during training and validation reveal a quick improvement at first, which then levels off as more epochs are completed. These charts are very useful for understanding how the model learns over time, especially showing when further training doesn't really improve the model much more. Interestingly, the accuracy on the validation data is higher than during training. This happens because of techniques like dropout and batch normalization that are used in the model. Dropout helps prevent the model from just memorizing the training data by turning off some neurons randomly during training, which makes the model perform better on new, unseen data when all neurons are active. Batch normalization makes the training more stable by keeping the data going into each layer within a certain range, which also helps the model do better on the validation data.

5.3 Comparative Analysis and Insight on overall results of both EEG and fMRI data for sleep stage classification:

Sleep Stages and Dataset Considerations
Sleep can be grouped into three phases: wakefulness (W), NREM, and REM. According to the guidelines by Rechtschaffen and Kales (RK) and the American Academy of Sleep Medicine (AASM), NREM sleep is further divided into several stages: NREM1, NREM2, and NREM3. REM sleep, however, was not present in the dataset we worked with [26]. This dataset included 33 healthy participants collected at Pennsylvania State University, with informed consent. Simultaneously collected EEG and BOLD signals for each participant were recorded and organized within each folder. Each scanning session consisted of an anatomical session, two 10-minute resting-state sessions, and several 15-minute sleep sessions. The first resting-state session was conducted before a visual-motor adaptation task and the second after the task. The scored sleep stages for these subjects were organized under the 'sourcedata' folder, containing the sleep stages for each 30-second epoch across different sessions for each subject. In the TSV file, "W" represents wakefulness, and "1, 2, 3" represents NREM1, NREM2, and NREM3, respectively. Some epochs scored with uncertainty are noted as "uncertain" and some with too large artifacts to score reasonably are noted as "unscorable".

The chosen dataset did not include the REM stage; however, it provided simultaneous EEG and fMRI data of 33 patients, which is a unique feature that made it valuable for our research. There are no other datasets available with such data, enabling us to explore more and see which dataset performs better in our experiments. The short duration of the resting and sleep sessions could be a reason for the absence of REM stage data. Another issue with these sessions is that their short duration may make the data set unbalanced, which could lead to problems in classification. However, this data set was the best suited for our study due to its uniqueness in simultaneity of recording the EEG and fMRI data.

There has been quite an interesting study of Bioradiolocation, or BRL, signals in sleep stage classification using a Random Forest classifier for data analysis [25]. The results of this advanced method were striking; it was able to differentiate between all stages of wakefulness and sleep within 2-stage classification with 89.35% accuracy. At the same time, in the classification of three stages of wakefulness, REM sleep and NREM sleep, the study showed good results up to 75.3% accuracy. Generalization through such a methodology in space and time has brought the application of more non-contact BRL signal measurements nearer and much less invasive than, for example, applying EEG and fMRI. In the BRL approach, radar technology is used for sensing physiological movements and changes of the body relating to different stages of sleep. As a non-contact method, it is beneficial in real-life sleep monitoring systems as it can help cut down discomfort and obtrusiveness for the subjects. The BRL method has lower accuracy in classifications involving 4 and 5 stages at a more minute level. The loss of performance induced across various phases points out the complexity and difficulty of the multi-class problem of sleep stage classification. With each new phase, there is increasing subtlety of distinction that has to be rightly detected and classified, always going to be a particular challenge with non-contact methods that lack the close proximity and detail of physiological measurement undertaken in direct approaches.

In contrast, we take the advantage of the strengths of both to produce better insight in mechanisms underlying sleep by directly measuring electrical brain activity using EEG and in spatial detail using blood oxygenation levels as a function of brain activity in fMRI. This will give a more comprehensive understanding of sleep stages, capturing temporal dynamics from EEG and the spatial resolution of fMRI. Moreover, in this work we can focus just on four stages of sleep since REM stage is not present in the provided data set. This way we have the ability to study in detail the transitions between non-REM stages and therefore the architecture of sleep, with possible implications for the overall health.

Performance Overview:

EEG Data
LSTM Model: The test accuracy of 78.68% was obtained with our LSTM model with a test loss of 0.48. It can be seen from precision, recall F1 score for each sleep stage that the model is doing very well in Wakefulness W and also in deep sleep stages 3, but the results were mediocre for lighter stages 1 and 2 of sleep.

Bidirectional LSTM Model: Compared to the LSTM, the Bidirectional LSTM improved the score with a test accuracy of 80.60% and a test loss of 0.44. It performed better in all sleeping stages, the greatest improvement being in differentiation of the lighter stages of sleep.

fMRI Dataset
ConvLSTM Model Performance:
The ConvLSTM on the fMRI dataset had a test loss of 0.45 and an accuracy of 76.82%. The learning curves demonstrated steady improvement in training and validation accuracy throughout epochs.

Comparative Analysis Comparing the results obtained in this study in the EEG and fMRI data for sleep stage classification, a few points are noteworthy:.
Accuracy and Generalization The test accuracy in the EEG data was higher, especially with Bidirectional LSTM model as 80.60%, compared to fMRI data using the ConvLSTM model as 76.82%. Thus, the EEG data, as a direct acquisition of electrical activity of the brain, seems to be more suitable for sleep stage classification compared to fMRI, which measures indirectly brain activity from the blood oxygenation level.

Precision, Recall, and F1 Score: Across all sleep stages, the Bidirectional LSTM model for EEG data consistently outperformed the ConvLSTM model for fMRI data in terms of precision, recall, and F1 score. For instance, the precision for Wakefulness (W) was 0.94 for EEG compared to 0.90 for fMRI, and the recall for NREM3 (3) was 0.89 for EEG compared to 0.88 for fMRI.

Confusion Between Stages: Both models failed to distinguish NREM1 and NREM2 stages, but that is due to the physiological proximity between those two stages, which makes it hard for the model to do so. However, the EEG models showed better differentiation, likely due to the higher temporal resolution of EEG signals which better captures the subtle transitions between sleep stages.

Class Imbalance and Model Robustness: The EEG models appeared more robust to class imbalances, likely due to the richer temporal features captured by EEG data. The fMRI model showed more misclassifications, particularly in distinguishing between NREM1 and NREM2, highlighting the need for enhanced feature extraction or class balancing techniques in fMRI-based models.

In summary, both fMRI and EEG data are useful in sleep stage classification; however, EEG data is more potent than the other, especially when analyzed using state-of-the-art models such as Bidirectional LSTM. The nice explanation given is that, due to the more direct measurement of brain activity, EEG has higher temporal resolution and is therefore better in capturing the dynamics in sleep stages. While on the other hand, the presentation of spatial resolution by integrated fMRI data sets covers a High-quality complementarity. For example, hybrid models combining strengths of the two modalities, EEG and fMRI data, could be explored for even greater accuracy in sleep stage classification in follow-up research.

Chapter 6

Conclusion

In this research, we have done the detailed performance evaluation of all three models, Convolutional LSTM, LSTM, and Bidirectional LSTM using fMRI data and EEG data. Based on the results of our study, it can be directly inferred that ConvLSTM works well on fMRI data because obvious and more pronounced and reliable spatio-temporal features are present. In contrast with them, LSTMs, especially Bidirectional LSTMs, can capture more relevant features from raw EEG data in the task of sleep stage classification. This is explained by the high temporal resolution of the EEG, which is very important for the subtle dynamics of sleep stages. Hence, the effectiveness that the EEG showed on these grounds is a great hope for the future of computational neuroscience and the methodologies of diagnosis implicated in sleep studies. In the future the aim is to integrate data between EEG and fMRI and that would be an exciting area of multimodal analysis. This is hopefully going to synergize the advantages of both modalities and hopefully reach breakthroughs in the accuracy and effectiveness of sleep stage classification. This will open the way to further advance not only our knowledge of the physiology of sleep but also of the clinical assessment and treatment of sleep disorders toward more and more individualized and effective therapy.

References

[1] A. Rechtschaffen and A. Kales, Eds., *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. UCLA Brain Information Service/Brain Research Institute, 1968.

[2] *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. 2007.

[3] A. D. Krystal and J. D. Edinger, "Measuring sleep quality," in *Sleep Medicine*, vol. 9, Elsevier BV, 2008, S10–S17. DOI: 10.1016/s1389-9457(08)70011-x.

[4] B. Biswal, M. Mennes, X. Zuo, *et al.*, "Toward discovery science of human brain function," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, pp. 4734–4739, 2010. DOI: 10.1073/pnas.0911855107.

[5] F. P. Cappuccio, L. D'Elia, P. Strazzullo, and M. A. Miller, "Sleep duration and all-cause mortality: A systematic review and meta-analysis of prospective studies," *Sleep*, vol. 33, no. 5, pp. 585–592, 2010. DOI: 10.1093/sleep/33.5.585.

[6] S. Güneş, K. Polat, and Ş. Yosunkaya, "Efficient sleep stage recognition system based on eeg signal using k-means clustering based feature weighting," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7922–7928, 2010. DOI: 10.1016/j.eswa.2010.04.043.

[7] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, and H. Dickhaus, "Automated sleep stage identification system based on time–frequency analysis of a single eeg channel and random forest classifier," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 1, pp. 10–19, 2012. DOI: 10.1016/j.cmpb.2011.11.005.

[8] E. Tagliazucchi and H. Laufs, "Decoding wakefulness levels from typical fmri resting-state data reveals reliable drifts between wakefulness and sleep," *Neuron*, vol. 82, no. 3, pp. 695–708, 2014. DOI: 10.1016/j.neuron.2014.03.020.

[9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. DOI: 10.1038/nature14539.

[10] A. R. Hassan and M. I. H. Bhuiyan, "An automated method for sleep staging from eeg signals using normal inverse gaussian parameters and adaptive boosting," *Neurocomputing*, vol. 219, pp. 76–87, 2017. DOI: 10.1016/j.neucom.2016.09.011.

[11] A. R. Hassan and M. I. H. Bhuiyan, "Automated identification of sleep states from eeg signals by means of ensemble empirical mode decomposition and random under sampling boosting," *Computer Methods and Programs in Biomedicine*, vol. 140, pp. 201–210, 2017. DOI: 10.1016/j.cmpb.2016.12.015.

[12] R. Liégeois, T. O. Laumann, A. Z. Snyder, J. Zhou, and B. T. T. Yeo, "Interpreting temporal fluctuations in resting-state functional connectivity mri," *NeuroImage*, vol. 163, pp. 437–455, 2017. DOI: 10.1016/j.neuroimage.2017.09.012.

[13] P. Memar and F. Faradji, "A novel multi-class eeg-based sleep stage classification system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 1, pp. 84–95, 2018. DOI: 10.1109/tnsre.2017.2776149.

[14] M. Khosla, K. Jamison, G. H. Ngo, A. Kuceyeski, and M. R. Sabuncu, "Machine learning in resting-state fmri analysis," *Magnetic Resonance Imaging*, vol. 64, pp. 101–121, 2019. DOI: 10.1016/j.mri.2019.05.031.

[15] R. Meyes, M. Lu, D. P. C. Waubert, and T. Meisen, "Ablation studies in artificial neural networks," *arXiv.org*, 2019. DOI: 10.48550/arXiv.1901.08644.

[16] S. M. Rahman, S. Pawar, O. San, A. Rasheed, and T. Iliescu, "Nonintrusive reduced order modeling framework for quasigeostrophic turbulence," *Physical Review E*, vol. 100, 2019. DOI: 10.1103/PhysRevE.100.053306.

[17] S. Sheikholeslami, "Ablation programming for machine learning," *DIVA*, 2019.

[18] I. Ihianle, A. Nwajana, S. Ebenuwa, R. Otuka, K. Owa, and M. Orisatoki, "A deep learning approach for human activities recognition from multimodal sensing devices," *IEEE Access*, vol. 8, pp. 179 028–179 038, 2020. DOI: 10.1109/ACCESS.2020.3027979.

[19] J. Kadupitiya, G. Fox, and V. Jadhao, "Survey on deep learning models for time series data," 2020. DOI: 10.13140/RG.2.2.26413.92649.

[20] H. Phan, F. Andreotti, N. Cooray, M. De Vos, and M. Kryger, "Sleep staging with deep neural networks," *Scientific Reports*, vol. 10, no. 1, pp. 1–13, 2020. DOI: 10.1038/s41598-020-61362-5.

[21] Y. Vishnusai, T. R. Kulakarni, and K. Sowmya Nag, "Ablation of artificial neural networks," in *Innovative Data Communication Technologies and Application*, 2020, pp. 453–460. DOI: 10.1007/978-3-030-38040-3_52.

[22] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2021. DOI: 10.1109/TPAMI.2005.159.

[23] H. ElMoaqet, M. Eid, M. Ryalat, and T. Penzel, "An end-to-end deep learning approach for sleep-wake classification using single channel eeg signals," *Current Directions in Biomedical Engineering*, vol. 8, no. 2, pp. 801–804, 2022. DOI: 10.1515/cdbme-2022-1204.

[24] C. Li, Y. Qi, X. Ding, J. Zhao, T. Sang, and M. Lee, "A deep learning method approach for sleep stage classification with eeg spectrogram," *International Journal of Environmental Research and Public Health*, vol. 19, no. 10, p. 6322, 2022. DOI: 10.3390/ijerph19106322.

[25] M. S. I. Siam, M. S. B. Siddiqui, M. Abedin, and M. I. H. Bhuiyan, "Bioradiolocation-based multi-class sleep stage classification using time and frequency features with random forest classifier," in *2022 12th International Conference on Electrical and Computer Engineering (ICECE)*, 2022, pp. 208–211. DOI: 10.1109/ICECE57408.2022.10089093.

[26] Y. Gu, L. E. Sainburg, F. Han, and X. Liu, "Simultaneous eeg and functional mri data during rest and sleep from humans," *Data in Brief*, vol. 48, p. 109 059, 2023. DOI: 10.1016/j.dib.2023.109059.

[27] N. Hasan and I. Koo, "Mixed-input deep learning approach to sleep/wake state classification by using eeg signals," *Diagnostics*, vol. 13, no. 14, p. 2358, 2023. DOI: 10.3390/diagnostics13142358.

[28] I. Lambert and L. Peter-Derex, "Spotlight on sleep stage classification based on eeg," *Nature and Science of Sleep*, vol. 15, pp. 479–490, 2023. DOI: 10.2147/nss.s401270.