SleepGCN-Transformer: A Hybrid Graph Convolutional and Transformer Network for Sleep Stage Classification

A Master's Thesis

Submitted in partial fulfillment of the requirement for the award of the degree of

Master of Technology in Artificial Intelligence

by

Tanmay Rathod

23MAI007

Under the guidance of

Dr. Santosh Kumar Satapathy

Department of Information and Communication Technology

School of Technology

Pandit Deendayal Energy University

Gandhinagar – 382426. Gujarat - India

May, 2025

Approval Sheet

This thesis entitled **"SleepGCN-Transformer: A Hybrid Graph Convolutional and Transformer Network for Sleep Stage Classification"** by **Tanmay Rathod** is recommended for the degree of **M.Tech** in **Artificial Intelligence**.

____

Guide : Dr Santosh Satapathy,

Assistant Professor

Dept. of ICT

____

Dr Nitin Singh Rajput,

Assistant Professor

Dept. of ICT

____

Dr. Paawan Sharma,

Head Of Department

Dept. of ICT

Date: ________________________

Place: ________________________

Acknowledgment

While my name appears as the sole contributor to the completion of this summer internship, it is important to recognize that the guidance and support of many individuals played a significant role in its success. This internship is the result of a collective effort, shaped by the contributions of numerous people, to whom I am deeply grateful.

I extend my heartfelt appreciation to my internal guide, Dr. Santosh Kumar Satapathy Author, for his consistent guidance, encouragement, and insightful suggestions throughout this internship. His belief in my abilities provided the confidence and motivation necessary to complete this project. Dr. Santosh Kumar Satapathy Author's support, along with his generous provision of time and access to resources, was instrumental in achieving the objectives of this internship.

I am also profoundly thankful for the unwavering support of my parents. Their continuous inspiration, moral backing, and blessings have been the cornerstone of my journey. Their understanding and encouragement have been invaluable, and I am truly fortunate to have had their steadfast support throughout this endeavor.

Tanmay Rathod

Student Declaration

I, **Tanmay Rathod**, hereby declare that this written submission represents my ideas in my own words, and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated, or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Pandit Deendayal Energy University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Tanmay Rathod

Roll No: 23MAI007

Date: ______________________

Abstract

Correct and effective sleep stage classification is crucial in the diagnosis of sleep disorders, conventionally done through manual interpretation of polysomnography (PSG) signals. Automated sleep stage classification is investigated in this thesis by a series of increasingly sophisticated models, from traditional machine learning to sophisticated deep learning architectures. Initial trials on the SleepEDF dataset compared models like Random Forest, XGBoost, and ensemble models, which obtained maximum accuracy of 85.35%, and set the foundation for automated systems to perform consumer-level sleep monitoring. Future work introduced deep learning models like BiLSTM and RNNs, and BiLSTM reached 81.13% accuracy, which proved to have better ability to model temporal sequences.

Based on these building blocks, we introduce *SleepGCN-Transformer*, a new hybrid model that combines Graph Convolutional Networks (GCNs) to learn spatial relations between multi-channel EEG, EMG, and EOG signals, and Transformer encoders to learn temporal patterns. The model applies Focal Loss to handle class imbalance and a CosineAnnealingLR scheduler for dynamic learning rate adaptation. Trained with the AdamW optimizer on preprocessed 30-second epochs, SleepGCN-Transformer is 93.12% training and 93.04% validation accurate. Feature attribution by LIME indicates EMG and EEG Pz-Oz to be the most significant channels, demonstrating the model's interpretability.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Sleep is a fundamental aspect of human health that greatly contributes to cognitive function, memory consolidation, and emotional regulation. Proper and quality sleep guarantees proper brain function and overall physical and mental health. Sleep disturbances result in several diseases, such as impaired concentration, mood disorders, and chronic conditions like cardiovascular diseases.

Sleep plays a critical function in sustaining physical and mental well-being. It is not a passive resting state but a sophisticated biological process crucial for memory consolidation, emotional homeostasis, metabolic health, and immune system function. Alterations in sleep quality or sleep patterns can reflect or lead to many health problems, such as insomnia, depression, cardiovascular disease, and neurodegenerative disorders. It is essential for clinicians and scientists to precisely assess sleep stages for the diagnosis of such conditions and the prescription of proper treatment plans.

Staging of the sleep stages is the basis of understanding the architecture of sleep and the diagnosis of sleep disorders. Sleep is divided into five major stages—N1, N2, N3, REM, and Wake—based on unique patterns seen in brain function. Every stage of sleep is associated with particular frequencies of brainwaves, which can be picked up using the technique of electroencephalography (EEG) signals. The precise staging is important to study the quality of sleep and to detect abnormalities like insomnia, sleep apnea, and narcolepsy.

1.0.1 The Role of EEG in Sleep Analysis

Electroencephalography (EEG) is among the major methods of tracking brain activity while asleep. EEG captures electrical activity created by neurons as they fire within the brain, recorded with electrodes applied to the scalp. These electrical signals are critical in distinguishing various stages of sleep since each stage has different patterns of brain wave activity. Whereas other physiological signals are indirect, EEG is a direct view into the electrical activity of the brain and hence the basis for research and diagnosis into sleep.

1.0.2 Understanding the SleepEDF Dataset

The Sleep-EDF (Sleep European Data Format) corpus, popular among researchers, comprises polysomnography (PSG) recordings taken from healthy subjects and mildly disordered sleep patients. Recordings are taken non-invasively and usually overnight in a laboratory setting. In the course of a session, several sensors are applied to the subject's body to record a number of different physiological signals such as EEG (electroencephalogram), EOG (electrooculogram), and EMG (electromyogram).

Specifically, the database contains two EEG channels, namely Fpz-Cz and

Pz-Oz, which record frontal and parietal brain activity. Additionally, it contains EOG signals to record eye movements and EMG signals to record muscle activity, particularly around the chin region. The recordings are manually annotated by experienced sleep technicians using visual patterns and set guidelines to mark sleep stages. The last annotation is retained in a hypnogram — a time series plot that shows the changes between various stages of sleep throughout the night.

1.0.3 Sleep Stages and Signal Patterns

Human sleep is commonly divided into five stages: Wake (W), Non-Rapid Eye Movement (NREM) stages N1, N2, and N3, and Rapid Eye Movement (REM). Each stage is characterized by specific frequency patterns in EEG signals:

TABLE 1.1
Brainwave Types and Their Characteristics in Sleep Staging

Wave Type	**Frequency Range**	**Associated Sleep Stage / Behavior**
Alpha waves	8–13 Hz	Relaxed wakefulness, especially with eyes closed
Beta waves	13–30 Hz	Alertness and active thinking; decrease during sleep
Theta waves	4–8 Hz	Light sleep (Stages N1 and N2)
Delta waves	0.5–4 Hz	Deep sleep (Stage N3); synchronized neuronal firing
Sleep spindles & K-complexes	12–15 Hz (spindles)	Characteristic of Stage N2; used for stage classification

These patterns are extracted from raw EEG signals through filtering and segmentation. Experts use these patterns, along with eye movement and muscle tone data, to classify each 30-second segment (epoch) into one of the five stages.

1.0.4 The Need for Automation

Manual scoring of sleep stages, while precise, is time-consuming and subject to inter-rater reliability. It takes hours of professional scoring for one night's worth of recording. Automated sleep stage scoring provides a quicker, more scalable, and more possibly consistent solution. Using machine learning and deep learning models, particularly those that can address intricate spatial and temporal patterns in physiological data, we can develop systems that emulate expert opinion and aid in clinical decision-making. This thesis suggests just such a system — *SleepGCN-Transformer* — that unifies graph-based representation of EEG sensor relationships with temporal learning from transformers. Our strategy not only targets high classification performance but also contributes to the increasing body of research in interpretable AI in medicine, moving us closer to clinically integrated, explainable, and fully automated sleep diagnostics.

Chapter 2

Literature Review

[1] The proposed architecture, Efficient Sleep Sequence Network (ESSN), has overcome the limitations of existing automatic sleep stage algorithms. This model addresses two main challenges. First, the model is quite complex, and often low-end systems are unable to process it; therefore, this model is designed to work on lightweight systems. The second challenge is the misclassification of the N1 stage, where models often confuse wake and REM stages. To address this, it introduces the N1 structure loss function. The ESSN model has achieved impressive metrics: 88.0% accuracy, 81.2% macro F1, and 0.831 Cohen's kappa. These results were obtained on the SHHS dataset. Additionally, it has reduced computational requirements, with only 0.27M parameters and 0.35G floating-point operations, and it claims to be faster than models like L-SeqSleepNet.

[2] The Multi-Domain View Self-Supervised Learning Framework (MV-TTFC) introduces a new approach to classify sleep stages by leveraging self-supervised learning (SSL) on unlabeled EEG data. By incorporating multi-view representation technology, this model enhances information exchange across different views. It also introduces the multisynchrosqueezing transform, which improves the quality of the time-frequency view. Ultimately, it captures the latent features within EEG signals. It was evaluated on two datasets (SleepEDF-78 and SHHS), and MV-TTFC achieved state-of-the-art performance with accuracies of 78.64% and 81.45%, and macro F1-scores of 70.39% and 70.47%, respectively.

[3] The proposed CNN-Transformer-ConvLSTM-CRF hybrid model presents a new integration method between local and global feature extraction to enhance the classification ability of sleep stages. The model can identify relationships among EEG features by applying a multi-scale convolutional neural network combined with a Transformer for encoding features of the EEG signal and a spatio-temporal encoder via ConvLSTM. Additionally, the adaptive feature calibration module improves the extracted features, and there is efficient learning of the transition relationships between the stages of sleep by the CRF module. Based on evaluations on three datasets, this hybrid model outperforms existing state-of-the-art methods, demonstrating its efficacy in sleep stage classification.

Chapter 3

Automated Sleep Staging System with EEG Signal using Machine Learning Techniques

3.1 Methodology

We begin by importing the Sleep-EDF data provided in EDF (European Data Format). The raw EEG recordings are loaded using MNE's `read_raw_edf` function, and corresponding annotations are aligned using `read_annotations`. We optionally exclude all non-EEG channels unless explicitly required, focusing on EEG signals for our analysis. The dataset includes two EEG channels: `Fpz-Cz` and `Pz-Oz`, which are the primary input sources in our work.

During loading, we crop the signal to reduce unnecessary wake-time data, retaining 30 minutes before the first sleep stage and 30 minutes after the last. This ensures that our training data remains focused around relevant sleep activity. The channel names are cleaned by stripping prefixes for standardization.

After loading and cropping, we segment the data into fixed-length 30-second epochs using the known sampling frequency of the recordings. Only epochs labeled with valid sleep stages — specifically stages W, N1, N2, N3, and REM — are retained. For each epoch, the raw data is extracted and structured into a format suitable for training.

Once all epochs are collected, we flatten the 3D EEG data (epochs $\times$ channels $\times$ time points) into 2D arrays (epochs $\times$ features). This is necessary for traditional machine learning classifiers. To address class imbalance in sleep stages, we apply SMOTE (Synthetic Minority Oversampling Technique) to generate balanced training samples.

The balanced dataset is then split into training and testing sets in an 80-20 ratio. We train a Machine Learning on the training data. After training, predictions are made on the test set, and the model is evaluated using standard metrics including classification report and confusion matrix. The confusion matrix is visualized using a heatmap to give a clear view of the stage-wise performance of the classifier.

3.2 Dataset Information

We have used the Sleep-EDF Dataset available at `https://www.physionet.org/content/sleep-edfx/1.0.0/`. For our experiments, we selected recordings from approximately 30 patients, resulting in 60 files — 30 EDF (European Data Format) recordings and their corresponding hypnogram annotation files. The EDF files include various bio-signals captured through multiple channels. The

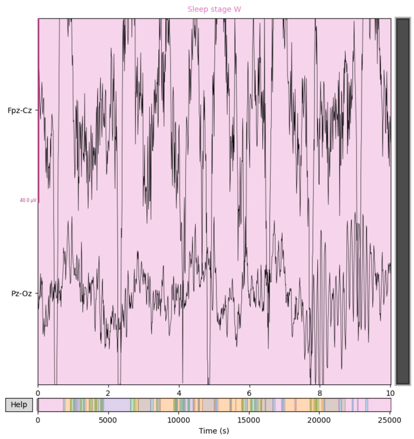

Figure 3.1. Filter EEG Chanels

hypnogram files contain sleep stage annotations, specifying the start time and type of each stage (Wake, N1, N2, N3, N4, REM) with timestamps in the HH:MM:SS format.

The EDF recordings consist of a range of physiological signals, including EEG (`Fpz-Cz` and `Pz-Oz`), EOG, EMG (submental), rectal temperature, and oro-nasal respiration. Each channel provides continuous signal recordings at a specific sampling frequency (Hz), which are used as input features for our downstream machine learning pipeline. From these signals, we derive two essential inputs: spatial features, used to construct graph-based representations, and temporal features, which capture the sequence and timing dynamics necessary for modeling the sleep process.

3.3 Preprocessing

This study applied machine learning techniques to classify sleep stages using EEG signals. The preprocessing pipeline included several crucial steps: channel mapping, data cropping, signal filtering, epoch creation, label mapping, and class balancing using SMOTE. These steps ensured clean, well-structured input for feature extraction and model training. The extracted features spanned time-domain, frequency-domain, and time-frequency representations. Multiple machine learning models—Gradient Boosting, Random Forest, Support Vector Machine (SVC), and Bagging—were trained and evaluated using metrics such as accuracy, precision, recall, and F1-score with cross-validation to assess generalization. The trained models were finally used to classify unseen EEG data into five sleep stages: Wake, N1, N2, N3, and REM.

A. Dataset Acquisition EEG signals were recorded using non-invasive electrodes placed on the scalp of human volunteers, capturing brain electrical activity during overnight sleep studies. Additional demographic data such as age, gender, medical history, and sleep habits were collected to provide contextual understanding. The data collection took place in controlled conditions to minimize noise and interference. Manual inspection of the raw EEG was performed to reject or correct artifacts where necessary, ensuring the quality of the input data. The final recordings were saved in standard EDF format for compatibility with downstream analysis. The dataset used in this study was sourced from publicly available Sleep-EDF recordings, which include EEG data and associated hypnogram annotations reflecting sleep stages.

B. Data Preprocessing Preprocessing was applied to prepare EEG data for modeling. Channel mapping was done to identify relevant EEG channels, particularly `Fpz-Cz` and `Pz-Oz`, and irrelevant or noisy channels were excluded. Data cropping was applied to isolate the sleep period from wake periods based on annotation markers. The signals were then filtered to remove noise and isolate specific frequency bands important for sleep analysis.

Epochs of 30-second duration were created from continuous EEG signals to structure the data for machine learning. Sleep stage labels corresponding to these epochs were derived from hypnogram annotations. Given the naturally imbalanced class distribution in sleep stage data (e.g., fewer REM and N1 stages), the SMOTE algorithm was employed to oversample the minority classes, thereby enhancing model robustness. After preprocessing, the dataset was split into training and testing subsets. This enabled reliable training of models and unbiased performance evaluation. The output from preprocessing was then used in the feature extraction and modeling stages.

3.4 Model Architecture and Learning Framework

The core modeling pipeline begins with transforming the raw EEG data into structured epochs suitable for machine learning classification. Each EEG recording is segmented into 30-second epochs using the sampling frequency and annotated sleep stages. Only valid sleep stages (Wake, N1, N2, N3, and REM) are retained for downstream processing. After epoching, each segment is extracted using the `mne.Epochs` function, resulting in structured EEG data with consistent time windows. These segments are then reshaped into two-dimensional feature vectors where each row corresponds to an epoch and each column to a time-series sample across all EEG channels.

To address the class imbalance inherent in sleep stage data, the SMOTE (Syn-

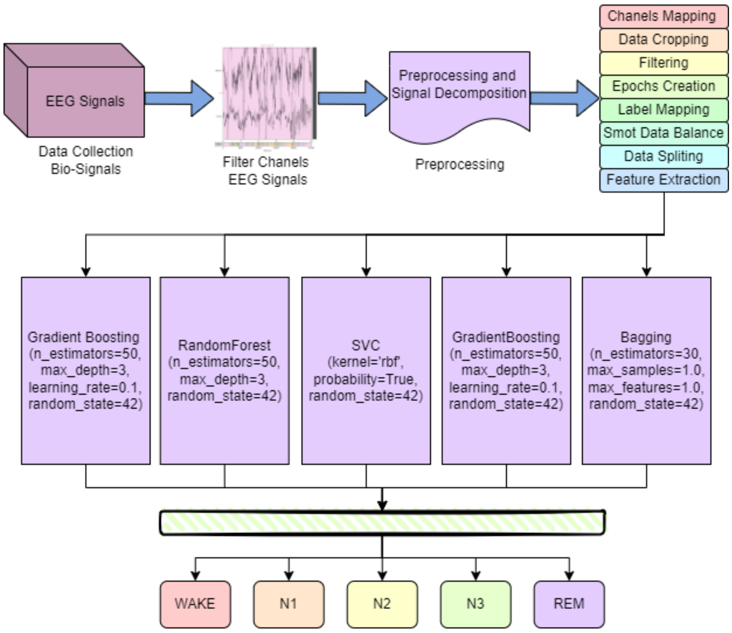

Figure 3.2. Architecture of Machine Learning Classifiers for Sleep Stage Classificatio

thetic Minority Over-sampling Technique) algorithm is applied. This balances the dataset by synthetically generating new examples in underrepresented classes. The balanced dataset is then split into training and testing sets in an 80:20 ratio.

3.4.1 Random Forest Classifier

We implemented a Random Forest classifier to establish a strong baseline model for sleep stage classification. The raw EEG recordings were segmented into 30-second epochs, and features were extracted by flattening the multidimensional epoch data. To handle class imbalance, the SMOTE algorithm was applied prior to training. The model was initialized with **100 decision trees** (`n_estimators=100`) and a fixed **random seed** (`random_state=42`) to ensure reproducibility. This configuration allowed the model to generalize well while maintaining robust performance across varying sleep stages.

3.4.2 Gradient Boosting Classifier with PCA

To improve computational efficiency and potentially boost model accuracy, we introduced Principal Component Analysis (PCA), retaining **95% of the explained variance** (`n_components=0.95`). The reduced feature set was passed to a **Gradient Boosting classifier** configured with **30 estimators** (`n_estimators=30`), a **maximum tree depth of 3** (`max_depth=3`), and a **learning rate of 0.1**. Class balancing was again addressed with SMOTE. The same 80-20 train-test split was maintained with `random_state=42`. This model leverages the sequential learning of weak classifiers to optimize classification performance over multiple iterations.

3.4.3 Ensemble Learning (Voting Classifier)

For further enhancement, we employed a soft voting ensemble that combines multiple classifiers, each with complementary strengths. The ensemble includes a **Gradient Boosting Classifier** (`n_estimators=50`, `max_depth=3`, `learning_rate=0.1`), a **Random Forest Classifier** (`n_estimators=50`, `max_depth=3`), a **Support Vector Classifier (SVC)** with an RBF kernel and probability estimates enabled, and a **Bagging Classifier** (`n_estimators=30`, `max_samples=1.0`, `max_features=1.0`). These classifiers were integrated using a **soft voting strategy** in the `VotingClassifier(voting='soft')` to average their probabilistic predictions.

Feature reduction was again performed using PCA with 95% variance retention. This ensemble approach leverages model diversity to achieve improved generalization and more stable predictions across sleep classes.

3.5 Results and Evaluation

This section presents the results and evaluations for the models trained, including Random Forest, Bagging, Ensemble Learning, and Gradient Boosting. The following figures summarize the performance metrics for each model.

3.5.1 Random Forest

For the Random Forest model, we present the following evaluation metrics:

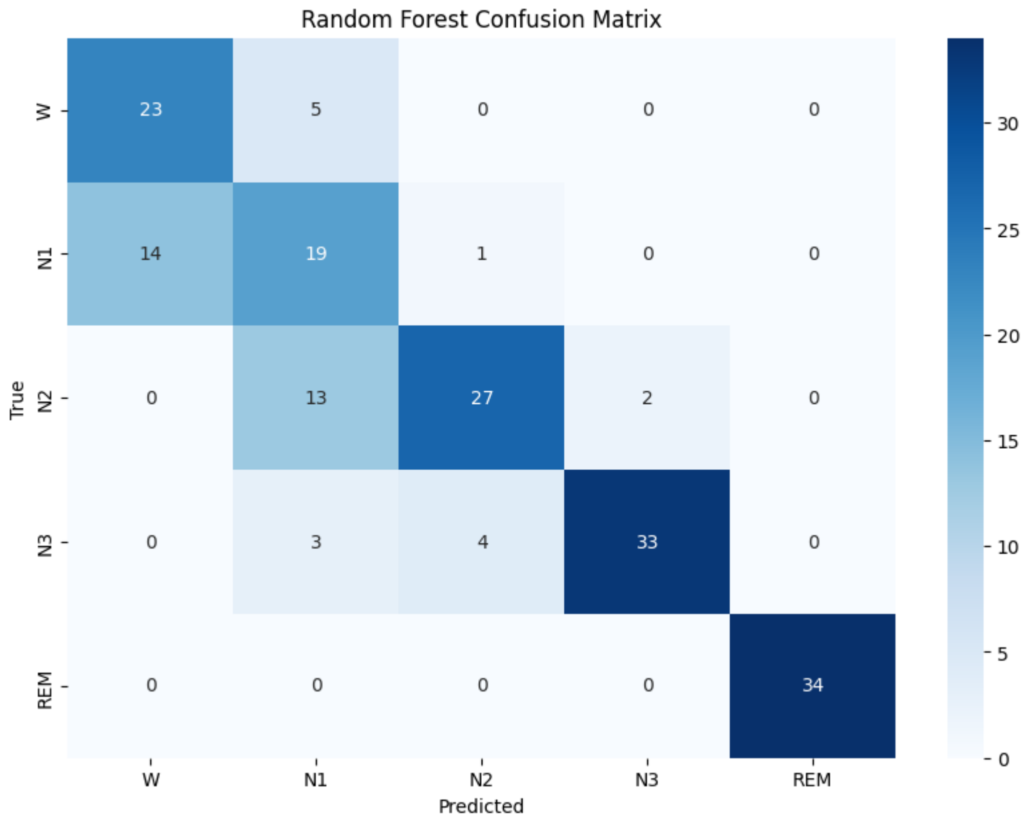

Figure 3.3. Random Forest Confusion Matrix

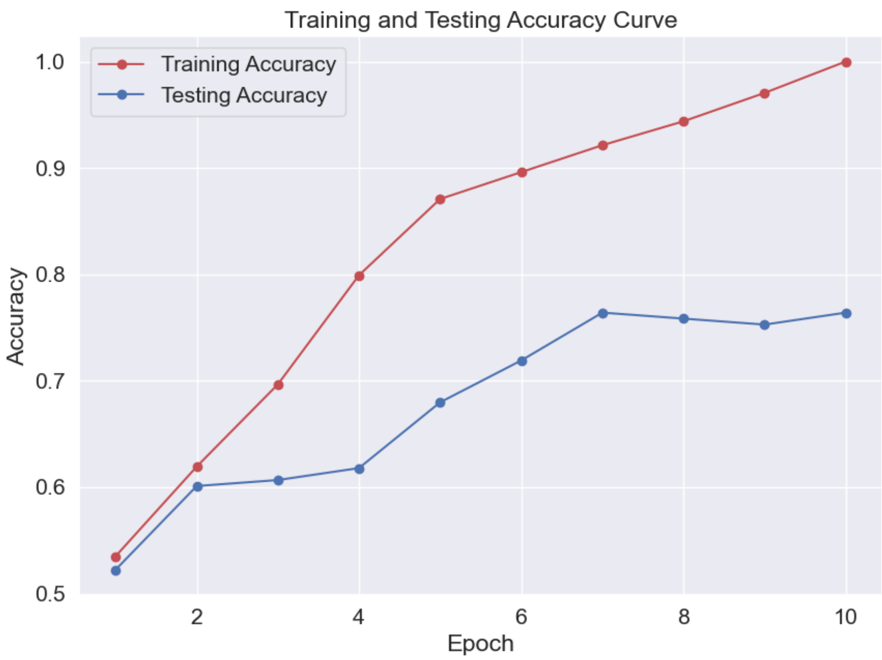

Figure 3.4. Random Forest Accuracy Curve

Figure 3.5. Random Forest Loss Curve

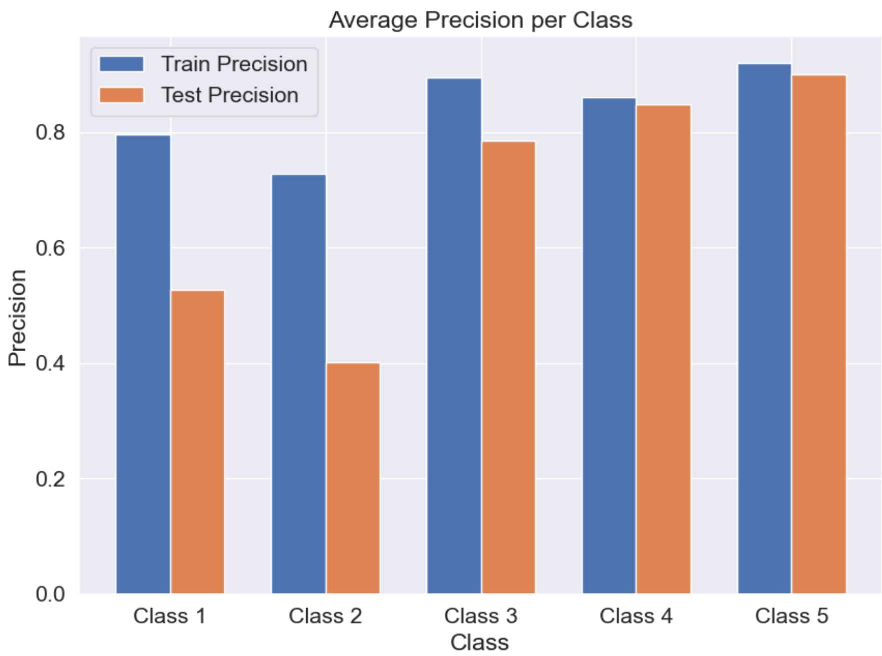

Figure 3.6. Random Forest Precision Per Class

3.5.2 Bagging Classifier

For the Bagging Classifier model, we present the following evaluation metrics:

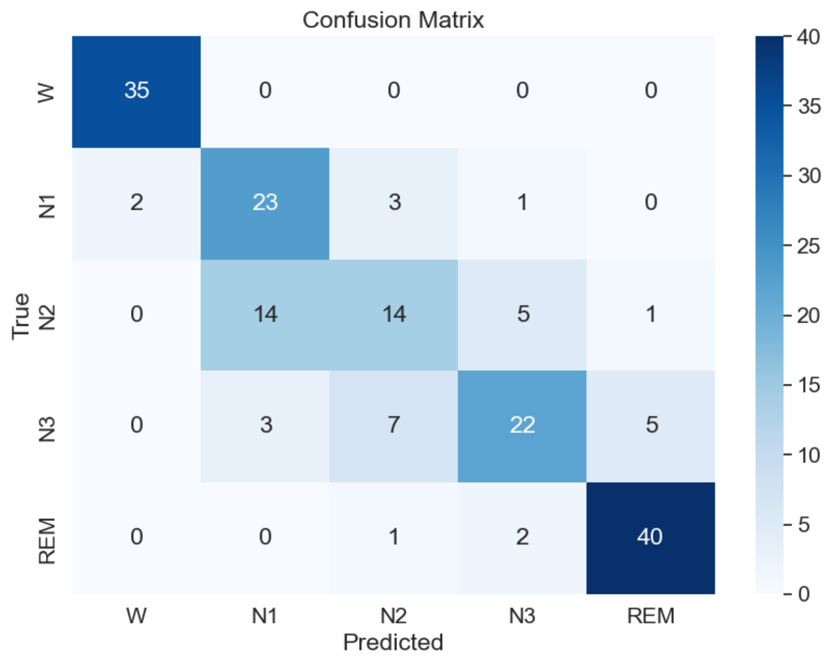

Figure 3.7. Bagging Classifier Confusion Matrix

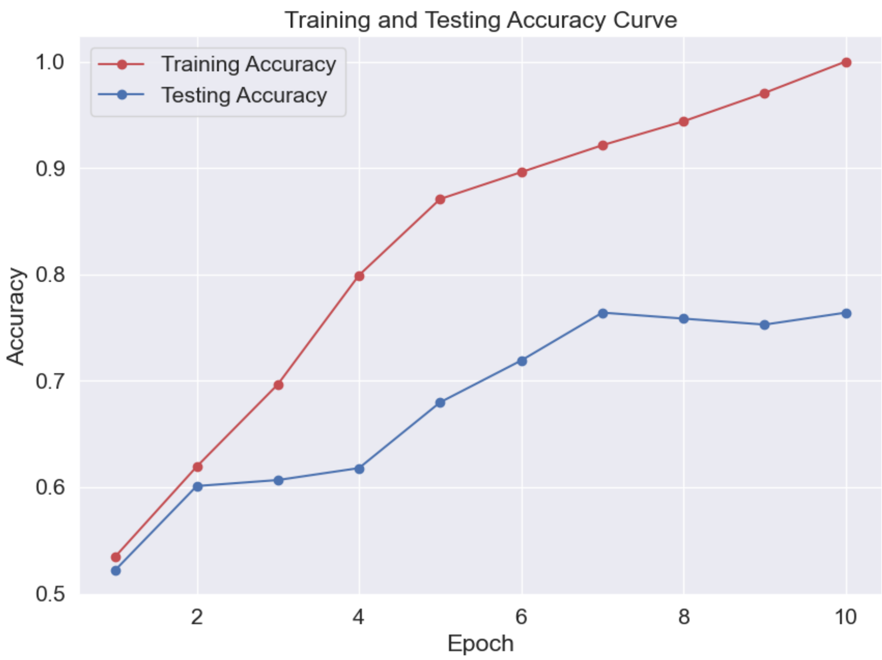

Figure 3.8. Bagging Classifier Accuracy Curve

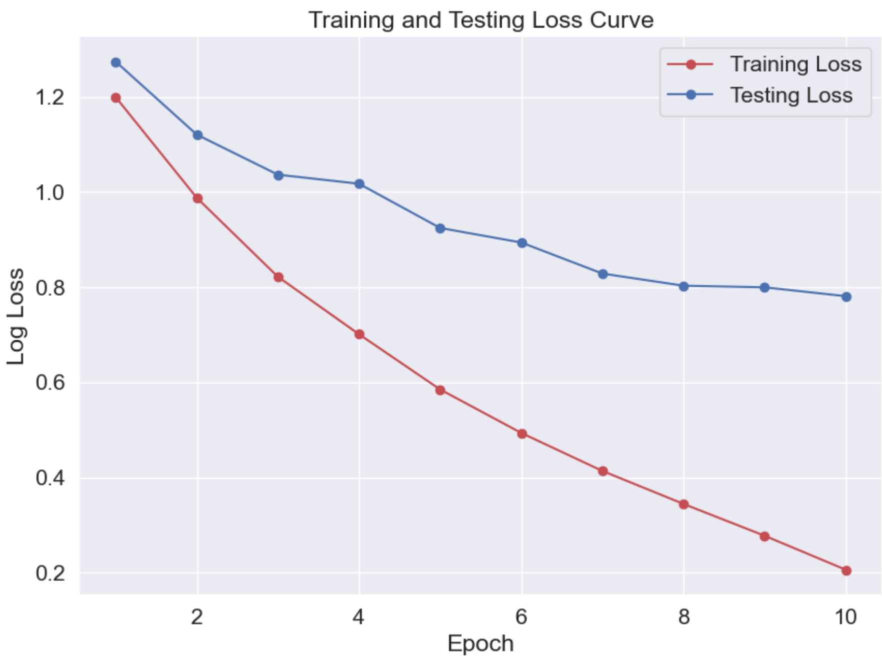

Figure 3.9. Bagging Classifier Loss Curve

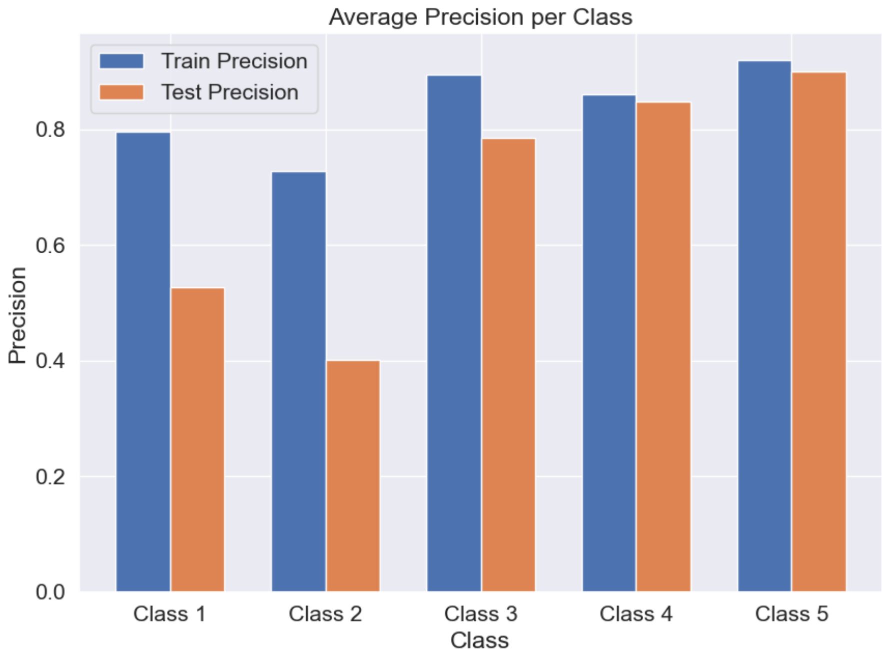

Figure 3.10. Bagging Classifier Precision Per Class

3.5.3 Ensemble Learning

For the Ensemble Learning model, we present the following evaluation metrics:

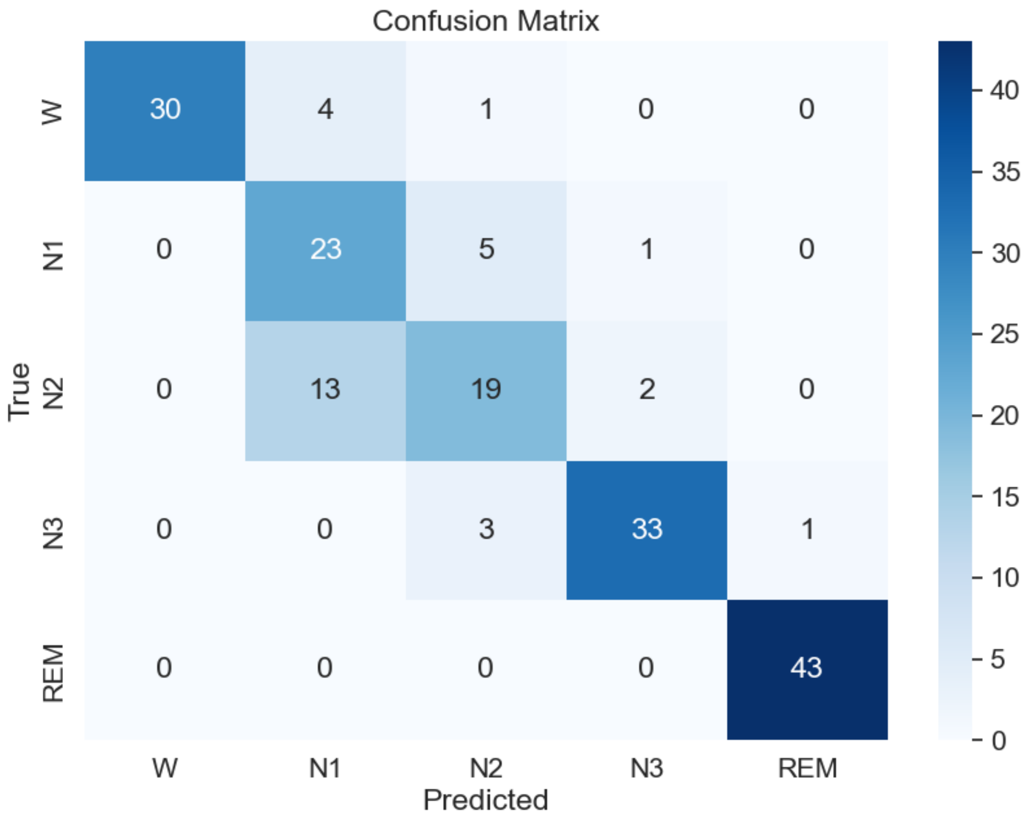

Figure 3.11. Ensemble Learning Confusion Matrix

Figure 3.12. Ensemble Learning Accuracy Curve

Figure 3.13. Ensemble Learning Loss Curve

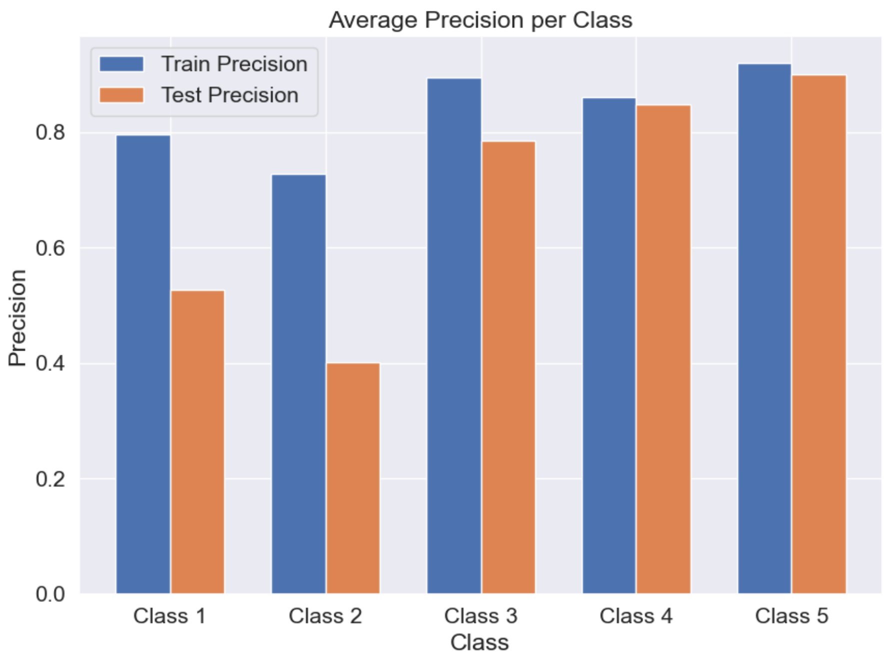

Figure 3.14. Ensemble Learning Precision Per Class

3.5.4 Gradient Boosting

For the Gradient Boosting model, we present the following evaluation metrics:

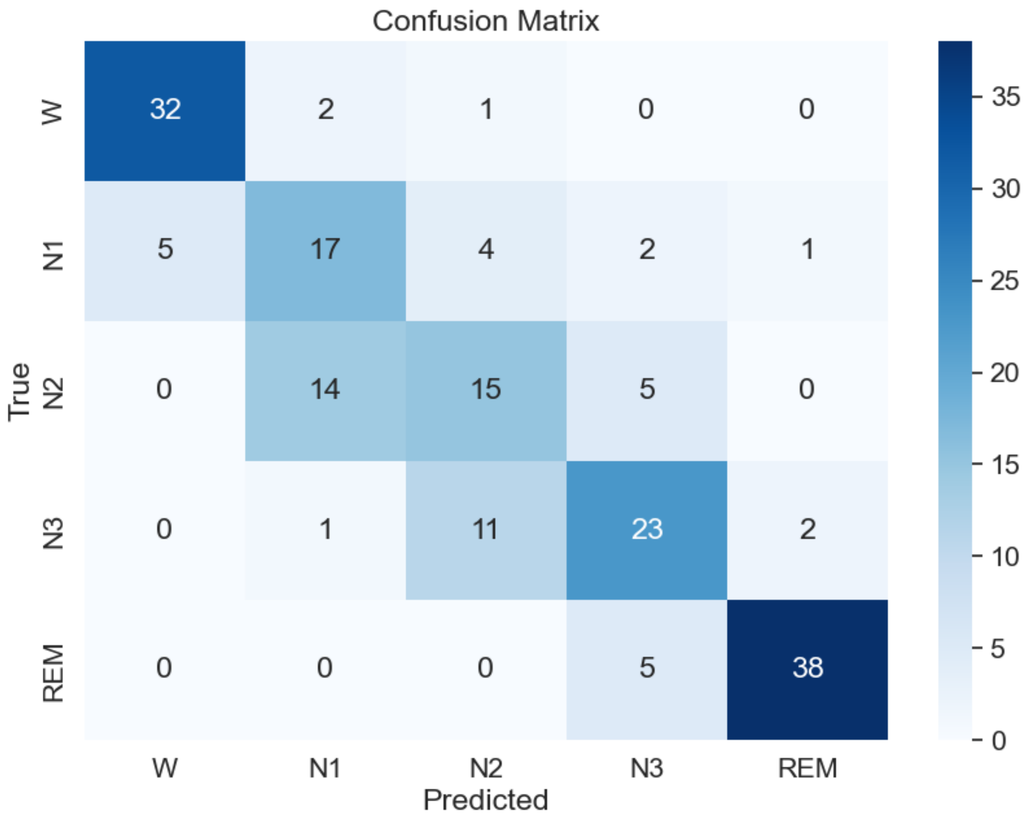

Figure 3.15. Gradient Boosting Confusion Matrix

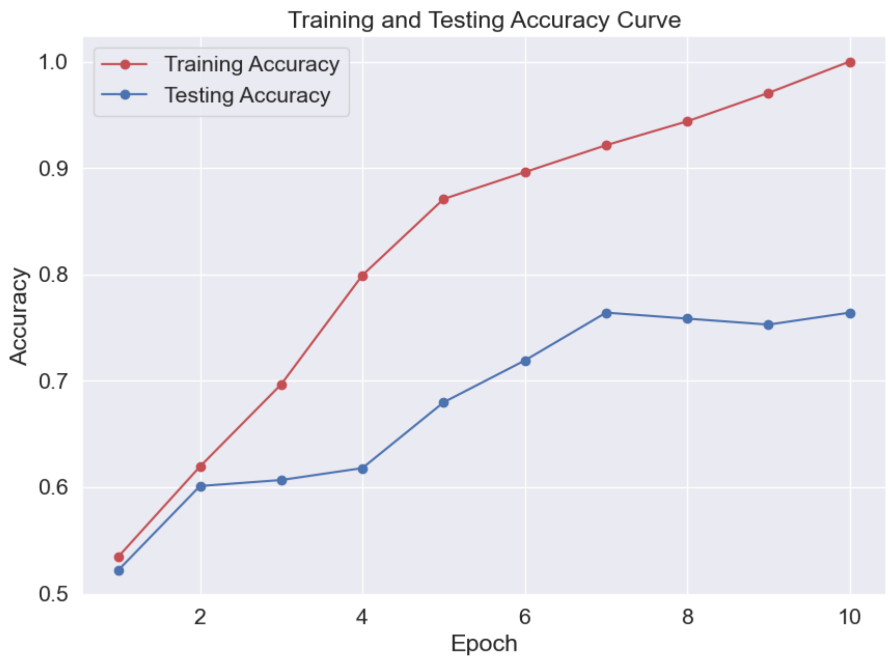

Figure 3.16. Gradient Boosting Accuracy Curve

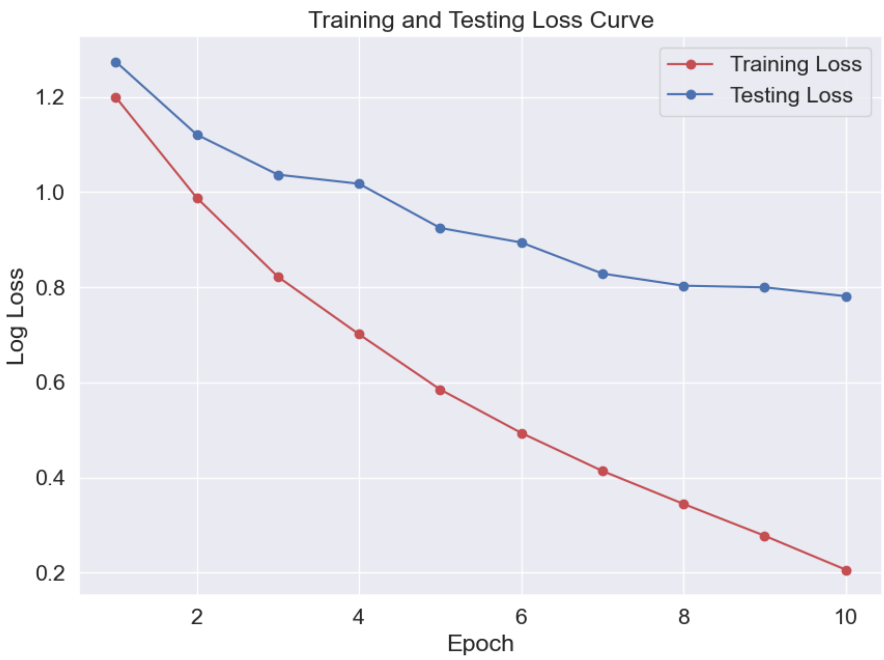

Figure 3.17. Gradient Boosting Loss Curve

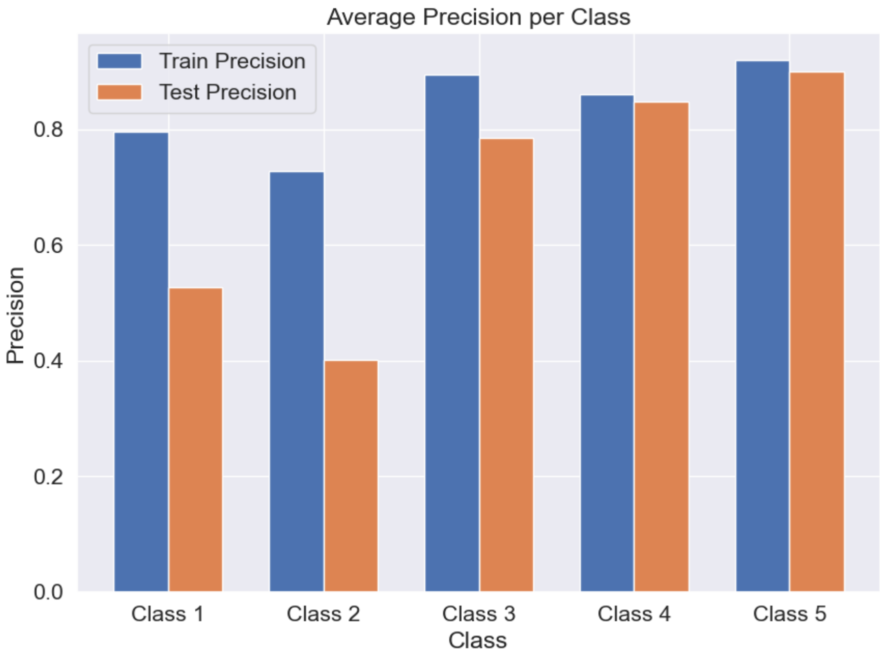

Figure 3.18. Gradient Boosting Precision Per Class

Chapter 4

Deep Neural Model for Automated Sleep Staging System using Single-Channel EEG Signal

4.1 Methodology

4.2 Dataset Information

4.3 Preprocessing Techniques

4.4 Model Architecture and Learning Framework

4.5 Results and Evaluation

Chapter 5

SleepGCN-Transformer: A Hybrid Graph-Convolutional and Transformer Network for Sleep Stage Classification

5.1 Methodology

5.2 Dataset Information

5.3 Preprocessing Techniques

5.4 Model Architecture and Learning Framework

5.5 Results and Evaluation

Chapter 6

Conclusion

Our proposed SleepGCN-Transformer model achieves 93.12% training accuracy and 93.04% validation accuracy, demonstrating its effectiveness in sleep stage classification. The integration of Graph Convolution Networks (GCN) captures spatial dependencies across EEG, EOG, and EMG channels, while the Transformer extracts temporal patterns. The use of Focal Loss enhances class balancing, improving performance on underrepresented sleep stages. Feature importance analysis highlights EMG and EEG Pz-Oz as key predictors. This robust approach lays the foundation for future work in Explainable AI, enabling medical professionals to interpret AI-driven sleep diagnostics effectively.

Chapter 7

Reference