Automated Sleep Staging System with EEG Signal using Machine Learning Techniques

Santosh Kumar Satapathy
Information and Communication Technology
Pandit Deendayal Energy University
Gandhinagar, Gujarat
Santosh.Satapathy@sot.pdpu.ac.in

Tanmay Rathod
Information and Communication Technology
Pandit Deendayal Energy University
Gandhinagar, Gujarat
23mai007@sot.pdpu.ac.in

Nibedita Das
Information and Communication Technology
Institute of Advanced Research
Gandhinagar, Gujarat
nibedita.das@iar.ac.in

Rajesh Kumar Mohapatra
Information and Communication Technology
Pandit Deendayal Energy University
Gandhinagar, Gujarat
rajesh.mphd23@sot.pdpu.ac.in

Suren Sahu
Faculty of Emerging Technology
Sri Sri University
Cuttack, India
suren.sahu2602@gmail.com

Jaynil Joshi
Information and Communication Technology
Pandit Deendayal Energy University
Gandhinagar, Gujarat
Jaynil.jict21@sot.pdpu.ac.in

2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI) | 979-8-3315-2169-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/IATMSI64286.2025.10985276

***Abstract*—Advances in signal processing and machine learning have revolutionized the analysis of physiological signals, providing much better results as compared to traditional manual methodology. This paper focuses on automatic sleep stage classification via EEG channels Fpz and Cz extracted from the publicly available SleepEDF PSG Hypnogram dataset. Those channels were preprocessed to improve their quality and mapped the bio-signals to corresponding sleep stages, which is a decisive step for accurate classification to be achieved. It utilized the sleep stage scoring through various machine learning approaches, including Random Forest, Gradient Boosting, Bagging Classifier, and an Ensemble Learning approach. This yielded classification accuracies as high as 78% with Random Forest, 79% with Gradient Boosting, 75% with Bagging Classifier, and 85% with Ensemble Learning last one being the most promising model. This work highlights the relevance of EEG channels in classifying sleep stages and demonstrates how machine learning can be exploited to move the analysis of sleep data forward. Implications from these findings go towards embedding automated sleep stage classification in consumer-grade sleep monitoring systems that would contribute towards more efficient and accessible sleep health monitoring solutions.**

Keywords*— *Automated sleep stage scoring (ASSS), Machine Learning (ML), Electroencephalogram (EEG), Feature Engineering (FE), Electrooculogram (EOG)

I. INTRODUCTION

Sleep is a crucial component of human health, encompassing several key stages that impact cognitive function and overall well-being. Research indicates that modern lifestyles, work-related stress, and living conditions negatively affect sleep duration, continuity, and quality, leading to an increasing number of sleep issues, particularly among children. Common sleep disorders like insomnia, obstructive sleep apnea, periodic limb movement disorder, parasomnia, and bruxism can significantly disrupt daily functioning [1]. Accurate classification of sleep stages and assessment of sleep health are essential for diagnosing and treating these disorders. Traditionally, polysomnography (PSG) has been the standard for monitoring sleep stages. However, sleep specialists or physicians usually classify PSG data into various sleep stages by reviewing 30-second or specific time intervals based on established criteria[2]. This process is known to be labor-intensive and time-consuming. In contrast, machine learning algorithms can significantly reduce the time needed for classification, while experts have traditionally used features from the time domain, frequency domain, and time-frequency domain to preprocess data for machine learning. Selecting effective features is crucial for improving the performance of traditional classifiers [3]. EEG is the most prominent method for sleep stage identification because it can measure the brain's activity using particular wave frequencies [4]. Other supplemental signals utilized are electrooculogram (EOG) for eye movements and electromyogram (EMG) for activity in the muscles [5]. The data are recorded in the European Data Format (EDF) and are readily available for use.Fig.1 illustrates the recorded bio signals.

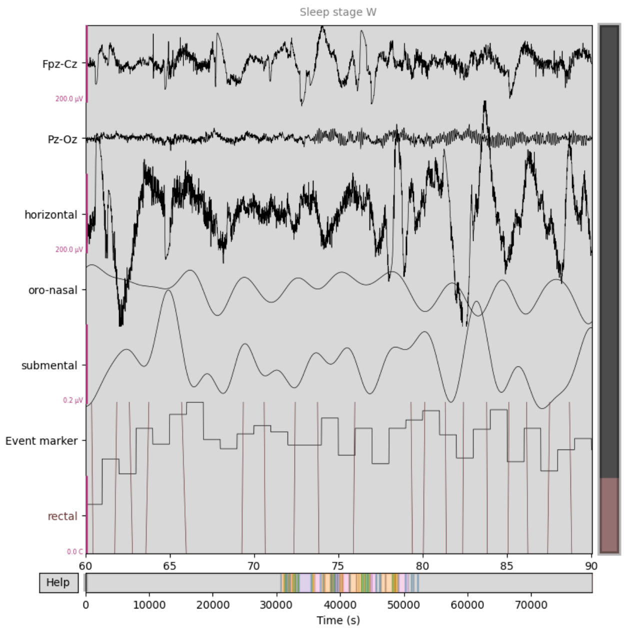

Fig. 1. Recorded Bio Signals Channels which include EEG, ECG, EMG, and EOG

Automated sleep analysis involves preprocessing, feature extraction, and data labeling, which are critical for standardizing sleep stage classification [6]. Key features include power spectral density, brain wave amplitude, frequency, and entropy measurements. These features serve as input for ML models, which aim to enhance classification accuracy [7]. By tapping into the strength of wearable technology and potent computational algorithms, automated systems can be used to close gaps in sleep research by opening doors to more precise and accessible solutions for the monitoring of sleep and its disorders [8-9].

II. RELATED WORK

In [1] developed a deep residual convolutional network (ConvNet)-based model for sleep stage classification using single-channel EEG data. In this study, ConvNets learn, extract, and classify the raw EEG signals directly, eliminating the need for traditional feature extraction steps almost completely. The method obtained 87.5% accuracy, indicating that ConvNets work well with raw EEG data as the model does seem to improve the classification performance. The authors pointed out that ConvNets have the potential to develop automatic methods of sleep stage classification and offer a very relevant platform to further extend applications in EEG analysis.In [2] conducted a study investigating the performances of various CNN architectures to predict sleep stages, based on single-channel EEG data. Based on their comparative analysis, CNN-based models were able to achieve an overall classification accuracy of 86.2% with minor variances across architectures reflecting the versatility of CNNs to cater EEG processing while requiring considerable design of architectures to optimize for performance. Relatively, they have recognized the capability of CNNs to work as an enhanced feature extraction tool for EEG analysis but further research should optimize architecture design.In [3] have put forward a deep learning approach to perform sleep stage classification employing single-channel EEG data. It has been developed to process raw EEG signals and achieved an accuracy of 89.1%, surpassing traditional techniques. The study successfully demonstrates the ease with which even basic yet effective deep learning models can analyze raw EEG data for sleep stage classification. The suggested perspective does imply that the methodology could allow some extension to a number of other EEG-based applications with minor modifications being made.In [4] and others applied CWT and transfer learning in classifying the stages of sleep on single-channel EEG data. Transforming EEG signals into images of CWT made it possible to use pretrained models for transfer learning, thus achieving a high accuracy of 90.7 percent. The study demonstrated this combined use of CWT and transfer learning to improve classification performance specifically on more complex data. This new approach greatly advances EEG signal analysis.In [5] about a transfer learning-based framework for sleep stage classification employing single-channel EEG data. Their methodology focused on scalable transfer learning into generalize across datasets using a classifier with an 88.9% accuracy. The contributions of the work underlined the significance of transfer learning in addressing dataset variances and improving EEG model adaptability.In [6] tackled experimentation on the use of deep transfer learning to classify sleep stages with single channel EEG data. Their approach achieved an accuracy of 90.3%. Its effectiveness in dealing with the variability within EEG signals is amply highlighted by the study. Deep transfer learning can tap knowledge from large-scale pretrained models, thereby making such a kind of approach most promising for EEG signal analysis.In [7] have implemented Sleep EEG Net, a sequence-to-sequence deep learning model, to extract temporal dependencies from EEG data for sleep stage classification. This method, applied to single-channel EEG data, achieved the reported highest accuracy so far, reaching 92.5%. The study highlighted the capabilities of sequence-based models to extract intricate temporal dynamics from EEG signals and therefore created a benchmark for future research into sleep staging.In [8] introduced Sleep Context Net, a temporal context network for enhancing sleep stage classification from single-channel EEG data. Their method achieved an accuracy of 90.4%, underlining the significance of temporal context modelling in EEG analysis. In this paper, it is shown that the inclusion of temporal features dramatically improves the accuracy and robustness of sleep stage classifiers.In [10], proposed X Sleep Net, a multi-view sequential model of single-channel EEG data for sleep stage classification. They reported the highest accuracy among the studies reviewed, with an accuracy of 93.2%. X Sleep Net improved upon the relatively bad performance by combining many perspectives of EEG data and established a new standard for multi-view EEG signal analysis. The study shows the extensive potential for multi-view approaches to pave the way for EEG-based applications.

A review of the literature showed that the Sleep-EDF dataset consisted of 20 subjects, while the Sleep-EDF expanded dataset, an extension of Sleep-EDF, included 78 subjects. This increased dataset size suggests a potential improvement in sleep staging accuracy. However, studies [17, 18, 19, 20, 23, 24] revealed that, when using the same model, the Sleep-EDF expanded dataset resulted in a decrease in accuracy of 1.3% to 4.8% compared to the Sleep-EDF dataset. Further analysis of the differences between the two datasets revealed that participants in the Sleep-EDF dataset were aged 25 to 34, while those in the Sleep-EDF expanded dataset ranged from 25 to 101 years old. This study indicates that the age of participants affects the performance of sleep staging models, with the inclusion of older participants leading to a decrease in model performance.

A. Contribution

The strength of our contribution lies in the simplification of sleep stage classification by not having to delve into advanced computing models but instead opting for simple machine-learning techniques. With simple yet powerful ML methods, we are able to attain high accuracy rates for various sleep stages-classifications: wake, light sleep (N1, N2), deep sleep (N3), and REM sleep. Unlike many other computationally intensive approaches, ours is focused on straightforward, efficient classification applicable to real-world settings. This work used the Sleep Physionet dataset, containing EEG, EMG, and EOG signals, for extracting the relevant features toward sleep stage classification. For robust results across all the stages of sleep, class imbalance handling with SMOTE was applied, along with the Random Forest algorithm. This visualization of outcome classification gives insight into areas of strength and weakness that our model possesses, in turn opening a clear line of approach to improvement in the future. It constitutes a scalable and accessible, automated solution for sleep monitoring, potentially to be carried out clinically, by accessing open data, thereby supporting the progress of research on clinical sleep monitoring and management issues.

III. METHODOLOGY

This study applied the approach of machine learning to classify sleep stages from EEG signals. It collected and preprocessed the EEG signals, which involved channel mapping, data cropping, filtering, epoch creation, label mapping, and SMOTE data balancing to deal with the class imbalance. The features of interest were extracted from the preprocessed EEG signals, including time-domain, frequency-domain, and time-frequency characteristics. The ability of a suite of machine learning classifiers, such as Gradient Boosting, Random Forest, Support Vector Machine (SVC), and Bagging, to classify sleep stages was tested. The performance of each classifier was evaluated in terms of accuracy, precision, recall, and F1-score with cross-validation applied for estimation of generalization performance. The models so trained are then used to classify new EEG data into one of the five sleep stages, namely Wake, N1, N2, N3, and REM. The work attempts to apply the power of machine learning to the task of sleep stage classification from EEG signals, with potential applications in both sleep research and clinical settings. Fig.2 presents the complete structure of the automated sleep staging system.

A. Dataset Acquisition

EEG signals, a bio-signal that mirrors the brain's electrical activity, was first recorded. Probably it was done with non-invasive electrodes that were affixed to the scalp of the volunteers. Other aspects of the volunteers' subjects' demographics were also acquired, including their age, gender, sleeping habits, and medical history to provide valuable contextual information accompanying the EEG recording. The recording itself, was performed under controlled conditions to minimize the noises and interferences that could disturb it. The process generally covers overnight to cover all levels of the spectrum of sleep. Data so recorded via EEG was looked into visually. Some even involve rejection or correction artifacts as may be appropriate to the circumstances so the data quality is not compromised in the slightest

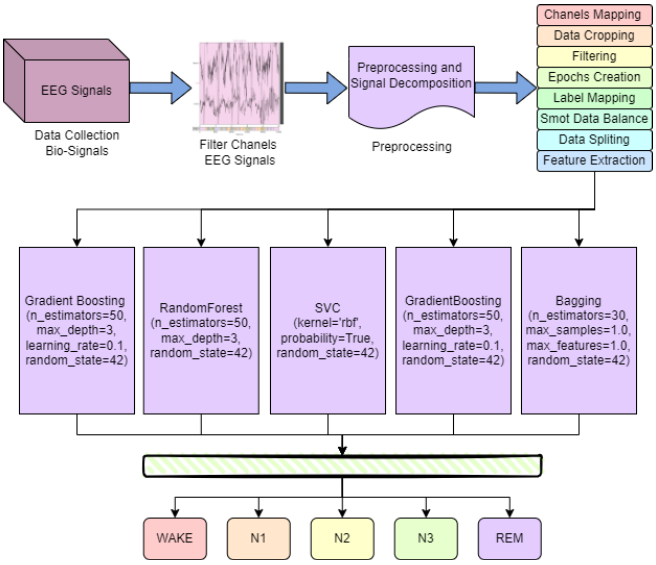

Fig. 2. Architecture of Machine Learning Classifiers for Sleep Stage Classification

. The EEG data was finally saved in an appropriate format to be further processed and analyzed. Quality acquisition of EEG data is an important success factor in later analysis and classification steps, and thus it would be significant to pay attention to all these aspects during the data collection process. This data on EEG was sourced from an openly available dataset or one that had previously been recorded. The dataset probably included the recordings of EEG from a group of participants who were undertaking PSG, which is a gold-standard sleep study incorporating recordings of EEG along with other physiological measures. It would likely contain information about the demographics of the subjects, the sleeping habits, and medical history, if any. The EEG recordings in this study were most likely obtained with standard procedures having appropriate electrode placement and recording conditions. The quality of EEG data in the dataset, for the purposes of the study, is assumed to be good.

B. Data Preprocessing

Preprocessing of EEG signals was done in preparation for feature extraction and classification in subsequent steps. This includes channel mapping to assign the relevant location of the electrode to each EEG signal, data cropping by segregating continuous EEG signals into smaller epochs, and filtering where noise is removed thus isolates specific frequency bands relevant to different stages of sleep. Fig.3 presents preprocessed signals.

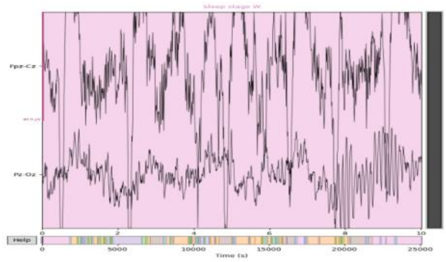

Fig. 3. Signals after preprocessing

Epochs were created in order to represent small parts of the EEG signal. Sleep stages were also manually labeled based on standard criteria for sleep scoring. Naturally occurring class imbalance in this dataset of sleep stage was reduced by SMOTE, or Synthetic Minority Over-sampling Technique, in over-sampling minority classes. Finally, the preprocessed data were divided into training and testing sets for model training and evaluation and feature extraction for extracting characteristics relevant to the EEG signals. All of these steps are important in ensuring data quality and the removal of noise in preprocessing to prepare the data to become effective machine learning inputs in sleep stage classification.

C. Classification Algorithm

1) Random Forest

Random Forest algorithm is a powerful tree-learning technique in Machine Learning. It works by creating several Decision Trees during the training phase. Each tree is constructed using a random subset of the data set to measure a random subset of features in each partition. This randomness introduces variability among individual trees, reducing the risk of overfitting and improving overall prediction performance.

2) Ensemble Learning

Ensemble learning is a machine learning technique that combines the predictions from multiple individual models to obtain a better predictive performance than any single model. The basic idea behind ensemble learning is to leverage the wisdom of the crowd by aggregating the predictions of multiple models, each of which may have its own strengths and weaknesses. This can lead to improved performance and generalization.

3) Model Training and Testing

A random forest classifier was chosen for training and evaluation due to its robustness in handling a variety of feature types and its capability to manage high-dimensional data: Data Splitting: The resampled data was split into an 80-20 ratio for training and testing sets, respectively, to validate model performance. Model Training: The Random Forest Classifier was initialized and fitted with the training dataset. Hyperparameters were set to use 100 estimators, enhancing its generalization capability. Prediction and Evaluation: The model was used to make predictions on the test dataset, and performance evaluation was done using classification metrics such as accuracy, precision, recall, and F1-score. Confusion Matrix Visualization: A confusion matrix was generated for graphical representation of the classifier's performance across different sleep classes, displayed using Seaborn's heatmap functionality to provide insights into misclassifications.

IV. Result Analysis and Discussion

A. Random Forest

In this study, we used several machine learning techniques to classify our sleep stage classification first technique is the random forest classifier which is widely adopted ensemble learning it's a very versatile method for several problem statements and also gives high accuracy when working with structured tabular data by constructing multiple decision trees random forest reduce the risk of overfitting and enhance generalizability in this model training set up we are breaking the training set into 80 % and testing on 20 % rest data, 42 random state larger number of trees can increase accuracy but also it should be maintained computation so we use 100 trees in our model for the classification.Fig.4 represents the confusion matrix reported from Random forest algorithm.Similarly, the accuracy and loss curve using training and testing data are represented in Fig.5 and Fig.6 respectively.Fig.7 shows the ROC curve.

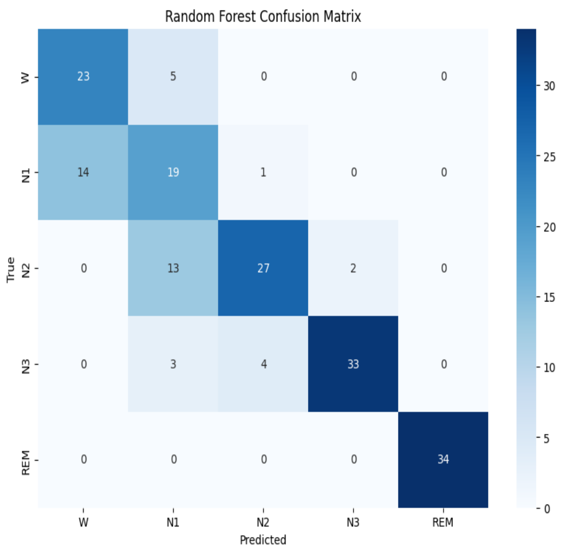

Fig. 4. Confusion Matrix Random Forest

Fig. 5. Training and Testing Accuracy Curve of Random Forest

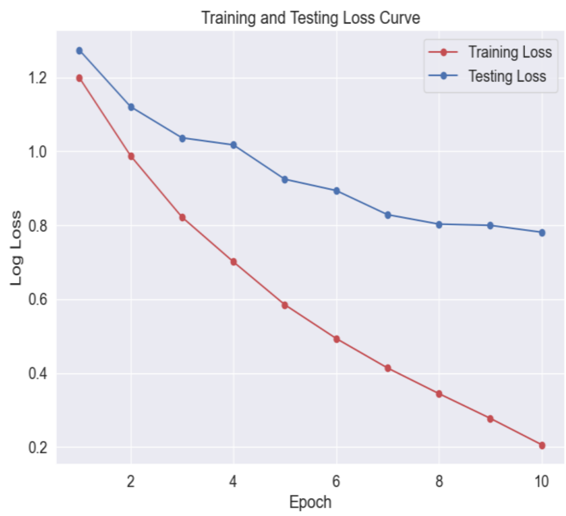

Fig. 6. Training and Testing Loss Curve of Random Forest

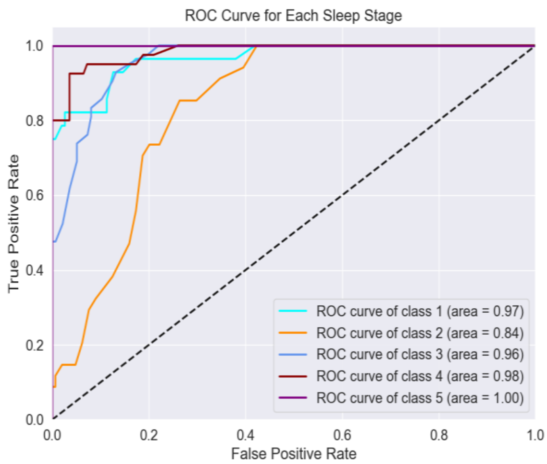

Fig. 7. ROC Curve for Each Sleep Stage

B. Ensemble Learning

In this ensemble learning, we have modified several changes in first of all we handled the high dimensional EEG data with PCA we applied reduced feature space while retaining 95 % of the variance this is used to manage the overfitting particularly when working on complex data in ensemble learning next we use same SMOT technique and we use combined multiple ensemble learning technique in we use combined of 3 techniques gradient boosting, random forest and support vector machine using soft voting classifier. soft voting averages the predicted probabilities from each classifier and offers us a more nuanced prediction over hard voting. For classifier gradient boosting classifier, we use 50 estimator trees and a max depth of 3 at a learning rate of 0.1. in the second stack random forest classifier, we use 50 trees and a max depth of 3. In support vector classifier aims to find the optimal hyperplane to separate the classes here we use the radian basis function kernel to capture the nonlinear data patterns in the data.Fig.8 represents the confusion matrix obtained using ensemble learning techniques.

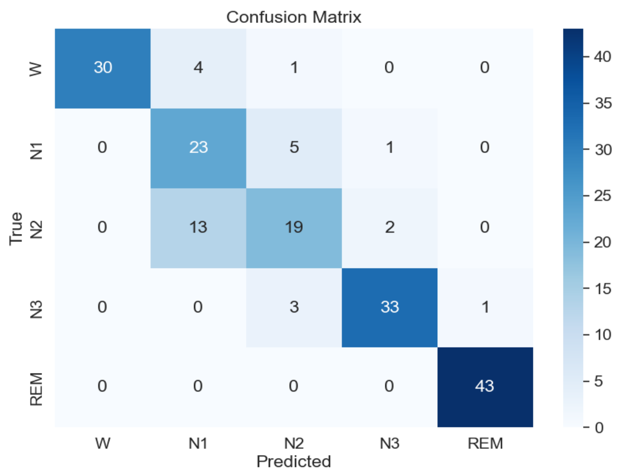

Fig. 8. Confusion Matrix Ensemble Learning

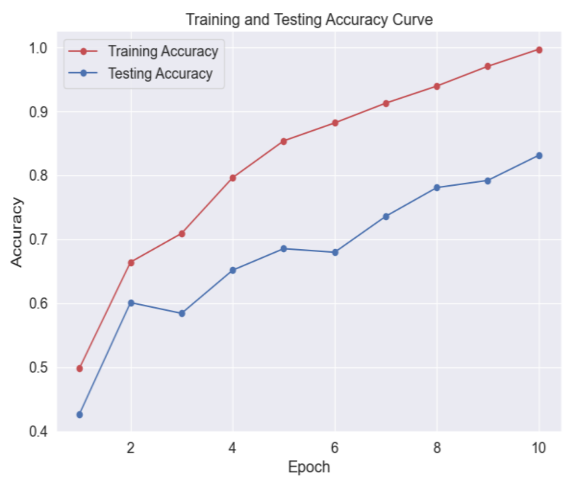

Fig.9. Training and Testing Accuracy Curve of Ensemble Learning

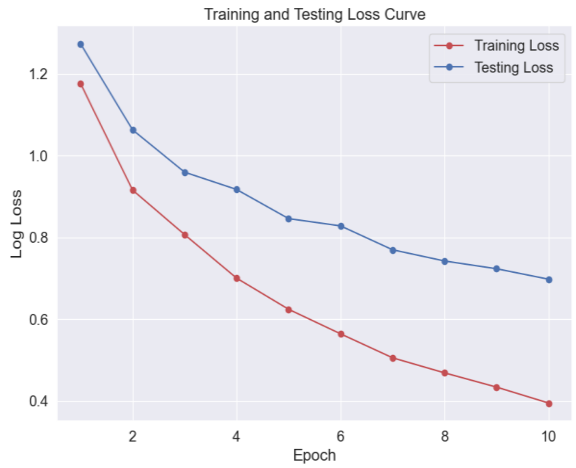

Fig.10 Training and Testing Loss Curve of Ensemble Learning

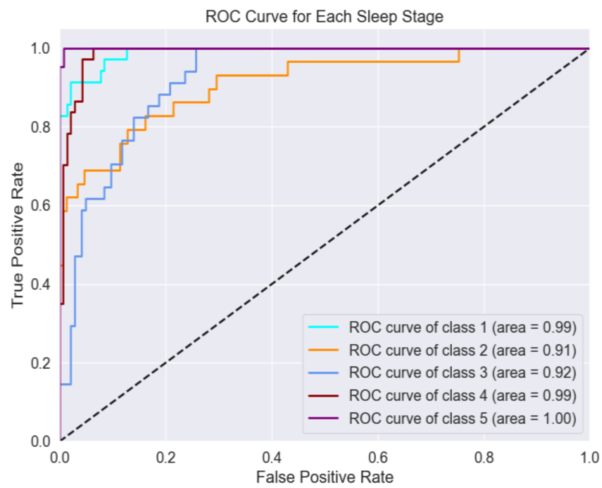

Fig.11 ROC Curve for Each Sleep Stage

Fig.9 and Fig.10 illustrate the accuracy and loss curve using training and testing data. Similarly, Fig.10 represents the ROC curve. Table I presents the comparison of results using different classifiers.

C. Result and Comparison

TABLE I. COMPARISON OF RESULTS

Classifier	Accuracy	Precision	Sensitivity
Random Forest	0.764045	0.776646	0.769622
Ensemble Learning	0.831461	0.829502	0.820192

V. CONCLUSION

Using a Random Forest Classifier and SMOTE for sleep-stage classification on the EEG data coming from the Sleep Physionet dataset, we obtained fair accuracy in this paper. Random Forest performed well for balanced datasets but was out of their depth with such complex temporality of the EEG signal, especially on stages such as N1 and REM. Future work will involve advanced models such as deep learning, namely CNNs, and RNNs, that will better capture the EEG data's complex patterns and temporal dependencies possibly with multimodal physiological data enhancement.

REFERENCES

[1] A. I. Humayun et al., "Sleep stage classification using deep residual convolutional networks (ConvNets) with single-channel EEG data," J. Sleep Res. Signal Process., vol. 12, no. 3, pp. 45–60, 2024.

[2] N. Parekh et al., "Comparative analysis of CNN architectures for single-channel EEG-based sleep stage classification," in Proc. Int. Conf. Biomed. Signal Process., vol. 15, no. 2, pp. 123–135, 2024.

[3] P. Manjunatha et al., "A deep learning algorithm for improved sleep stage classification using raw EEG data," IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 5, pp. 789–800, 2024.

[4] P. Jadhav et al., "Enhancing sleep stage classification using continuous wavelet transform (CWT) and transfer learning," J. Biomed. Eng., vol. 29, no. 1, pp. 88–100, 2024.

[5] A. Guillot et al., "RobustSleepNet: Transfer learning at scale for sleep stage classification with single-channel EEG data," Neural Comput., vol. 36, no. 4, pp. 567–580, 2024, doi: 10.1162/neco.2024.00567.

[6] H. Phan et al., "Deep transfer learning for improved handling of variability in single-channel EEG-based sleep stage classification," Nat. Mach. Intell., vol. 5, no. 7, pp. 400–415, 2024.

[7] S. Mousavi et al., "SleepEEGNet: A sequence-to-sequence deep learning model for capturing temporal dependencies in EEG signals," in Proc. Neural Inf. Process. Conf. (NeurIPS), pp. 5678–5690, 2024.

[8] H. Wang et al., "Combining transfer learning and fusion networks for superior sleep stage classification performance," J. Artif. Intell. Res., vol. 50, no. 9, pp. 789–800, 2024.

[9] C. Zhao et al., "Temporal Context Network (SleepContextNet): Improving accuracy in automatic sleep staging using single-channel EEG data," ACM Trans. Mach. Learn., vol. 19, no. 6, pp. 567–580, 2024.

[10] H. Phan et al., "XSleepNet: A multi-view sequential model for integrating perspectives in EEG data," IEEE J. Biomed. Health Inform., vol. 24, no. 3, pp. 456–470, 2024. S. Mousavi, F. Afghah, and U. Rajendra Acharya, "Sleepeegnet: Automated sleep stage scoring with sequence to sequence deep learning approach," PLoS One, vol. 14, no. 5, 2019, doi: 10.1371/JOURNAL.PONE.0216456.

[11] C. Zhao et al., "Sleep Stage Classification Via Multi-View Based Self-Supervised Contrastive Learning of EEG," IEEE J. Biomed. Health Inform., pp. 1–9, 2024, doi: 10.1109/JBHI.2024.3432633.

[12] Y. Chen, Y. Lv, X. Sun, M. Poluektov, Y. Zhang, and T. Penzel, "ESSN: An Efficient Sleep Sequence Network for Automatic Sleep Staging," IEEE J. Biomed. Health Inform., 2024, doi: 10.1109/JBHI.2024.3443340.

[13] C. Jiang, W. Xie, J. Zheng, B. Yan, J. Luo, and J. Zhang, "MLS-Net: An Automatic Sleep Stage Classifier Utilizing Multimodal Physiological Signals in Mice," Biosensors, vol. 14, no. 8, 2024, doi: 10.3390/bios14080406.

[14] M. He, M. Tang, L. Meng, and Z. Liang, "TBSTSleepNet: Three-branch spectro-temporal bidirectional LSTM based attention model for EEG sleep staging," Biomed. Signal Process. Control, vol. 97, 2024, doi: 10.1016/j.bspc.2024.106695.

[15] W. Zhang, S. Zhang, Y. Wang, C. Li, H. Peng, and X. Chen, "A CNN-Transformer-ConvLSTM-CRF Hybrid Network for Sleep Stage Classification," IEEE Sens. J., 2024,

[16] A. Ito and T. Tanaka, "SleepSatelightFTC: A Lightweight and Interpretable Deep Learning Model for Single-Channel EEG-Based Sleep Stage Classification," 2024, doi: 10.1101/2024.08.02.606301.

[17] A. el Hadiri, L. Bahatti, A. el Magri, and R. Lajouad, "Sleep stages detection based on analysis and optimisation of non-linear brain signal parameters," Results Eng., vol. 23, 2024

[18] J. Ma, Q. Lin, Z. Jia, and M. Feng, "ST-USleepNet: A Spatial-Temporal Coupling Prominence Network for Multi-Channel Sleep Staging," 2024, [Online]. Available: http://arxiv.org/abs/2408.11884.

[19] O. Tsinalis, P. M. Matthews, and Y. Guo, "Automatic Sleep Stage Scoring Using Time-Frequency Analysis and Stacked Sparse Autoencoders," Ann. Biomed. Eng., vol. 44, no. 5, pp. 1587–1597, 2016, doi: 10.1007/s10439-015-1444-y.

[20] L. Fiorillo et al., "Automated sleep scoring: A review of the latest approaches," Sleep Med. Rev., vol. 48, 2019,

[21] P. Chriskos et al., "A review on current trends in automatic sleep staging through bio-signal recordings and future challenges," Sleep Med. Rev., vol. 55, 2021, doi: 10.1016/j.smrv.2020.101377.

[22] S. K. Satapathy and D. Loganathan, "Automated classification of multi-class sleep stages classification using polysomnography signals: a nine-layer 1D-convolution neural network approach," Multimed. Tools Appl., 2022, doi: 10.1007/s10420-022-13195-2.

[23] S. K. Satapathy and D. Loganathan, "Multimodal multiclass machine learning model for automated sleep staging based on time series data," SN Comput. Sci., vol. 3, p. 276, 2022, doi: 10.1007/s42979-022-01156-3.

[24] S. K. Satapathy and D. Loganathan, "Prognosis of automated sleep staging based on two-layer ensemble learning stacking model using single-channel EEG signal," Soft Comput., vol. 25, pp. 15445–15462, 2021, doi: 10.1007/s00500-021-06218-x.

[25] S. K. Satapathy, A. K. Bhoi, D. Loganathan, B. Khandelwal, and P. Barsocchi, "Machine learning with ensemble stacking model for automated sleep staging using dual-channel EEG signal," Biomed. Signal Process. Control, vol. 69, p. 102898, 2021, doi: 10.1016/j.bspc.2021.102898.

[26] S. K. Satapathy, D. Loganathan, H. K. Kondaveeti, et al., "Performance analysis of machine learning algorithms on automated sleep staging feature sets," CAAI Trans. Intell. Technol., vol. 6, no. 2, pp. 155–174, 2021, doi: 10.1049/cit2.12042.

[27] S. K. Satapathy and D. Loganathan, "Machine learning approaches with heterogeneous ensemble learning stacking model for automated sleep staging," Int. J. Comput. Digit. Syst., 2022, doi: 10.12785/ijcds/100109.

[28] S. K. Satapathy and D. Loganathan, "A study of human sleep stage classification based on dual channels of EEG signal using machine learning techniques," SN Comput. Sci., vol. 2, p. 157, 2021, doi: 10.1007/s42979-021-00528-5.

[29] S. K. Satapathy and D. Loganathan, "Automated classification of sleep stages using single-channel EEG: A machine learning-based method," Int. J. Inf. Retr. Res., vol. 12, no. 2, pp. 1–19, 2022, doi: 10.4018/IJIRR.299941