Improving Efficiency by using Dilated Causal Convolutions to extract Hierarchical Temporal Pattern of Time Series ECG Signals for identifying Sleep Apnea Syndrome

BhanuSree.Y
Assistant Professor
SVKM's NMIMS University
Hyderabad, India
Bhanusree.Y@nmims.edu

Thrisha Reddy.J
Department of CS-DS
SVKM's NMIMS University
Hyderabad, India
thrisha.reddy026@nmims.edu.in

Snehil.A
Department of CS-DS
SVKM's NMIMS University
Hyderabad, India
snehil.a030@nmims.edu.in

Vijaya Laxmi.S
Department of CS-DS
SVKM's NMIMS University
Hyderabad, India
svijaya.laxmi044@nmims.edu.in

Abstract:
A prevalent but underdiagnosed condition linked to serious cardiovascular risks is sleep apnea. Automated ECG-based detection is necessary because traditional polysomnography is expensive and time-consuming. This study developed and evaluated several models, including Random Forest (RF), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), Squeeze-and-Excitation Multi-Scale CNN with Transformer (SE-MSCNN+Transformer), and Temporal Convolutional Network with Transformer (TCN+Transformer), using single-lead ECG signals from the PhysioNet Apnea-ECG database. To extract pertinent features, preprocessing included temporal segmentation, R-peak detection, and FIR filtering. With an accuracy of 90.59%, precision of 88.9%, recall of 87.2%, and an AUC of 0.9665, the suggested TCN+Transformer outperformed other baseline models in the experiments. These results highlight the potential for integration into wearable and real-time diagnostic systems by demonstrating the efficacy of hybrid temporal convolution and attention mechanisms for reliable and effective sleep apnea detection.

Keywords: ECG Signal Analysis, Temporal Convolutional Network, Deep Learning, Multi-Scale Processing.

I. Introduction:

Sleep apnea syndrome, the most common and serious sleep disorder, means that breathing stops and starts during sleep in a cyclical way. This ends with causing persistent irregular sleep patterns and less oxygen being delivered to the human body. Given the high prevalence of sleep apnea and its well-established associations with cardiovascular morbidity, cognitive dysfunction, and decreased quality of life, early diagnosis is crucial [14]. Therefore, there is an important requirement for non-invasive diagnostic techniques [17].

However, previous machine learning efforts on physiological signals are promising yet suffer from low accuracy, scalability, and generalization across populations [19]. Notably, existing studies show a gap in comprehensive comparisons of classical and deep learning models, especially those integrating temporal convolutional networks and transformer architectures for ECG analysis.

To overcome these shortcomings, this study investigates the effectiveness of classical ensemble methods and Deep Learning algorithms, such as Random Forest (RF), temporal convolutional networks-transformers (TCN+Transformer), SEMSCNN-Transformers, Long Short-Term Memory networks (LSTM), and Bidirectional LSTM (Bi-LSTM) , then implemented these algorithms on PhysioNet ECG data to extract rich temporal dynamics and morphology-based features to predict sleep apnea. The key objectives include closely reviewing and contrasting model performance in recognizing sleep apnea from ECG signals and determining the most ideal architecture for advanced usage by the patients of that [19]. The methodology entails pre-processing ECG signals, extracting features where appropriate, training of the models, and evaluating the model performance through standard metrics.

By performing a balanced and reproducible comparison, it proves that transformer-augmented convolutional models, specifically the TCN+Transformer, have better accuracy, sensitivity, and specificity compared to recurrent and traditional baselines[13]. The findings identify the possibility of using precise, computationally light, and interpretable models in low-cost, handheld screening devices, thus facilitating early detection and enhanced clinical management of sleep apnea.

II. Literature Review:

A number of studies have attempted the use of ECG signals for the automatic detection of sleep apnea using machine learning and deep learning methods with varying levels of success. Anbalagan et al. applied Bi-LSTM and LSTM models on the PhysioNet challenge data and reported encouraging apnea detection performance, where Bi-LSTM performed better than traditional recurrent models [2]. Nevertheless, the lack of convolutional or attention mechanisms was a limitation to effective capture of complicated temporal dependencies [4].

Yue et al. introduced a squeeze-and-excitation multi-scale convolutional neural network (SE-MSCNN) for obstructive sleep apnea diagnosis based on ECG signals. The light model had high specificity but was tested on a fairly small data set, which doesn't allow it to be validated in different patient populations [6].

Liu et al. proposed an EfficientNet-based approach to OSA detection from single-lead ECG signals with an AUC of 91.7% [12]. Although effective, the research called for additional clinical validation prior to widespread implementation [11].

Despite such findings in ECG-based sleep apnea detection, most have constraints in terms of dataset size, incorporation of deep architectures, or efficiency in computation [18]. The current work is unique in that it uses a hybrid Temporal Convolutional Network with Transformer attention to reach high accuracy of 90.59% without compromising computational efficiency for real-time use in clinics on a large and heterogeneous publicly available dataset [15].

III. Proposed Method:

The methodology proposed entails automated sleep apnea event detection from single-lead ECG recordings using a wide variety of machine learning and deep learning models. The system begins with intensive signal preprocessing and feature extraction, employs multi-scale temporal representation, and then passes these representations into different classification models for comparison in a robust manner [7].

A. Dataset

This work exclusively utilized the Electrocardiogram data(ECG) from the Apnea-ECG Database, meticulously curated by George Moody and Roger Mark and released on February 10, 2000. The dataset contained 70 records, of which 35 were for the learning set labelled a01 through A20, b01 through B05, and c01 through C10, and an additional 35 were allocated to the test set labelled x01 through X35). The recordings varied from slightly less than 7 h to nearly 10 h. Continuous digital ECG signals and apnea annotations generated by human experts based on the simultaneous recording of respiration and associated signals are included in every file. Additionally, machine-generated QRS annotations are present, where all beats, regardless of their type, are uniformly labeled as normal. While eight recordings included supplementary signals such as chest and abdominal respiratory effort signals (Resp C and Resp A), oronasal airflow (Resp N), and oxygen saturation (SpO2), our focus was solely on extracting and utilizing the digitized ECG data [9]. The dataset encompasses multiple files associated with each recording, with files such as rnn.dat containing digitized ECGs. hea files specifying the names and formats of the associated signal files. apn files serving as binary annotation files indicating the presence or absence of apnea per minute (available exclusively for the learning set), and qrs files representing machine-generated annotation files for QRS detection.

B. Data Preprocessing

In the process of data preparation , the ECG signals were used which went through a rigorous preprocessing pipeline to enhance signal quality,while preserving physiological relevance and prepare multi-scale representations of features for model building Single-Lead ECG at 100Hz was bandpass filtered with a FIR filter and cut-off frequencies of 3-45Hz .This filtering strategy well reduced baseline wander and high-frequency artifacts , such as muscle activity and powerline interference , without removing the morphological integrity of QRS complexes critical for accurate cardiac event detection [10].

The R-peaks were identified by Biosppy's application of the Hamilton segmentation algorithm, which is selected for its stability over a wide range of ECG morphologies and noise levels-and then further cleaned with the correct-rpeaks function employing a ±0.1S window of tolerance. Temporal adjustment to this level protected against peak times yielding high-fidelity RR-interval time series necessary for Heart rate variability calculation and subsequent feature extraction [7].

After R-peak detection and enrichment, the preprocessed ECG signals were broken down into overlapping five-minute windows with two minutes of prior context, one minute as the middle target interval, and two minutes of posterior context. To enrich the training data set and learn apnea occurrences in varied temporal contexts, a sliding-window strategy with a step of 15 seconds was utilized. Every segment went through physiological plausibility screening, which excluded any having mean heart rates outside the 20–300 bpm range, less than 40 or greater than 200 beats per window, or overabundant artifacts [10].

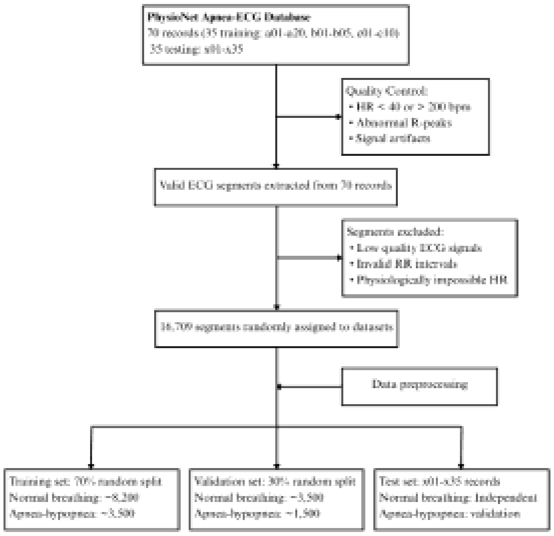

Figure 1. Data preprocessing and partitioning workflow for the PhysioNet Apnea-ECG dataset.

The remaining eligible segments were resampled through cubic-spline interpolation to generate multi-scale temporal representations at three scales: full scale (900 samples) over the entire five-minute window to time autonomic trends over the long-term; medium scale (540 samples) over the central three-minute segment for event-related cardiac responses; and fine scale (180 samples) over the central one-minute segment for high-amplitude cardiac fluctuations [6].This multi-scale strategy allowed the models to learn in parallel gradual and fleeting cardiac responses typical of sleep apnea.

Following preprocessing, 41,162 segments were left in the learning set, which were divided into 70% training and 30% validation subsets with even apnea–normal class distribution. Validation data were kept entirely unseen throughout training to allow for an objective test of model generalization [9].

C. Model Selection & Architecture:

After ECG data preprocessing and multi-scale feature extraction, then compared five different model architectures combining both classical and deep learning methodologies for sleep Apnea screening [3].

The study aimed to find the comparative effectiveness of these models to capture the temporal and morphological dynamics present in ECG signals during apnea event. Initially employed a Random Forest classifier, which uses an ensemble of decision trees learned from multi-scale handcrafted features extracted from temporal and amplitude variations in the ECG. The traditional approach is very good at extracting difficult nonlinear patterns in the feature spaces while ensuring in models [18].

The second and third assessed architectures were sequential models based on—Long Short-Term Memory (LSTM) and Bidirectional LSTM (Bi-LSTM) networks, Utilizing keras to analyze the preprocessed multi-scale ECG segments. The LSTM model modelled temporal dependencies unidirectionally, learning past to present patterns within each segment. Conversely, the Bi-LSTM network read the sequence in both the forward and backward directions and thus captured the full temporal context around apnea events [2].In both architectures, recurrent layers were followed by FC layers with Rectified Linear Unit (ReLU) activation functions for the purpose of introducing nonlinearity as well as the dropout regularization to prevent overfitting.

The fourth model was the Squeeze-and-Excitation Multi-Scale CNN with a Transformer (SEMSCCNN+Transformer). This hybrid model combines several convolution streams learning ECG signals at very temporal scales, strengthened by squeeze-and-excitation blocks that adaptively rescale channel-wise feature responses to highlight the most representative signals. After feature extraction, transformer encoder layers capture long term temporal dependencies along the ECG segment [4].

The last model, Temporal Convolutional Network with Transformer (TCN+Transformer) utilized dilated causal convolutions to efficiently extract hierarchical temporal pattern over large receptive fields, of the ECG time series[15].

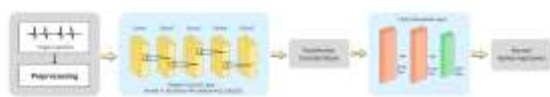

Figure 2. Architecture of the TCN+Transformer model.

The TCN+Transformer model combines stacked one-dimensional convolutional layers with increasingly larger dilation rates (e.g., 1, 2, 4, 8, and 16), which allows effective extraction of both short- and long-term temporal features from ECG time series. The resulting hierarchical feature representations are further processed by transformer encoder blocks with multi-head self-attention, enabling the model to detect global dependencies and contextual patterns over the entire input segment[5], This hybrid architecture is well-suited to leverage the advantages of temporal convolutions and attention mechanisms and thus to enhance the detection of apnea events on both localized waveform features and global temporal patterns.

All models were trained over class- balanced datasets with early stopping for overfitting and hyperparameters optimized through the performance of the validation set. Iterative experimentation was employed to find optimal learning rates, dropout rates, and specific architectural parameters. Iterative optimization aimed not just for increase in for accuracy, precision, recall, and F1-score but also for a balance between predictive performance, computational efficiently, and interpretability essential necessities for eventual clinical deployment of ECG-based sleep apnea systems [18].

D. Training and Evaluation Protocol

The training was carried out on preprocessed multi-scale electrocardiogram database with data stratified to preserve the native class ratio between apnea and normal. Deep learning models were trained using the Adam optimizer with categorical cross-entropy loss, employing early stopping based on the loss of validation to avoid overfitting. Batch sizes were chosen to maximize GPU efficiency. A held-out validation set of training data was employed for hyperparameter optimization, including learning rate (originally fixed at 0.001) optimization, dropout rate (between 0.2 and 0.5), and layer dimensions [16]. The data was split at the subject level into training and test sets with the ratio 80:20, with the training set also being divided in the ratio of 80:20 for training and validation to facilitate proper model development and assessment. Traditional models like Random Forest were trained with 5-fold cross-validation with 100 estimators for good parameter selection and accurate performance estimation. Computational metrics such as average training time per epoch and inference latency per segment were quantified using an NVIDIA GPU (e.g., RTX 3060) to assess the models' appropriateness for real-time clinical use.

E. Evaluation Metrics:

The model performance was measured by accuracy, precision, recall, F1-score and the Area Under ROC Curve (AUC-ROC).Confusion matrices were also employed to study the true and false classification ,distribution and common patterns of errors between classes [10].Computational metrics of training time and inference latency were also captured to gauge the viability of deploying the models in real-time (or) resource-limited clinical settings.

Upon validation, the models were externally tested on a new dataset of unseen subjects to evaluate their capacity for generalizing to novel patient ECG profiles. This external testing phase also helped ensure real-world applicability through the maintenance of high predictive accuracy while satisfying practical limitations like low inference latency and moderate computational demand — both being imperative for automated ECG-based sleep apnea screening on devices used clinically and portably [6].

Iv. Results and Comparative Analysis

The five models were compared— random forest (RF), long short-term memory (LSTM), bi-directional LSTM (Bi-LSTM), SEMSCNN+Transformer, and temporal convolution network (TCN)+ transformer.

Table III provides the classification results of each architecture on various metrics, such as accuracy, precision, F1-score, and AUC-ROC[1].

Table 2. Performance evaluation of sleep apnea detection approaches of the proposed architecture.

Model	Accuracy	Precision	F1-Score	AUC-ROC
Random Forest	0.781	0.691	0.687	0.833
LSTM	0.849	0.812	0.819	0.923
Bi-LSTM	0.867	0.825	0.837	0.938
SEMSCNN+ Transformer	0.902	0.885	0.876	0.962
TCN+ Transformer	0.905	0.889	0.877	0.965

The hybrid deep learning architectures with attention models such as SEMSCNN+Transformer and TCN+Transformer models achieved higher accuracy and AUC than the classical and recurrent models, demonstrating that combining multi-scale convolutional or temporal convolutions with transformer-based attention improves ECG signal Apnea detection [4].

The confusion matrices for the TCN+Transformer and SEMSCNN+Transformer are provided in Table V[15].

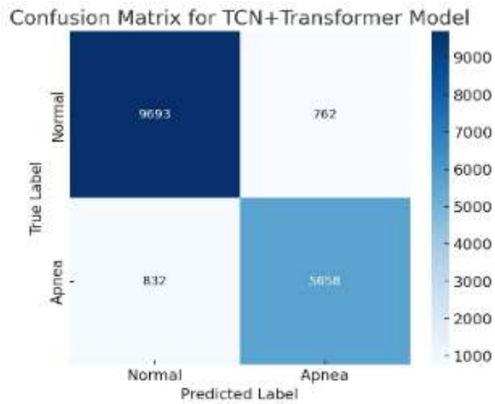

Figure 3. Confusion matrix for the TCN+Transformer model.

In the case of the TCN+Transformer model, the test set contained a total of 12,285 normal events, with 9,693 correctly labeled as normal (true negatives) and 762 wrongly identified as apnea (false positives). For apnea events, of the 6,490 actual occurrences, 5,658 were correctly classified (true positives), while 832 were missed (false negatives) [5].

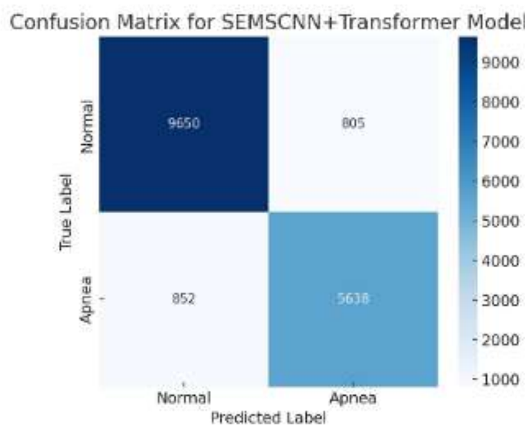

Figure 4. Confusion matrix for the SEMSCNN+Transformer model.

In the same vein, the SEMSCNN+Transformer accurately classified 9,650 normals and 5,638 apnea events. Its corresponding false positives and false negatives were 805 and 852, respectively.[3].

A number of misclassified apnea events are probably from mild apnea that triggers minor ECG changes and are hard to identify against normal respiration. False positives among the normals, on the other hand, can come from noisy or of low quality ECG signals impacting feature extraction[7].

According to these findings, TCN+Transformer reported an accuracy of 90.59%, precision of 88.9%,F1-score of 87.7%, and an AUC-ROC of 0.9665 (as indicated in Table VI). In relation to the SEMSCNN+Transformer model, TCN+Transformer exhibited greater sensitivity and specificity, indicating its application efficacy for precise and trustworthy ECG-based sleep apnea screening in real-world environments [8].

V.Conclusion:

This study conducted a thorough comparison of five models— random forest (RF), long short-term memory (LSTM), bi-directional LSTM (Bi-LSTM), automatically detecting sleep apnea from single-lead ECG signals. The work contrasted a traditional ensemble method with sequential, convolutional, and attention-based deep learning designs, using a structured process that included multi-scale ECG preprocessing, balanced dataset splitting with cross-validation, and strict evaluation on previously unseen subjects.In our experiments, we showed that hybrid models which combine hierarchical temporal feature extraction with Transformer attention performed better over more classical approaches. Of these, the TCN+Transformer achieved the highest accuracy, sensitivity, and specificity with efficient inference times and low misclassification rates, as it modeled short- and long-range temporal dependencies through dilated causal convolutions and multi-head self-attention for clinical real-time screening. In summary, our results validate the proposed architecture as a reliable, scalable, and interpretable model for ECG-based apnea detection, and further research will address the generalization of this classifier to larger and more diverse patient cohorts necessary for broad clinical deployment.

Acknowledgments

References

[1] Anbalagan, Thivya, et al. "A novel deep learning approach for detection of sleep apnea from ECG signal." 2023 4th International Conference on Signal Processing and Communication (ICSPC). IEEE, 2023.

[2] Biswas, Pranta, and Mohammad Abu Yousuf. "Leveraging Transformer Models for Accurate Detection of Obstructive Sleep Apnea from Single-Lead ECG Signals." Proceedings of the 3rd International Conference on Computing Advancements. 2024.

[3] Devashania, Ielin Daisy, et al. "Temporal Fusion Transformer Based Temporal Convolutional Network Model for Sleep Apnea Detection Using Single-Lead ECG Signals." 2024 9th International Conference on Communication and Electronics.

[4] Elsheikh, Asmaa Ali, Doaa Ezzat, and Manal Tantawi. "Efficient Sleep Apnea Detection via ECG Signals Using a 1D Convolutional Neural Network." International Conference on Advanced Intelligent Systems and Informatics. Cham: Springer Nature Switzerland, 2025.

[5] Fahruzi, Iman, Vicky Ardiansyah, and Muhammad Habib Hibrizi. "Develop a Deep Learning Framework of 1D-CNN to Practical Application of Sleep Assessment Method." 2024 4th International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS). IEEE, 2024.

[6] Gladence, L. Mary, et al. "Obstructive Sleep Apnea Detection Using Machine Learning." 2024 4th International Conference on Advancement in Electronics & Communication Engineering (AECE). IEEE, 2024.

[7] Kolhar, Manjur, Manahil Muhammad Alfridan, and Rayan A. Siraj. "AI-Driven Detection of Obstructive Sleep Apnea Using Dual-Branch CNN and Machine Learning Models." Biomedicines 13.5 (2025): 1090.

[8] Kumar, Vinodh, et al. "Enhanced Sleep Apnea Diagnosis Using Hybrid CNN-LSTM Framework and Machine Learning Classifier." 2024 International Conference on Emerging Research in Computational Science (ICERCS). IEEE, 2024.

[9] Lee, Cheol-Hui, et al. "Toward Foundational Model for Sleep Analysis Using a Multimodal Hybrid Self-Supervised Learning Framework." arXiv preprint arXiv:2502.17481 (2025).

[10] Liu, Meng-Hsuan, et al. "EfficientNet-based machine learning architecture for sleep apnea identification in clinical single-lead ECG signal data sets." BioMedical Engineering OnLine 23.1 (2024): 57.

[11] Pan, Hui, et al. "MobileNetV2: A lightweight classification model for home-based sleep apnea screening." arXiv preprint arXiv:2412.19967 (2024).

[12] Pham, Duc Thien, and Roman Mouček. "Efficient sleep apnea detection using single-lead ECG: A CNN-Transformer-LSTM approach." Computers in Biology and Medicine 196 (2025): 110655.

[13] Prajitha, M. V., G. Naveen Sundar, and D. Narmadha. "Detection of Sleep Apnea: A Comparative Analysis of Advanced Machine Learning and Deep Learning Approaches." 2024 11th International Conference on Advances in Computing and Communications (ICACC). IEEE, 2024.

[14] Q. Zhang, J. He, and J. Li. "Transformer-Based Multimodal Framework for Accurate Sleep Apnea Detection Using ECG and SpO2 Signals." 2024 3rd International Conference on Health Big Data and Intelligent Healthcare (ICHIH), Zhuhai, China, 2024, pp. 361-365.

[15] Rolant Gini, J., and K. Dhanalakshmi. "Apnea detection using ECG signal based on machine learning classifiers and its performances." Journal of Medical Engineering & Technology 47.7 (2023): 344-354.

[16] Scarpetta, Marco, et al. "A Tiny Deep Learning Model for Sleep Apnea Detection Based on ECG Signals." 2025 IEEE Medical Measurements & Applications (MeMeA). IEEE, 2025.

[17] Vani, Koneru Suvarna, et al. "Sleep Apnea Prediction Using Real Time ECG Data." 2024 Intelligent Systems and Machine Learning Conference (ISML). IEEE, 2024.

[18] Yamane, Takahiro, Masanori Fujii, and Mizuki Morita. "Clinical-level screening of sleep apnea syndrome with single-lead ECG alone is achievable using machine learning with appropriate time windows." Sleep and Breathing 29.2 (2025): 156.