Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal **Vol. 7, No. 1, January 2025, pp: 105-116; eISSN: 2656-8632**

RESEARCH ARTICLE **OPEN ACCESS**

Manuscript received September 10, 2024; revised October 6, 2024; accepted October 12, 2024; date of publication November 20, 2024
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i1.573

How to cite: Fadil Pratama, Wiharto, Umi Salamah, "Sleep Apnea Detection Model Using Time Window and One-Dimensional Convolutional Neural Network on Single-Lead Electrocardiogram", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 7, no. 1, pp. 105-116, January 2025.

Sleep Apnea Detection Model Using Time Window and One-Dimensional Convolutional Neural Network on Single-Lead Electrocardiogram

Fadil Pratama, Wiharto, and Umi Salamah

Department of Informatics, Faculty of Information Technology and Data Science, Universitas Sebelas Maret, Indonesia

Corresponding author: Wiharto (e-mail: wiharto@staff.uns.ac.id)

ABSTRACT Sleep apnea is an important disorder that involves frequent disruptions in breathing during sleep, which can result in numerous serious health issues, such as cognitive deterioration, cardiovascular illness, and heightened mortality risk. This study introduces a detailed model designed for the detection of sleep apnea using single-lead electrocardiogram signals, providing an accurate detection method. We can use single-lead ECG signals to get ECG-Derived Respiration (EDR). EDR combines important respiratory signals with RR intervals to help find sleep apnea more accurately. We structure the research process into seven systematic stages, ensuring a comprehensive approach to the issue. The process commences with the acquisition of data from the "Apnea-ECG Database" accessible on the PhysioNet platform, which underpins the ensuing analysis. Subsequent to data collection, we execute a sequence of preprocessing procedures, including segmentation, filtering, and R-peak detection, to enhance the ECG data for analysis. After that, we do feature extraction, which gives us 12 unique features from the RR interval and 6 features from the R-peak amplitude, which are both necessary for the model to work. The research subsequently utilizes feature engineering, implementing a Time Window methodology to encapsulate the temporal dynamics of the data. To ensure the results are robust, we conduct model evaluation using stratified K-fold cross-validation with five folds. The modeling technique employs a 1D Convolutional Neural Network (1D-CNN) utilizing the Adam optimizer. Ultimately, the performance assessment shows an accuracy score reaching 89.87%, sensitivity at 86.16%, specificity at 92.30%, and an AUC score of 0.96, attained with a Time Window size of 15. This model signifies a substantial improvement in performance relative to previous studies and serves as a feasible option for the detection of sleep apnea.

INDEX TERMS Electrocardiogram, One-Dimensional Convolutional Neural Network, Sleep Apnea, Stratified K-Fold Cross-Validation, Time Window

I. INTRODUCTION

Sleep apnea is a prevalent sleep disorder marked by repeated interruptions in breathing disturbances lasting over 10 seconds during sleep. This illness may result in numerous adverse impact on a person's health and quality of life [1]. If inadequately managed, sleep apnea may result in severe problems, including cognitive impairment [2], cardiovascular disorders [3], and potentially fatal outcomes. Approximately 14% of males and 5% of women in the United States are estimated to suffer from sleep apnea, with prevalence growing throughout diverse groups globally [4]. Given the significant increase in sleep apnea cases, it is critical to identify people at risk and implement prompt and effective treatments for monitoring, managing, and treating this condition.

Polysomnography (PSG) is a prevalent technique for identifying sleep apnea. This approach entails the examination of several biological data, including electrooculograms (EOG), electroencephalograms (EEG), electromyograms (EMG), pulse oximetry, and electrocardiograms (ECG). This technique necessitates the patient to spend the night in a sleep clinic, where an expert will affix numerous electrodes and cables and oversee the entire testing process [5]. These constraints may impede its utilization in clinical environments, underscoring the necessity to develop other

methods that are more pleasant and practical for sleep apnea detection.

The utilization of an electrocardiogram (ECG) is an effective method for identifying sleep apnea, as it offers critical insights into the cardiorespiratory system [6]. The ECG signal extracts the ECG-Derived Respiration (EDR) metric, which reflects the respiratory pattern. Heart Rate Variability (HRV) refers to the variations in the time intervals between heartbeats, determined by isolating the R peaks from the QRS complex within the ECG signal [7]. Single-lead ECG recordings can be employed to extract EDR, incorporate respiratory signals, and assess RR intervals, which aid in the detection of sleep apnea [8]. As a result, single-lead ECG data has emerged as the primary focus in the detection of sleep apnea, given its efficacy and accessibility.

In recent years, researchers have introduced many techniques for identifying sleep apnea utilizing single-lead ECG readings. The study by Rizal et al. [9] has established a classification algorithm for Obstructive Sleep Apnea (OSA) utilizing Support Vector Machine (SVM) techniques. The research findings demonstrate a segment accuracy of 89.5%. This work demonstrates effective results in sleep apnea identification by conventional machine learning techniques; nonetheless, this methodology has inherent limits in its operation. Conventional algorithms depend on a straightforward learning mechanism, wherein the model directly correlates processed input to output, lacking the capacity to comprehend intricate data patterns. This approach often struggles to detect non-linear relationships or interactions that might be hidden within the data. The advent of deep learning has significantly expanded possibilities with its far greater ability to capture complex and dynamic patterns that earlier machine learning methods could not handle.

A study by Wang et al. [10] has established a method for detecting sleep apnea utilizing variables derived from the amplitudes of R-peaks and RR intervals. This study employs a Multilayer Perceptron (MLP) integrated with a Time Window (TW) that varies in size from 0 to 15. In this study, the integration of TW effectively encapsulates the temporal dynamics of the ECG signal. This model attained a segment accuracy score reaching 87.3%, a sensitivity score reaching 85.1%, a specificity score reaching 88.7%, and an AUC (Area Under the Curve) score reaching 0.945. Nonetheless, while TW with an interval of 0–15 demonstrates optimal performance at size 5, the MLP model employed has not yet fully optimized the augmented number of retrieved features. Increased feature extraction often enhances performance; nevertheless, the MLP model has yet to fully assimilate the intricate patterns of the ECG data. This signifies a necessity for extensive TW testing and the implementation of more sophisticated models, such as One-Dimensional Convolutional Neural Networks (1D-CNN), which may excel in identifying sleep apnea and discerning intricate patterns from biological inputs.

Recent advancements in deep learning have shown favorable outcomes in numerous classification problems, particularly in picture classification [11], [12], signal analysis [13], [14], and its applications across diverse industries. Convolutional Neural Networks (CNN) are a deep learning methodology that emulates the hierarchical architecture of human visual processing [15]. CNN can analyze biological signals and conduct biometric signal analysis [16] and performing biometric signal analysis through the automatic extraction of salient aspects [17]. The deployment of CNN models, including 1D Convolutional Neural Networks (1D-CNN), may improve precision and dependability for detecting sleep apnea [18].

A study by Bahrami and Forouzanfar [19] created multiple models for sleep apnea detection through ECG data. This research employs machine learning methods, deep learning, and hybrid models. With an accuracy score reaching 88.13%, a sensitivity score reaching 84.26%, and a specificity score reaching 92.27%, the hybrid Model 1D-CNN ZFNet-BiLSTM did the best job of segmentation. To ensure result reliability, this research used a stratified K-fold cross-validation procedure with five folds, partitioning the data into numerous segments for model training and testing [20]. Although it delivers excellent results, the hybrid model is highly intricate due to the combination of several algorithms.

A study by Wang et al. [21] devised a sleep apnea detection algorithm utilizing features derived from the amplitudes of R-peaks and RR intervals. This study utilized a 1D-CNN based on the LeNet-5 architecture. It had an AUC score reaching 0.950, a segment accuracy score reaching 87.6%, a sensitivity score reaching 83.1%, a specificity score reaching 90.3%, and a sensitivity of 83.1%. Different deep learning models have been shown to be good at finding sleep apnea in the past, but none of them have looked into how Time Window (TW) can be used to improve the temporal context in ECG signal processing.

The implementation of the Time Window (TW) approach is a crucial technique in signal analysis employed to capture the temporal dynamics of data. This technique separates the ECG data into defined time intervals, enabling the model to examine variations in the signal within a systematic period. Employing TW enables the extraction of pertinent information from each time segment, yielding enhanced insights into the patterns of the ECG signal [10].

This research offers a sleep apnea detection model utilizing the 1D Convolutional Neural Network (1D-CNN) methodology alongside modifications of the Time Window (TW), based on numerous prior studies. This model tries to work as well as possible by using the best parts of 1D-CNN to find complicated ECG signal patterns and TW to deal with how the signal changes over time. The suggested model greatly enhances sleep apnea detection performance relative to prior models, especially for accuracy, sensitivity, and specificity, by the integration of 1D-CNN and Time Window modifications. The steps in this research method include getting ECG signal data, preprocessing the signals, and taking features from RR intervals and R-peak amplitude. We then implement the Time Window (TW) approach and evaluate the model using stratified K-fold cross-validation. The modelling uses One-Dimensional Convolutional Neural Networks (1D-

CNN) for classification purposes. Consequently, this research may serve as a more accurate prediction for the detection of sleep apnea.

II. MATERIALS AND METHODS

The research technique encompasses a comprehensive examination process consisting of seven sequential procedures: data collection, data preprocessing, feature extraction, feature engineering, modeling, model evaluation, and performance assessment. Data preparation encompasses activities, including segmentation, filtering, and R-peak identification. Feature extraction emphasizes the amplitudes of R-peaks and RR intervals. This research utilizes stratified K-fold cross-validation, including 5 folds. This system integrates TW with 1D-CNN and employs the Adam optimizer. Performance assessment evaluates the model's efficacy by examining metrics including accuracy, specificity, sensitivity, and the confusion matrix. FIGURE 1 delineates the successive phases of the research process.

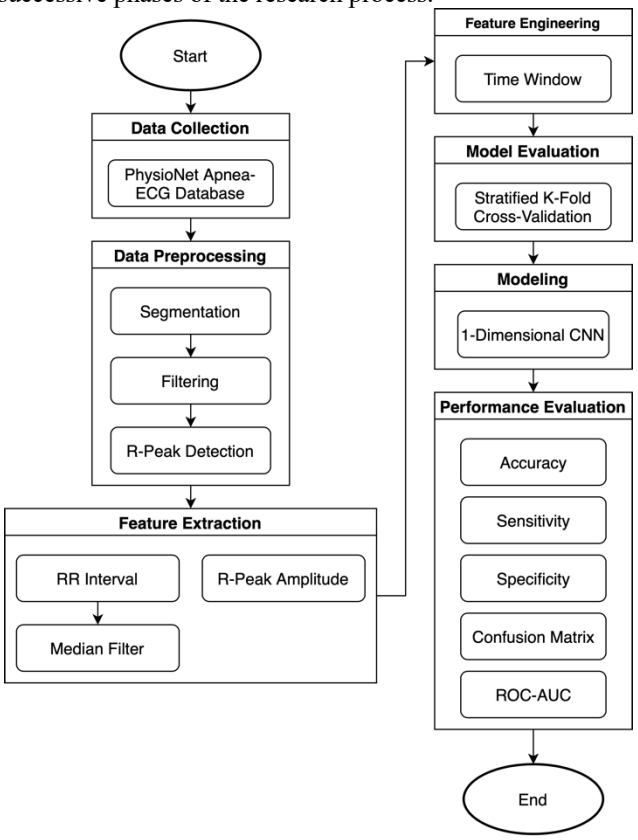

FIGURE 1. Research Process Flow Diagram

A. DATASET

This research utilized the "Apnea-ECG Database" [22], accessible via the PhysioNet platform [23]. This dataset comprises recordings from a heterogeneous group of 32 individuals, specifically 25 males and 7 women, aged between 27 and 63 years. This dataset lacks comprehensive information on the limitations or particular criteria for participant selection. Consequently, this study employs the dataset in its original form, without stratification or supplementary exclusion criteria. This dataset is a publicly accessible resource, provided users adhere to the terms and conditions of the specified license. Anyone who supplies the requisite reference can access the file. TABLE 1 presents the characteristics of the participants, encompassing diverse demographic information.

TABLE 1
Characteristics of the Apnea-ECG Dataset Participants

Characteristic	Value / Mean ± SD
Subjects (M:F)	32 (25:7)
Age (years)	48.0 ± 9.9
Height (cm)	175.4 ± 6.5
Weight (kg)	90.1 ± 17.6
Record Length (minutes)	493.5 ± 27.5
Non-apnea minutes	295.2 ± 126.1
Apnea minutes	198.4 ± 126.0
Hours with apnea	6.8 ± 2.7
AHI (Apnea-Hypopnea Index)	35.3 ± 23.9

The dataset comprises 70 records, divided into two categories: 35 records of training set and 35 records of a testing set. Each recording has a total duration of around 8 hours at a sampling rate of 100 Hz. We have labeled each recording with "A" for apnea and "N" for non-apnea at one-minute intervals. The training set comprises 17045 labeled segments, including 6514 segments identified as apnea and 10531 segments categorized as non-apnea. The test set consists of 17268 labeled segments, including 6550 classified as apnea and 10718 classified as non-apnea. Further details concerning the quantity of labeled segments for non-apnea and apnea are available in TABLE 2.

TABLE 2
Number of Label Distributions in Dataset

Dataset	Apnea	Non-Apnea	Total
Training set	6514	10531	17045
Test set	6550	10718	17268
Total	13064	21249	34313

B. DATA PREPROCESSING

Previous research has demonstrated that the R-peak amplitude and RR interval provide crucial insights into the occurrence of sleep apnea. Therefore, we employ a data preprocessing technique to extract characteristics from R-peak amplitude and RR interval. We divide the signal into 1-minute segments with a sampling rate of 100 Hz, as illustrated in FIGURE 2. We utilize a Finite Impulse Response (FIR) bandpass filter with an order of 30 and a passband of 3~45 Hz, effectively reducing noise. The input-output relationship of a FIR filter is determined by Eq. (1) [24].

$$y[n] = \sum_{k=0}^{N-1} h(k).x[n-k] \quad (1)$$

Here, $x[n-k]$ and $y[n]$ denote the filter input and output, $h(k)$ signifies the filter coefficients, and $N-1$ indicates the quantity of filter coefficients (filter order). As illustrated in FIGURE 3, this range guarantees the retention of pertinent ECG components while eliminating extraneous frequencies. FIGURE 4 illustrates the proper detection of R-peaks using the Hamilton method [25], which has a tolerance of 0.1 seconds.

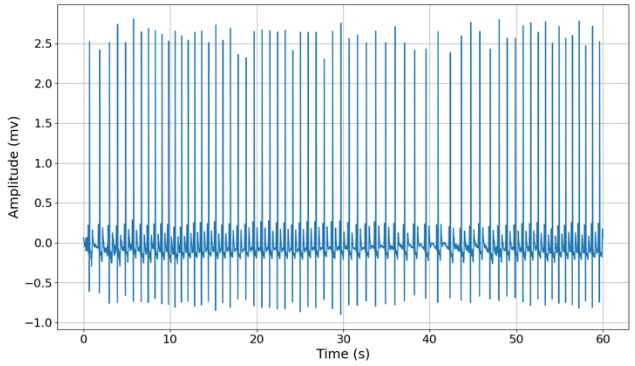

FIGURE 2. Segmentation Result

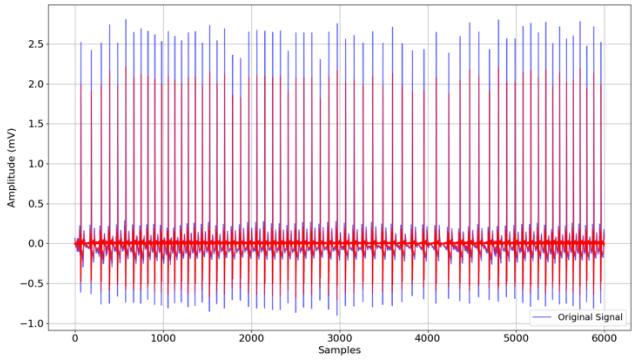

FIGURE 3. Filtering Result

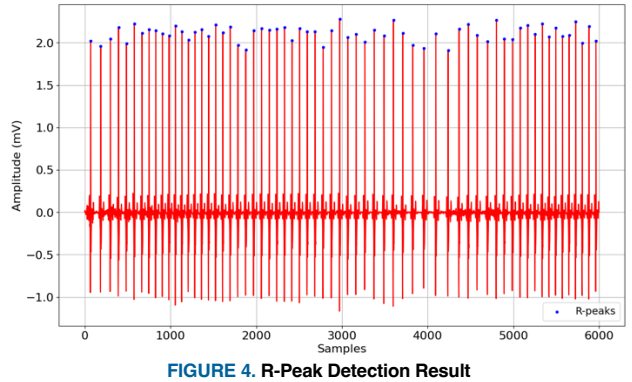

FIGURE 4. R-Peak Detection Result

C. FEATURE EXTRACTION

The RR interval represents the duration between two successive R-peaks [4]. The characteristics comprise six time domain attributes and six frequency domain attributes. Earlier studies have demonstrated that the amplitudes of R-peaks and RR intervals provide valuable information are important for diagnosing sleep apnea because they show how the heart reacts to breathing problems [10], [19], [21]. We applied a median filter, as described in [26], to the RR interval data to eliminate aberrant values that lacked physiological interpretation. This study employs a strategy that substitutes each data value with the median of an ordered group of nearby values. The set's kernel size is three. This computation considers three sequential values (i.e., the preceding value, the current value, and the subsequent value) for each data point, thereafter determining the median of these three values. TABLE 3 illustrates the amalgamation of 12 variables derived from the RR interval and 6 characteristics obtained from the R-peak amplitude.

TABLE 3
Feature Extraction

Name	RR Interval	R-Peak Amplitude	Details
MRR	✓		Average of RR intervals
MHR	✓		Average of heart rates
RMSSD	✓		Square root of the mean of successive RR intervals differences
SDNN	✓		Standard deviation of the differences RR intervals
NN50	✓		Number of adjacent RR intervals exceeding 50 ms
pNN50	✓		Proportion NN50 to total RR intervals
VLF	✓	✓	Very low frequencies with 0~0.04 Hz
LF	✓	✓	Low frequencies with 0.04~0.15 Hz
HF	✓	✓	High frequencies with 0.15~0.4 Hz
LF/HF	✓	✓	The ratio of low to high frequency
LF/(LF + HF)	✓	✓	Proportion of low frequency to the sum of low and high frequencies.
HF/(LF + HF)	✓	✓	Proportion of high frequency to the total of low and high frequencies.

FIGURE 5 illustrates one of the results of the RR interval feature extraction, encompassing the MRR (Mean RR Interval). FIGURE 6 depicts the comparison after the median filter is applied.

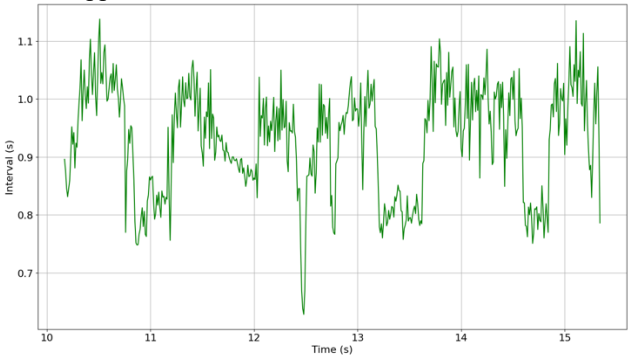

FIGURE 5. RR Interval Result (MRR)

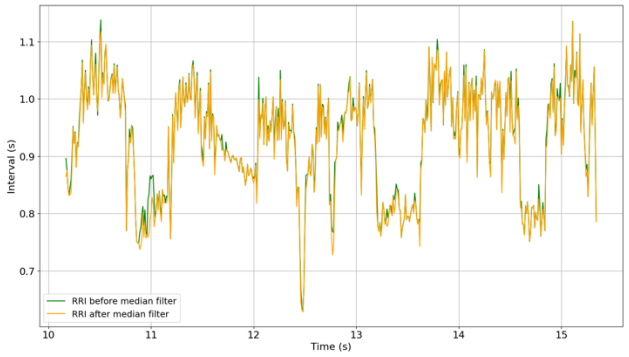

FIGURE 6. Comparison RR Interval Before and After Median Filter (MRR)

The R-peak amplitude denotes the apex of the R-wave, indicating respiration. Consequently, there are six frequency aspects in addition to those derived from the RR interval. FIGURE 7 presents the results of feature extraction from R-peak amplitude, encompassing attributes like VLF.

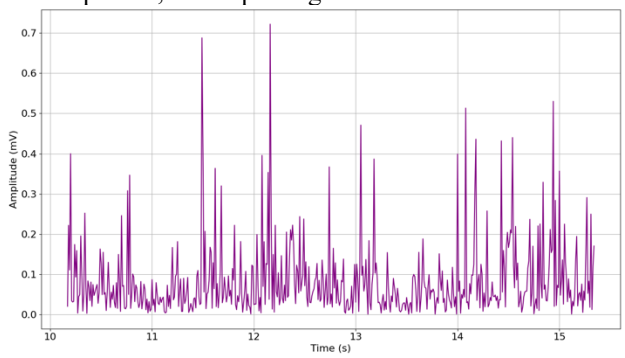

FIGURE 7. R-Peak Amplitude Result (VLF)

D. FEATURE ENGINEERING

Previous research has established a temporal correlation among several components of ECG data. In earlier studies, characteristics derived from contemporary parts of the ECG signal were used as input. This indicates that the algorithm's output relied solely on the characteristics of the present ECG signal segments. To address this issue, a Time Window (TW) was implemented to improve the attributes by incorporating temporal correlations, as illustrated in FIGURE 8.

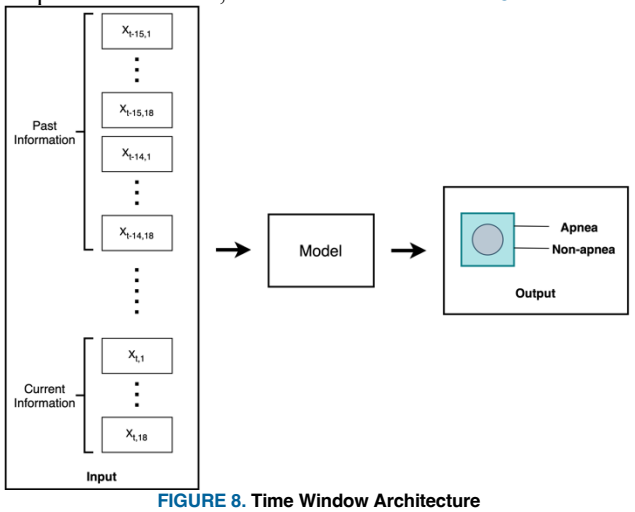

FIGURE 8. Time Window Architecture

The relevant type of TW is the Moving Time Window. The TW size indicates the number of data points or time steps included in each analysis window. For this inquiry, the specified TW size ranges from 0 to 15.

E. MODEL EVALUATION

This research employs stratified K-fold cross-validation with five folds to divide the training data into five distinct subgroups. Bahrami and Forouzanfar [19] cited the selection of five folds to equilibrate the need for a rigorous evaluation of model performance with the requirement for adequate training data in each iteration. We train the model on four subsets in each iteration and validate it on one subset. We evaluate the model's performance using the validation data, with accuracy serving as the primary metric, after training on each fold.

F. MODELING

1) 1D-CNN

Convolutional Neural Networks (CNN) have demonstrated efficacy in signal processing. One variant is the 1D Convolutional Neural Network (1D-CNN), designed primarily for processing signal data. FIGURE 9 depicts the 1D Convolutional Neural Network (1D-CNN) architecture.

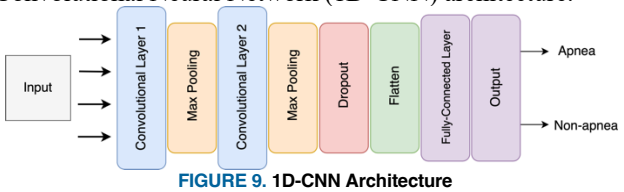

FIGURE 9. 1D-CNN Architecture

The model comprises several layers, including convolutional layers that employ multiple filters to extract features from the input data. Each filter possesses distinct weights and biases, enabling the network to acquire various features. This layer encapsulates spatial hierarchies within data, rendering it efficacious for signal processing tasks [27]. The mathematical representation of the processes conducted in this layer is provided in Eq. (2) [27].

$$x_j^l = f\left(\sum_{i \in M_j} x_j^{l-1} * w_{ij}^l + b_j^l\right) \quad (2)$$

Here, x_j^l represents the feature vector generated by the j^{th} convolution kernel of layer l, M_j represents the feature vector set input to layer l, x_i^{l-1} indicates the feature vector produced by the i^{th} convolution kernel in layer l-1, while w_{ij}^l refers to the weight vector associated with the i^{th} kernel. This study uses maximum pooling, which is usually put after the convolutional layer, to cut down on parameters, speed up processes, and stop overfitting [27]. This study utilizes maximum pooling as delineated in Eq. (3) [27].

$$P_{ij}^l = \max_{(j-1)w+1 \le k \le jw} \{x_{ik}^l\} \quad (3)$$

We denote P_{ij}^l as the result of the pooling process from the j^{th} neuron in the i^{th} channel in the l^{th} layer, where w as the breadth of the pooling region, and x_{ik}^l as the k^{th} data input from the i^{th} channel of the l^{th} layer. The activation layer of a CNN augments its learning capacity by incorporating

nonlinear elements. The convolution procedure is linear, producing a linear combination of information. We utilize activation functions, such as the Rectified Linear Unit (ReLU), to more effectively represent nonlinear interactions [27]. This formula is shown in Eq. (4) [27].

$$ReLU = \begin{cases} x \; if \; x > 0 \\ 0 \; if \; x \leq 0 \end{cases} \quad (4)$$

We employ a dropout layer as a regularization technique to reduce the risk of overfitting. The flattening layer transforms the output of preceding layers from a three-dimensional format into a one-dimensional vector, which is essential for connecting pooling and convolutional layers to fully linked layers [28]. The fully connected layer positioned after the last pooling layer, links every neuron to all feature maps from the preceding layer, extracting advanced features for the classifier, as illustrated in Eq. (5) [29].

$$F_j^l = \sum_{i=1}^{n} x_i^{l-1} * w_{ij}^l + b_j^l \quad (5)$$

In this context, F_j^l represents the neuron's output, whereas w_{ij}^l and b_j^l are the weights and biases, respectively. The output layer often employs the softmax function to generate the final diagnostic result. The computation is defined in Eq. (6) [30].

$$y_k = \frac{e^{a_k}}{\sum_{i=1}^{n} e^{a_i}} \quad (6)$$

In this context, y_k denotes the output of the k^{th} neuron, a_k signifies the k^{th} input signal, and a_i indicates the i^{th} input signal.

2) ADAM OPTIMIZER

Optimization algorithms are crucial in establishing the foundational framework for robots to learn from experience. These methods compute gradients and aim to minimize the loss function. We can employ various optimization algorithms to implement diverse learning strategies. Adam, an acronym for adaptive moments. This method is a synthesis of the RMSprop and momentum optimization strategies. Adam's updating mechanism only considers the smoothed gradient and incorporates a method to correct for bias [31].

G. PERFORMANCE EVALUATION

The research assesses the effectiveness and performance of classification algorithms through a confusion matrix, an essential instrument for machine learning algorithms. The matrix compares he system's classification results against the anticipated outcomes, assessing the model's capability in accurately classifying and predicting features [32]. The confusion matrix in TABLE 4 provides a clear understanding of the model's strengths and weaknesses, revealing its ability to accurately classify data [33].

TABLE 4
Confusion Matrix

	Predicted Label	
True Label	TN	FP
	FN	TP

The study used a confusion matrix to compute performance metrics, which are essential for evaluating the model's efficacy in identifying sleep apnea and differentiating between positive and negative intervals. Below are the equations for these calculations [32].

1. Accuracy (ACC)
   The ratio of accurately identified samples determines the accuracy of sleep apnea detection. It indicates the model's ability to appropriately categorize apnea and non-apnea intervals. Accuracy is determined by summing True Negatives (TN) and True Positives (TP), then dividing by the overall number of samples, as illustrated in Eq. (7)

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (7)$$

2. Sensitivity (SN)
   Sensitivity is the model's capacity to accurately detect sleep apnea cases from all positive samples. It's crucial to avoid false negatives, as low sensitivity can lead to missed diagnoses. High sensitivity reduces false negatives and allows the model to identify as many existing apnea cases as possible. Sensitivity is calculated by dividing True Positives (TP) by total positive samples, as illustrated in Eq. (8)

$$Sensitivity = \frac{TP}{TP + FN} \quad (8)$$

3. Specificity (SP)
   Specificity is defined as a model's ability to accurately detect non-apnea samples from all negative samples. Low specificity can lead to incorrect diagnoses, while high specificity prevents false alarms. Specificity is calculated by dividing True Negatives (TN) by the overall number of negative samples, ensuring an accurate sleep apnea diagnosis, as shown in Eq. (9)

$$Specificity = \frac{TN}{TN + FP} \quad (9)$$

III. RESULTS

This paper presents the findings from assessing the effectiveness of an apnea detection model. The model is predicated on ECG readings and employs Time Window (TW) alongside a 1D Convolutional Neural Network (1D-CNN), utilizing the Adam optimizer for optimization. We execute data normalization using the MinMax Scaler, as described in Eq. (10), which modifies the feature range to span from 0 to 1. This guarantees that the values for each sample of a feature remain within permissible limits [34].

$$x_{sc} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (10)$$

Let x denote the initial value of the feature that requires normalization, with x_{min} and x_{max} representing the minimum and maximum values, respectively. This study employed comprehensive model training approaches, as detailed in TABLE 5. The parameters included model design, hyperparameter configurations, and optimization techniques,

all intended to attain optimal results. This particular information provides essential insights into the methodological technique used.

TABLE 5
Training Parameters for the Model

Hyperparameter	Value
Activation Function	Relu, SoftMax
Learning Rate	0.0001
Optimizer	Adam
Epochs	100
Batch Size	128

A detailed overview of the 1D-CNN architecture used in this research is presented in TABLE 6. This table presents a comprehensive overview of the essential components, including the organization and layout of layers, as well as key parameters utilized in the architecture. This table provides a detailed examination of the individual contributions of each factor influencing the model's overall efficacy, thereby increasing comprehension of how design choices affect sleep apnea diagnosis using ECG data.

TABLE 6
1DCNN Architecture Model

Layer	Parameter
Input	-
1D Conv 1	Number of kernels: 32, kernel size: 5, stride: 1
1D MaxPool	Pooling size: 2
1D Conv 2	Number of kernels: 64, kernel size: 5, stride: 1
1D MaxPool	Pooling size: 2
Dropout	0.7 rate
Fully-Connected Layer	Number of neurons: 32, activation: relu
Output	Number of neurons: 2, activation: softmax

This study utilized a Learning Rate Scheduler to enhance training efficiency. The objective of this scheduler is to progressively reduce the learning rate following the 70th epoch. Every 10 epochs, a factor of 0.1 will diminish the learning rate. We specifically developed this method to enhance the model's stability and boost performance on validation and test data. The findings of this thorough performance analysis are presented in TABLE 7.

TABLE 7
Performance Evaluation

Time Window Size	*Accuracy* (%)	*Sensitivity* (%)	*Specificity* (%)
0	81.65	75.23	85.62
1	85.22	79.10	89.02
2	86.56	80.54	90.32
3	87.67	81.66	91.45
4	88.26	82.76	91.73
5	88.55	83.73	91.60
6	88.83	83.68	92.11
7	89.12	84.35	92.16
8	89.17	84.27	92.31
9	89.38	84.75	92.35
10	89.41	85.04	92.25
11	89.54	84.99	92.49
12	89.61	85.00	92.60
13	89.73	85.41	92.55
14	89.80	85.19	**92.82**
15	**89.87**	**86.16**	92.30

The model's performance evaluation results in this study demonstrate that the 1D-CNN model with a Time Window (TW) combination has improved performance compared to previous research. This model achieves an accuracy score reaching 89.87%, a specificity score reaching 92.30%, and a sensitivity score reaching 86.16% with a TW size of 15. Compared to the study by Rizal et al. [9], which used standard classification methods with Support Vector Machine (SVM) and got an accuracy of 89.5%, these results show a big improvement in performance. Wang et al. [10] conducted research with a Multilayer Perceptron (MLP) in conjunction with a Time Window, attaining an accuracy score reaching 87.3%, sensitivity score reaching 85.1%, and specificity score reaching 88.7%. Conversely, the hybrid ZFNet-BiLSTM model utilized by Bahrami and Forouzanfar [19] achieved an accuracy score reaching 88.13%, sensitivity score reaching 84.26%, and specificity score reaching 92.27%. Wang et al. [21] conducted research using the LeNet-5 model on 1D-CNN, achieving an accuracy score reaching 87.6%, a sensitivity score reaching 83.1%, a specificity score reaching 90.3%, and an AUC score reaching 0.950. Not only does this study add to what has already been written, but it also makes new discoveries by showing how using TW techniques and 1D-CNN can greatly enhance the ability of ECG signals to detect sleep apnea, especially when it comes to the model's capacity to identify intricate patterns in biological signals. A succinct summary of the model architecture is presented in TABLE 8.

TABLE 8
Summary of the 1DCNN Architecture

Layer	Output Shape	Parameter
Input	(None, 288, 1)	0
1D Conv 1	(None, 284, 32)	192
1D MaxPool	(None, 142, 32)	0
1D Conv 2	(None, 138, 64)	10304
1D MaxPool	(None, 69, 64)	0
Dropout	(None, 69, 64)	0
Fully-Connected Layer	(None, 32)	141344
Output	(None, 2)	66

We assessed the system's classification performance using a confusion matrix, as illustrated in FIGURE 10. The model successfully classified 5815 out of 6469 data segments, with 654 misclassified. This indicates the model's robustness in

detecting apnea events. The True Positive (TP) is 2208, and the False Negative (FN) is 354, resulting in a high sensitivity of 86.16%. Conversely, the False Positive (FP) is 300, leading to a specificity of 92.30%. While the model performs well in detecting both apnea and non-apnea cases, it shows slightly more misclassification of negative cases (non-apnea), possibly due to the dataset imbalance. A potential dataset imbalance could influence the model's slight tendency to misclassify negative cases (non-apnea), despite its strong ability to accurately classify positive cases of apnea. This imbalance may cause the model to slightly favor one class over the other to minimize statistical error, though overall performance remains consistent.

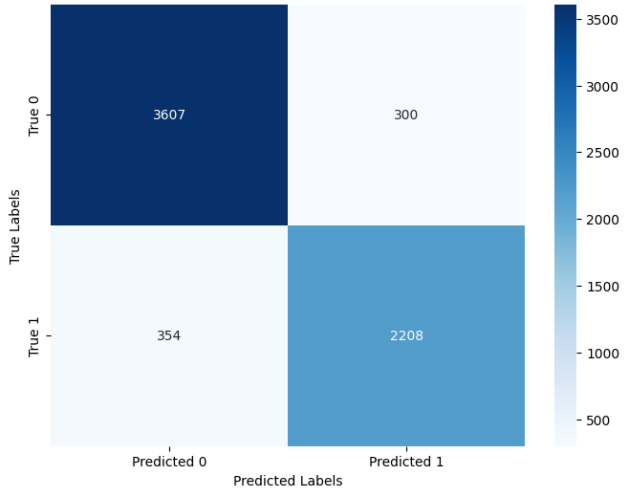

FIGURE 10. Confusion Matrix Result

FIGURE 11 and FIGURE 12 present the graphs illustrating the training loss and validation loss, respectively. FIGURE 11 illustrates the progression of training loss over the epochs, reflecting the model's advancement in error minimization throughout the training process. An even and progressively declining loss curve indicates the model's efficacy in learning from the training data and approaching an optimal solution.

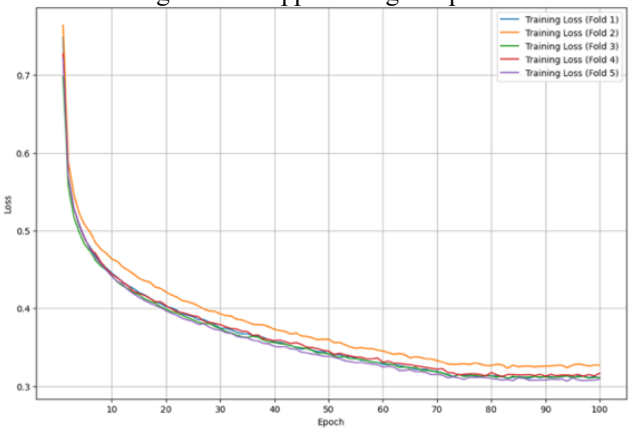

FIGURE 11. Training Loss Result

Conversely, FIGURE 12 illustrates the validation loss, which monitors the model's efficacy on previously unobserved validation data. This picture is especially beneficial for assessing the model's generalization capability. A significantly lower validation loss suggests that the model is not overfitting and has the potential to generalize well to new data, whereas a validation loss curve that levels out or rises may signify potential overfitting.

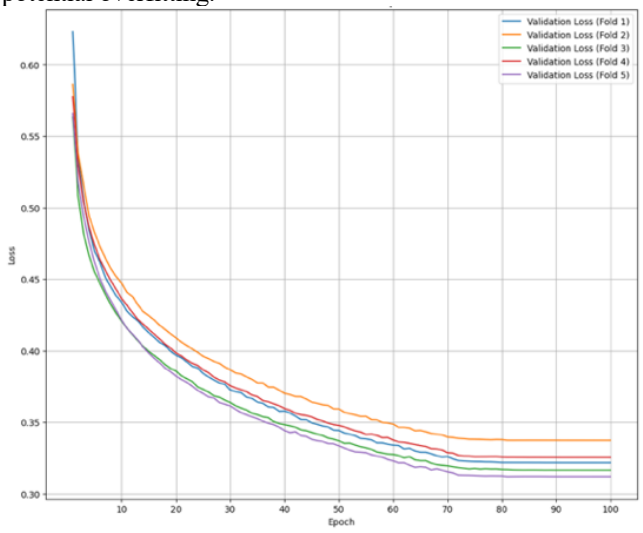

FIGURE 12. Validation Loss Result

Both data demonstrate a consistent decrease in training and validation losses, signifying decreased errors and enhanced performance. Beginning about epoch 71, the rate of decline in both training loss and validation loss diminishes, and the loss values begin to stabilize, indicating that the model has reached a state of convergence.

We assess the model's ability to differentiate between positive and negative classes utilizing the Area Under the Curve (AUC) and the Receiver Operating Characteristic (ROC), as illustrated in FIGURE 13. This graphic depicts the balance between true positive and false positive rates at various threshold levels, providing a thorough assessment of the classifier's effectiveness. The AUC score functions as a singular metric to evaluate the model's ability to differentiate across classes, with a higher AUC signifying superior classification performance.

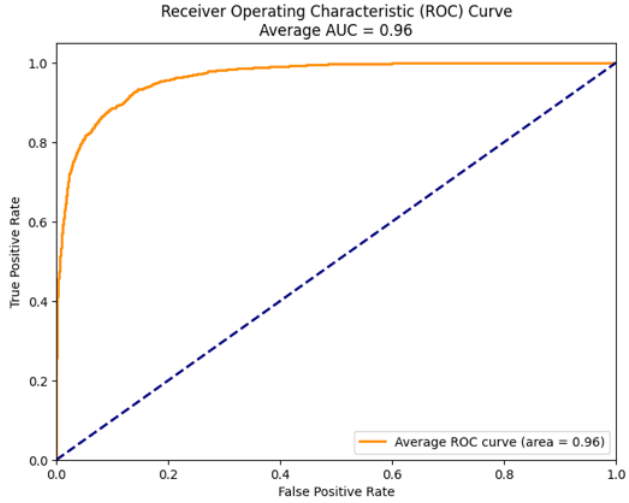

FIGURE 13. ROC & AUC Result

IV. DISCUSSION

This research sought to enhance sleep apnea detection through the integration of 1D-CNN and Time-Wave (TW) methodologies. We employed the Adam optimizer and modified the TW size from 0 to 15. The objective was to identify the optimal configuration to improve the model's accuracy and stability while capturing significant temporal trends. The amalgamation of 1D-CNN and TW methodologies is crucial as it allows the model to discern complex temporal correlations between ECG signals. Conventional techniques frequently fail to detect nuanced differences signifying apnea occurrences. The study provides a thorough methodology to improve detection accuracy by utilizing the advantages of 1D-CNN and TW.

The results expand the understanding of sleep apnea diagnosis by illustrating how the amalgamation of 1D-CNN and TW methodologies significantly improves the model's capacity to grasp intricate temporal characteristics in ECG signals. Current diagnostic frameworks may incorporate the proposed model, enabling healthcare providers to employ improved algorithms for more accurate and timely detection of sleep apnea. This development enhances detection accuracy and has therapeutic implications for improved patient outcomes. FIGURE 14 depicts the correlation between the feature count and the growing TW size.

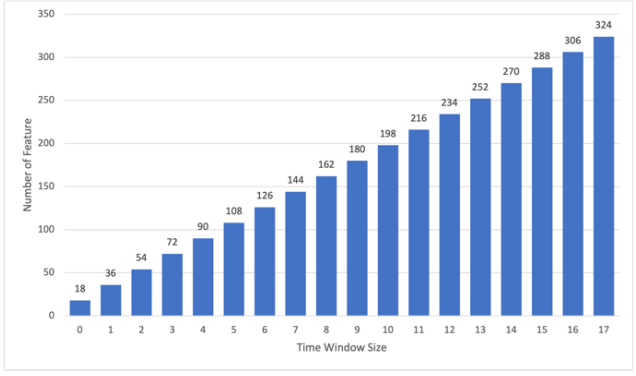

FIGURE 14. Feature Distribution

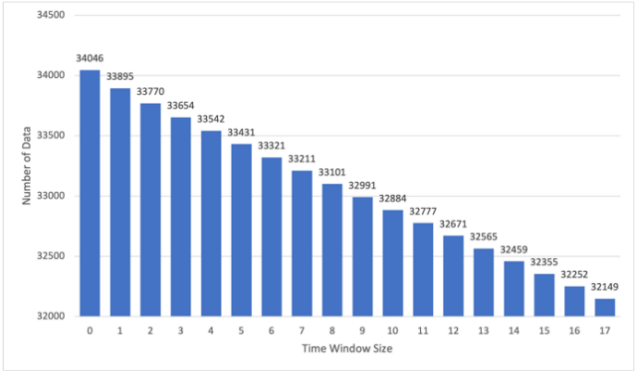

FIGURE 15. Data Distribution

FIGURE 14 illustrates that an increase in TW size to 15 results in a rise in the number of features from 18 to 288. This is attributable to the model's function in encompassing a broader spectrum of data variances and patterns. The enhanced feature extraction capability facilitates a more thorough study of the ECG data, hence augmenting the accuracy of sleep apnea event identification. By manipulating and integrating data from several TW segments, the model generates additional features that allow it to assess temporal information in the ECG signal, resulting in improved performance metrics. Refer to FIGURE 15 to see the decrease in data volume as the TW size increases. FIGURE 15 demonstrates a reduction in data volume from 34046 at TW size 0 to 32355 at TW size 15, underscoring the impact of TW on data organization. As the time window size expands, the emergence of Not a Number (NaN) values during the shifting process leads to the discarding of an increased volume of data, leaving only valid data for subsequent analysis. FIGURE 16 presents the accuracy results.

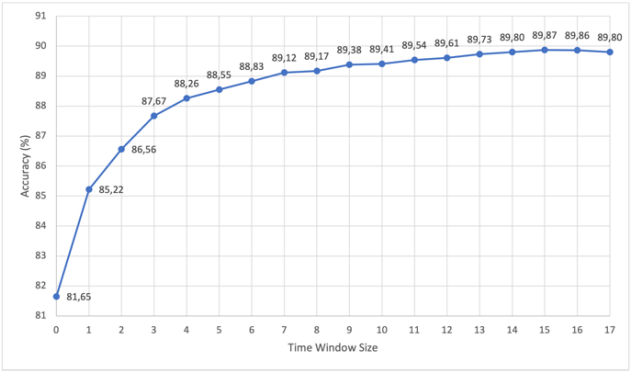

FIGURE 16. Accuracy Result

The model achieves peak accuracy of 89.87% during trials with a temporal window size of 15, as illustrated in FIGURE 16. Increasing the size of the temporal frame leads to more acquired characteristics, which makes it easier to find small changes in ECG signals that are linked to sleep apnea events. The model can collect more complex and important temporal data, which improves its ability to recognize patterns and find changes. This lets it tell the difference between segments that show signs of apnea and those that don't, leading to higher accuracy. For an analysis of sensitivity concerning TW size, please refer to FIGURE 17.

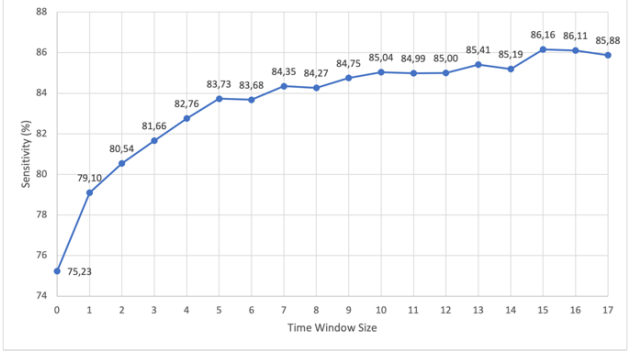

FIGURE 17. Sensitivity Result

The model's sensitivity data indicates a peak of 86.16% with a TW size of 15, as illustrated in FIGURE 17. This is due to its ability to utilize features from a larger dataset, thereby improving its proficiency in recognizing complex patterns and temporal data. As the model improves its ability to identify genuine apnea cases, its sensitivity increases, resulting in a

higher rate of accurate positive case detection. Please refer to FIGURE 18 for an analysis of the TW size specificity results.

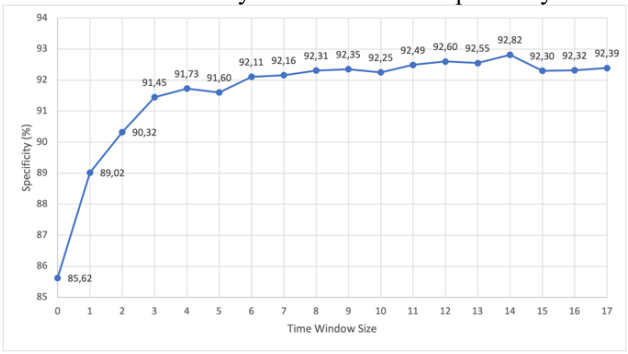

FIGURE 18. Specificity Result

FIGURE 18 shows that a TW size of 14 achieves the maximum specificity of 92.82%. A TW size of 15 exhibits enhanced accuracy and sensitivity, although it fails to achieve ideal specificity. Increased TW sizes can improve detection capabilities, but may also create complexity that impacts specificity. Consequently, a TW size of 14 provides a superior balance between sensitivity and specificity. The performance of this study is assessed through a systematic analysis and comparison of results with prior investigations, as presented in TABLE 9.

TABLE 9
Comparison of Result with Past Research

Research	Algorithm	ACC (%)	SN (%)	SP (%)	AUC
[9]	SVM	89.5	-	-	-
[10]	TW-MLP	87.3	85.1	88.7	0.945
[19]	1DCNN-DRNN	88.13	84.26	92.27	-
[21]	1DCNN	87.6	83.1	90.3	0.950
Proposed Research	TW-1DCNN	**89.87**	**86.16**	**92.30**	**0.96**

The research indicates that the Time Window approach utilizing the TW-1DCNN with Time Window size of 15 is the most efficacious for identifying patients with apnea. The TW-1DCNN achieved the maximum accuracy and sensitivity required for detecting individuals experiencing apnea events, enabling prompt intervention. The approach attained an accuracy score reaching 89.87%, a sensitivity score reaching 86.16%, a specificity score reaching 92.30%, and an AUC score reaching 0.96, it was observed that increasing the TW size leads to a larger feature set, allowing the model to capture more variation and patterns in the data.

The 1D-CNN with a time window size of 5 attained accuracy score reaching 88.55%, sensitivity core reaching 83.73%, and specificity score reaching 91.60%, respectively. Although both the TW-MLP [10] and TW-1DCNN had an identical Time Window size of 5, the 1D-CNN consistently surpassed the MLP regarding accuracy and specificity. The improved performance is because the 1D CNN can extract deeper, more complex information through convolutional layers. This is important for finding the delicate patterns connected to sleep apnea events. However, the MLP's rudimentary feature extraction technique constrains it, resulting in diminished accuracy.

The model's ability to accurately classify positive cases of apnea is strong, but it struggles more with detecting negative cases due to the dataset imbalance. This imbalance can cause the model to favor one class over another to minimize statistical error, leading to inconsistent performance. Future research should prioritize improving sleep apnea detection systems by developing more effective feature selection approaches to find the most relevant features. This may diminish computer complexity while preserving high accuracy, rendering the method more viable for clinical applications. Enhancing feature selection may assist healthcare professionals in more accurately recognizing sleep apnea, ultimately improving patient outcomes through earlier and more precise diagnosis.

V. CONCLUSION

This research employs feature extraction methods and deep learning algorithms the detection of sleep apnea. The method involves seven stages: data collection, data preprocessing, feature extraction, feature engineering, model evaluation, modelling, and performance evaluation. The experiment merges 12 characteristics from the RR interval and 6 from the R-peak amplitude. The 1D Convolutional Neural Network (1D-CNN) with the Adam optimizer is implemented, and a Time Window (TW) varies from 0 to 15. The results show that the model with a TW size of 15 achieved the maximum performance, achieving an accuracy score reaching 89.87%, sensitivity score reaching 86.16%, specificity score reaching 92.30%, and AUC value reaching 0.96. The goal of this study is to present a different way to find sleep apnea by analyzing single-lead ECG data. The results show that the proposed model produces better outcomes than previous studies. The broader implications of this research extend to medical informatics and sleep medicine, suggesting that advanced algorithms could facilitate more precise diagnostics and personalized treatment plans for patients with sleep apnea. The model could be implemented in clinical settings, such as sleep clinics and cardiology departments, to assist healthcare professionals in accurately diagnosing sleep apnea. The model effectively classifies positive apnea cases, but struggles to identify negative ones due to a dataset imbalance, suggesting the need for a balanced dataset for improved efficacy. Future research should focus on enhancing sleep apnea detection methods by focusing on more efficient feature selection methods to determine the most relevant features.

REFERENCES

[1] S. Taran and V. Bajaj, "Sleep Apnea Detection Using Artificial Bee Colony Optimize Hermite Basis Functions for EEG Signals," *IEEE Trans Instrum Meas*, vol. 69, no. 2, pp. 608–616, Feb. 2020, doi: 10.1109/TIM.2019.2902809.

[2] J. Vanek *et al.*, "Obstructive sleep apnea, depression and cognitive impairment," Aug. 01, 2020, *Elsevier B.V.* doi: 10.1016/j.sleep.2020.03.017.

[3] C. H. Lin, R. C. Lurie, and O. D. Lyons, "Sleep Apnea and Chronic Kidney Disease: A State-of-the-Art Review," Mar. 01, 2020, *Elsevier Inc*. doi: 10.1016/j.chest.2019.09.004.

[4] C. Song, K. Liu, X. Zhang, L. Chen, and X. Xian, "An Obstructive Sleep Apnea Detection Approach Using a Discriminative Hidden Markov Model from ECG Signals," *IEEE Trans Biomed Eng*, vol. 63, no. 7, pp. 1532–1542, Jul. 2016, doi: 10.1109/TBME.2015.2498199.

[5] S. Q. Ali, S. Khalid, and S. B. Belhaouari, "A Novel Technique to Diagnose Sleep Apnea in Suspected Patients Using Their ECG Data," *IEEE Access*, vol. 7, pp. 35184–35194, 2019, doi: 10.1109/ACCESS.2019.2904601.

[6] M. Bahrami and M. Forouzanfar, "Detection of sleep apnea from single-lead ECG: Comparison of deep learning algorithms," in *2021 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2021 - Conference Proceedings*, Institute of Electrical and Electronics Engineers Inc., Jun. 2021. doi: 10.1109/MeMeA52024.2021.9478745.

[7] M. De Zambotti *et al.*, "Impact of evening alcohol consumption on nocturnal autonomic and cardiovascular function in adult men and women: A dose-response laboratory investigation," 2020, doi: 10.1093/sleep/zsaa135/5871424.

[8] R. B. Shouldice, L. M. O'Brien, C. O'Brien, P. De Chazal, D. Gozal, and C. Heneghan, "Detection of Obstructive Sleep Apnea in Pediatric Subjects using Surface Lead Electrocardiogram Features," *Sleep*, vol. 27, no. 4, pp. 784–792, 2004, doi: 10.1093/sleep/27.4.784.

[9] A. Rizal, F. D. A. A. Siregar, and H. T. Fauzi, "Obstructive Sleep Apnea (OSA) Classification Based on Heart Rate Variability (HRV) on Electrocardiogram (ECG) Signal Using Support Vector Machine (SVM)," *Traitement du Signal*, vol. 39, no. 2, pp. 469–474, Apr. 2022, doi: 10.18280/ts.390208.

[10] T. Wang, C. Lu, and G. Shen, "Detection of Sleep Apnea from Single-Lead ECG Signal Using a Time Window Artificial Neural Network," *Biomed Res Int*, vol. 2019, 2019, doi: 10.1155/2019/9768072.

[11] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: An astounding baseline for recognition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, IEEE Computer Society, Sep. 2014, pp. 512–519. doi: 10.1109/CVPRW.2014.131.

[12] Y. Wei *et al.*, "HCP: A flexible CNN framework for multi-label image classification," *IEEE Trans Pattern Anal Mach Intell*, vol. 38, no. 9, pp. 1901–1907, Sep. 2016, doi: 10.1109/TPAMI.2015.2491929.

[13] B. Sedighi, I. Palit, X. S. Hu, J. Nahas, and M. Niemier, "A CNN-Inspired Mixed Signal Processor Based on Tunnel Transistors," EDA Consortium, 2015. doi: 10.7873/DATE.2015.0053.

[14] Y. H. Kwon, S. B. Shin, and S. D. Kim, "Electroencephalography based fusion two-dimensional (2D)-convolution neural networks (CNN) model for emotion recognition system," *Sensors (Switzerland)*, vol. 18, no. 5, May 2018, doi: 10.3390/s18051383.

[15] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network," in *Neural Networks*, Elsevier Ltd, 2003, pp. 555–559. doi: 10.1016/S0893-6080(03)00115-1.

[16] R. Haidar, I. Koprinska, and B. Jeffries, "Feature Learning and Data Compression of Biosignals Using Convolutional Autoencoders for Sleep Apnea Detection," *International Conference on Neural Information Processing*, pp. 162–174, 2019, doi: 10.1007/978-3-030-36708-4_14.

[17] U. Erdenebayar, Y. J. Kim, J. U. Park, E. Y. Joo, and K. J. Lee, "Deep learning approaches for automatic detection of sleep apnea events from an electrocardiogram," *Comput Methods Programs Biomed*, vol. 180, Oct. 2019, doi: 10.1016/j.cmpb.2019.105001.

[18] K. McClure, B. Erdreich, J. H. T. Bates, R. S. McGinnis, A. Masquelin, and S. Wshah, "Classification and detection of breathing patterns with wearable sensors and deep learning," *Sensors (Switzerland)*, vol. 20, no. 22, pp. 1–13, Nov. 2020, doi: 10.3390/s20226481.

[19] M. Bahrami and M. Forouzanfar, "Sleep Apnea Detection from Single-Lead ECG: A Comprehensive Analysis of Machine Learning and Deep Learning Algorithms," *IEEE Trans Instrum Meas*, vol. 71, 2022, doi: 10.1109/TIM.2022.3151947.

[20] S. Widodo, H. Brawijaya, and S. Samudi, "Stratified K-fold cross validation optimization on machine learning for prediction," *Sinkron*, vol. 7, no. 4, pp. 2407–2414, Oct. 2022, doi: 10.33395/sinkron.v7i4.11792.

[21] T. Wang, C. Lu, G. Shen, and F. Hong, "Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network," *PeerJ*, vol. 2019, no. 9, 2019, doi: 10.7717/peerj.7731.

[22] T. Penzel', G. B. Moody2j3, R. G. Mark2j3, A. L. Goldberges, and J. H. Peter', "The Apnea-ECG Database," 2000. doi: 10.1109/CIC.2000.898505.

[23] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.," *Circulation*, vol. 101, no. 23, 2000, doi: 10.1161/01.cir.101.23.e215.

[24] W. H. Ali and E. S. Kohiwole, "Design and implementation of low-pass, high-pass and band-pass Finite Impulse Response (FIR) filters using FPGA," in *Digital Signal Processing (DSP): Fundamentals, Techniques and Applications*, Nova Science Publishers, Inc., 2016, pp. 131–151. doi: 10.4236/cs.2015.62004.

[25] P. Hamilton and E. P. Limited, "Open Source ECG Analysis," 2002. doi: 10.1109/CIC.2002.1166717.

[26] L. Chen, X. Zhang, and C. Song, "An automatic screening approach for obstructive sleep apnea diagnosis based on single-lead electrocardiogram," *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 1, pp. 106–115, Jan. 2015, doi: 10.1109/TASE.2014.2345667.

[27] B. Hu, J. Liu, R. Zhao, Y. Xu, and T. Huo, "A New Fault Diagnosis Method for Unbalanced Data Based on 1DCNN and L2-SVM," *Applied Sciences (Switzerland)*, vol. 12, no. 19, Oct. 2022, doi: 10.3390/app12199880.

[28] K. Yang, Z. Huang, X. Wang, and X. Li, "A Blind Spectrum Sensing Method Based on Deep Learning," *Sensors*, vol. 19, no. 10, May 2019, doi: 10.3390/s19102270.

[29] X. Li, Y. Li, Y. Cao, S. Duan, X. Wang, and Z. Zhao, "Fault Diagnosis Method for Aircraft EHA Based on FCNN and MSPSO Hyperparameter Optimization," *Applied Sciences (Switzerland)*, vol. 12, no. 17, Sep. 2022, doi: 10.3390/app12178562.

[30] F. Xu, Z. Sui, J. Ye, and J. Xu, "Ternary Precursor Centrifuge Rolling Bearing Fault Diagnosis Based on Adaptive Sample Length Adjustment of 1DCNN-SeNet," *Processes*, vol. 12, no. 4, Apr. 2024, doi: 10.3390/pr12040702.

[31] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014.

[32] Ž. Đ. Vujović, "Classification Model Evaluation Metrics," *IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, p. 2021, 2021, doi: 10.14569/IJACSA.2021.0120670.

[33] R. Shoghi, L. Morand, D. Helm, and A. Hartmaier, "Optimizing machine learning yield functions using query-by-committee for support vector classification with a dynamic stopping criterion," *Comput Mech*, vol. 74, pp. 447–466, 2024, doi: 10.1007/s00466-023-02440-6.

[34] V. R. Prasetyo, M. Mercifia, A. Averina, L. Sunyoto, and Budiarjo, "Prediksi Rating Film pada Website IMDB Menggunakan Metode Neural Network," *NERO (Networking Engineering Research Operation)*, vol. 7, no. 1, pp. 1–8, 2022.

AUTHOR BIOGRAPHY

Fadil Pratama continued his education since 2020 at Universitas Sebelas Maret, Surakarta. He has been pursuing his studies in the Department of Informatics within the Faculty of Information Technology and Data Science. Currently, he is dedicating his efforts to conducting research in an area of data science, which coincidentally aligns with his primary academic and professional interests. This program has provided him with the opportunity to explore the intricacies of data science and its applications, particularly in the healthcare domain. The main goal of his final task is to conduct innovative research on the detection of sleep apnea using advanced machine learning techniques.

Wiharto is a senior lecturer in the Informatics Department, Faculty of Information Technology and Data Science at Universitas Sebelas Maret, Surakarta. He has a strong academic background and extensive experience in the field of Computational Science & Engineering. Wiharto earned his doctoral degree from Universitas Gadjah Mada, where he specialized in Bio-Medical Informatics. His research primarily focuses on Bio-Medical Informatics, Artificial Intelligence, and Computational Intelligence, with an emphasis on solving real-world problems through innovative computational approaches. His contributions extend to mentoring students, publishing impactful research, and collaborating on interdisciplinary projects to advance the fields of health informatics and artificial intelligence.

Umi Salamah is a dedicated lecturer at the Department of Informatics within the Faculty of Information Technology and Data Science at Universitas Sebelas Maret (UNS) in Surakarta, Indonesia. Her academic journey reflects a strong commitment to the field of Computational Science and Engineering, where she has made significant contributions. She earned her doctoral degree in Computer Science from Institut Teknologi Sepuluh Nopember (ITS), one of Indonesia's most prestigious technological universities. Her research interests primarily revolve around fuzzy logic, a field where she has developed innovative methods and applications. In addition to her research, she actively mentors students, fostering the next generation of computer scientists.