ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

Optimized CNN Model with Derived Kernels for Apnea Classification Application

Smruthy A[1] and Suchetha M[2]

ABSTRACT

Sleep related disorders are common in both men and women, irrespective of their age. In recent years, there has been an abrupt increase of interest in the prediction of sleep related disorders among researchers. The most common data analysis approach utilizes a large data set to predict sleep apnea episodes. The computational complexity of such predictions is high and most of the techniques fail to select the optimal features. The recent trends show convolution neural networks gaining in popularity because of their ability to select the optimal features and the dimension reduction property. On the other hand, it is also important to address the noise issues in the acquired physiological data. Variational Mode Decomposition (VMD) is an adaptive way of decomposing the different frequency levels, such that the noise level can be easily separated into levels. This paper presents a novel technique for optimal feature extraction from the Variational Mode Function (VMF) levels by using the concept of a Convolution Neural Network (CNN). The proposed approach provides an accuracy of 98.4375% with and average time complexity of 0.7471 seconds.

Article information:

Keywords: Variational Mode Decomposition, Computational Complexity, Sleep Apnea, Convolution Neural Network

Article history:
Received: April 7, 2021
Revised: August 28, 2021
Accepted: December 6, 2021
Published: July 30, 2022
(Online)

DOI: 10.37936/ecti-cit.2022163.246994

1. INTRODUCTION

The name *Apnea conveys the meaning want of breath*. Similar to the name, sleep apnea defines the continuous pausing of breath more than 10 seconds which occurs several times in an hour. During such situations, the patient exhibits some snoring kind of symptoms. The direct effect of this pausing leads to blood oxygen reduction and indirectly it may cause Intelligent Quotient (IQ) loss and heart attacks. One of the common symptom of this disease is day-time sleepiness. The screening of sleep apnea is done using Polysomnography (PSG). Polysomnography is the traditional process of recording and analysing the physiological signals of a patient during sleep time. This process of evaluation is time consuming and expensive. The accuracy of detection is also reduced due to human errors. Recent research trends show adoption of an evaluation of sleep related disorder detection based on feature extraction and classification techniques that gives more accuracy. In such cases the features extracted from the nasal airflow signal [1–4], finger photoplethismoraphy, and actigraphy are used for classification [2, 5–9]. The features derived from Electrocardiogram (ECG) [10, 11], Electroencephalogram (EEG) [12], Electromyogram (EMG) [14], Peripheral Capillary Oxygen Saturation [15, 16], and Electrooculogram (EOG) [12] are separately classified using a specific classifier model. The computational complexity of the detection is high in such situations. The popularity of sleep apnea detection with an ECG signal has increased because it is more accurate than the other types of detection techniques. Even though this method is popular among recent researchers, there are difficulties in the optimal feature extraction from the ECG signal segments caused by noise. The Variational mode decomposition algorithm was introduced by Konstantin Dragomiretskiy et al.[17][18] as an effective method for optimal feature extraction from the different decomposed modes. In addition to this, variational mode decomposition is a mathematical technique for extracting the different decomposed modes. Hence, the computational complexity of the classification is high because of the large data set obtained after the VMD process. The convolution neural network is the most recent classification method that selects optimal features from the signals and performs the classification simultaneously [19] [20]. The main advantage of CNN is the dimension reduction property. In earlier

[1,2]The authors are with VIT University, Chennai, Tamilnadu, India, E-mail: smruthy.a2014@vit.ac.in and suchetha.m@vit.ac.in

research work, CNN was used only for image classification problems [21][22]. Later the properties of CNN have been utilized for 1D classification problems also [23][24]. This paper introduces novel feature extraction from noise free VMD levels using the properties of CNN.

1.1 Related Works

This subsection explores the most recent works related to the proposed method. As mentioned in the previous section, the main advantage of the proposed methodology is optimal feature extraction by reducing or down sampling the feature set. Thus the computation burden of the entire system will be reduced. There are different techniques previously published which are related to optimal feature extraction. One of the above-mentioned work is the deep neural network based method utilized for detecting apnea episodes. In this technique, the authors tried to find the optimal hyper-parameters of the DNN structure. This was achieved by using the Differential Evolution (DE) concept [25]. The contextual information from EEG epochs can also be considered as a biomarker for sleep apnea detection. Kun Chen et al. proposed a technique similar to this [26]. They developed a two-stage screening model to detect apnea episodes using a multi-scale CNN architecture. The Sleep EDF database and OSA database are used for the apnea detection process. In 2019, Fu-Tai Wang et.al proposed a new technique to detect the respiratory drops associated with sleep apnea syndrome [27]. Two models were developed based on the Obstructive Sleep Apnea (OSA) severity index. One is the regression model that finds the Apnea Hypopnea Index and the other one is the severity measuring technique using Respiratory Fluctuation Index (RFI) values. In a recent work, Viswabhargav et al. adopted a novel feature extraction technique from ECG derived Respiratory Signal (EDR). The spars entropy features are segregated from the EDR and ECG signals [28]. The classification of the apnea subjects is performed using fuzzy K-means clustering and SVM classifier models. In a similar way, the temporal features from the ECG signals can be extracted using the deep recurrent neural network concept. The classification is performed using a one dimensional convolution neural network along with a fully connected layer [29]. The important shortcoming of these papers is the failure to consider the noise issues present in the ECG signal. In order to avoid that, in our proposed work, we develop an adaptive kernel which removes the unwanted noise from the ECG signal.

2. MATERIALS AND METHODS

This section provides a detailed explanation of the data acquisition and classification techniques. The ECG signal is mixed with diverse noises like baseline and power line interference. The accurate segregation of apnea and healthy subjects requires a powerful signal conditioning algorithm. The variational mode decomposition is a signal conditioning algorithm which decomposes the entire signal into different modes of specific frequencies.

2.1 VMD Algorithm

In Variational Mode Decomposition, the ECG signal is decomposed into different variational modes. Each of these modes exhibits a particular sparsity property when recreating the signal [17][18]. The Variational Mode Decomposition technique can be explained using three steps. In the first step, the Hilbert transform is utilized to obtain a distinct frequency spectrum of the signal. This spectrum of the particular decomposed mode is then drifted to its base band with an exponential function. This function can be created by using its respective center frequency. In the final step, the estimation of bandwidth is carried out by utilizing Gaussian smoothing of the detected signal. The corresponding variational equation for optimization is represented in equation (1).

$$\min_{\{u_l\},\{w_l\}} \left\{ \sum_l ||(\partial_t[\delta(t) + j/\pi t) * u_l]e^{(-jw_l t)}||_2^2 \right\} \tag{1}$$

s.t

$$\sum_l u_l = f$$

In the equation (1), $u_l := u_1, u_2, \ldots, u_l$ represents the different extracted modes and $w_l := w_1, w_2, \ldots, w_l$ represents their respective center frequencies.

2.2 The Bandwidth Minimization Technique

The bandwidth optimization procedure is implemented by applying the Alternate Direction Method of Multipliers (ADMM)technique [17][18]. The first step of this procedure involves initializing the primary iteration values of each decomposed mode. This primary iteration value is incremented for the calculation of successive iteration values. The convergence of the modes is checked in between the mode calculations. If the convergence criteria are satisfied, then the corresponding value of l is incremented. If this convergence is not happening, then the corresponding mode value is incremented for calculating the next mode and center frequency. The above mentioned procedure is repeated until all of the modes are converged. Equations (2) and (3) represent the various decomposed modes and their respective center frequencies.

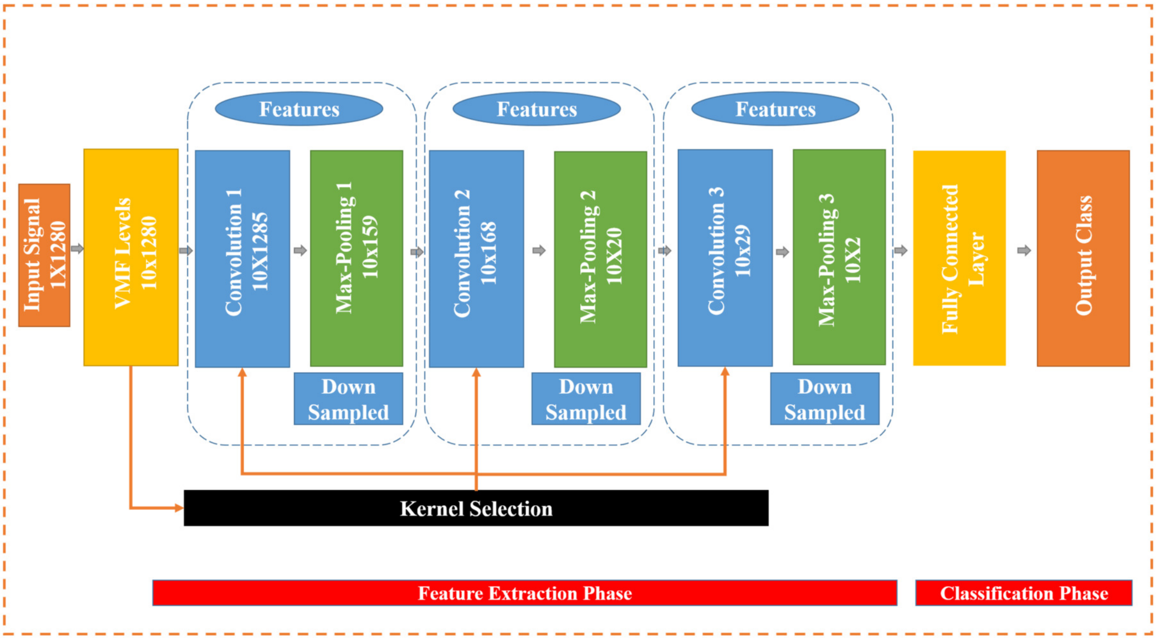

Fig.1: *The proposed improved CNN model architecture.*

$$\mathrm{u}_l^{(n+1)}(w) = \frac{\left(f(w) - \sum_{i<l} u_l^{(n+1)}(w) - \sum_{i>l} u_i^n(w) + \frac{(\lambda^n(w))}{2}\right)}{(1+2\propto(w-w_l^n)^2)} \quad (2)$$

$$w_l^{n+1} = \frac{\int_0^\infty w|u_l^{n+1}(w)|^2 dw}{\int_0^\infty |u_l^{n+1}(w)|^2 dw} \quad (3)$$

The dual ascent for the values, whose enter frequency is equal or greater than zero, can be expressed as

$$\lambda^{(n+1)}(w) = \lambda^n(w) + \tau(f(w)) - \sum_k u_l^{n+1}(w) \quad (4)$$

The mathematical representation for convergence of each mode is given by equation (5).

$$\sum_l \frac{||u_l^{n+1} - u_l^2||_2^2}{||u_l^2||_2^2} < \epsilon \quad (5)$$

3. THE IMPROVED CNN ARCHITECTURE

The CNN architecture is a variant of a multilayer perceptron which resembles the visual cortex of an animal. The visual cortex contains the complex arrangements of cells. These cells helps to identify the different patterns in the input images. This architecture is designed to select the optimal feature values without utilizing the separate feature extraction and classification algorithms. Also, it helps to reduce the computational time of the analysis. The main architecture of CNN comprised of five layers. These are input layer, some convolution layers, few max-pooling layers, a fully connected multilayer perceptron and an outcome or output layer [31]. In our proposed work, the raw data samples from the patient are filtered using a convolution layer. The corresponding filtered samples are then down sampled by utilizing a max-pooling layer. Thus the feature set is reduced and the optimal features are obtained for classification. These primary layers are representing the feature extraction method. A fully connected multilayer perceptron concept is utilized for the classification purpose. The entire framework of this modified convolution model is explored in the imminent sections. The architecture of the proposed improved convolution neural network model is represented in fig.1

3.1 Improved convolution layer with derived kernels

The convolution layer usually works with the principles of the convolution operation. This layer is depicted to convert the input values to a particular reduced feature set. This process of conversion is achieved by taking the summation of the dot product values of two functions. The first function represents the input signal and the second one is a kernel function. The dot product of these functions are taken after the shifting and reversing of one of the functions. The above mentioned kernel function is used for filtering the input signals.

3.1.1 Derived kernels

In our proposed work, the filtering of the input signals during the convolution operation is achieved

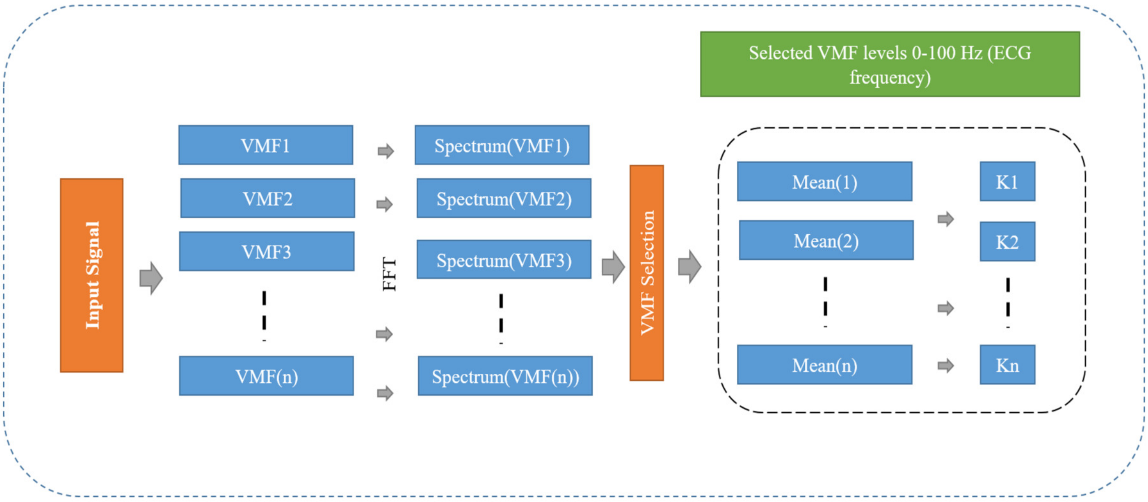

Fig.2: *The kernel selection method for improved CNN network.*

using a kernel which is automatically derived from the noise free VMF levels. The mean of these selected VMF levels is utilized as the kernel of the convolution operation. The main advantage of this type of kernel selection is that the feature set obtained after the convolution is noise free and resembles the original ECG signal features. The improved architecture of our derived kernel is represented in Fig. 2. Let us consider the convolution of two functions, input and the kernel. The inputs to the convolution layer are nothing but the different VMF levels (vmf_l of length i) obtained after performing the VMD algorithm on raw ECG signals. A kernel of length n is the other function, which represents the mean of different noise free VMF levels. The outcome of the convolution layer with a length of $(i+n-1)$ can be expressed as

$$o_l(1) = vmf_l(1)k_l(1)$$

$$o_l(2) = vmf_l(1)k_l(2) + vmf_l(2)k_l(1)$$

$$o_l(3) = vmf_l(1)k_l(3) + vmf_l(2)k_l(2) + vmf_l(3)k_l(1)$$

$$o_l(n) = vmf_l(n)k_l(1) + vmf_l(n-1)k_l(2) + \cdots + vmf_l(0)k_l(n)$$

where, vmf_l represents the different levels obtained after VMD. Thus, these equations can be generalized as

$$o_l(n) = \sum_{r=-i}^{i} (vmf_l(p+1)k_l(n-p+1)) \tag{6}$$

3.2 The Max-pooling Layer

The dimension of the feature sets obtained after the convolution become large. It is important to minimize the feature set to obtain the optimal feature map. This is achieved by down sampling the entire feature set. The procedural idea of down sampling is the calculation of the maximum value of this entire feature set in a short window span. The window length is calculated by considering the non-overlapped windowing method. Fig. 3 represents the max-pooling operation with non-overlapped windowing method. A fully connected multilayer perception is utilized for classifying the down sampled feature sets. The max-pooling operation can be mathematically represented as

$$O_l = \max(o_l^*) \tag{7}$$

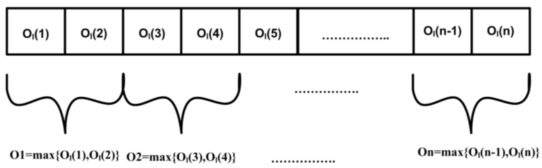

Fig.3: *The schematic representation of Max-pooling operation with non-overlapped window of length 2.*

3.3 Fully Connected Layer

The fully connected multilayer perceptron is the last layer in the CNN architecture. It gives the prediction and classification of the feature map into different classes. A fully connected layer works similarly to the conventional neural network architecture. The network layer maps the different feature sets from the previous layers into different classes, which represents the output layer. Finally, the output layer is represented as

$$y_l = \tanh\left(\sum_{l=1}^{N} O_l w_l + b_l\right) \quad (8)$$

The mapping is done by an activation function called tanh function. This function can be expressed as

$$\tanh(l) = \frac{e^l - e^{-l}}{e^l + e^{-l}} \quad (9)$$

The outcome of this activation function lies between [-1 1]. The back propagation learning algorithm is implemented for training the network. It will reduce the error between the estimated outcome and the actual outcome.

3.3.1 Training the Proposed Improved CNN Architecture

The training phase of the CNN architecture adopts the back propagation learning algorithm along with the stochastic gradient descent technique [30]. During the forward propagation phase, the mean square error of the proposed CNN increases because of the random selection of the weights and biases. This mean square error can be represented as

$$E = \sum (d_l - y_l) \quad (10)$$

where d_l is the actual output and y_l is the estimated output. The estimated error reduction can be achieved by utilizing the concept of gradient descent. In the gradient descent technique, an iterative estimation of weight and biases is performed. Thus the classification error is reduced. The error reduction is achieved by using the partial derivatives of this error in terms of weight and bias. In the case of the ordinary back propagation algorithm, this partial derivative of the error in terms of weight is represented by multiplying and dividing this value with the partial derivative of estimated output (y_l). The above mentioned error reduction procedure can be mathematically represented as

$$\frac{\partial E}{\partial w_l} = \frac{\partial E}{\partial y_l}\frac{\partial y_l}{\partial w_l} = \frac{\partial E}{\partial y_l} y_l \quad (11)$$

Where

$$\frac{\partial y_l}{\partial w_l} = y_l \quad (12)$$

Hence

$$\frac{\partial E}{\partial w_l} = \triangle E y_l \quad (13)$$

and

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial y_l}\frac{\partial y_l}{\partial b} = \frac{\partial E}{\partial y_l} \quad (14)$$

So

$$\frac{\partial E}{\partial b} = \triangle E \quad (15)$$

In a convolution neural network, these weight and bias estimations are mathematically derived by

$$\frac{\partial E}{\partial w_l} = conv(y_l, \triangle E) \quad (16)$$

and

$$\frac{\partial E}{\partial b} = \sum \triangle E \quad (17)$$

3.3.2 Testing the Improved CNN Model

The testing of the feature sets is done using the improved CNN model with the optimized weight and bias values obtained after the training phase. The MLP algorithm with reduced error rate gives the best classification accuracy compared to the other networks.

4. RESULT AND DISCUSSION

4.1 Data Collection

4.1.1 Online Data Collection

In this study, the data collection from the subjects was performed with both online and offline processes. The collected online and offline data sets are treated separately for classification purposes. The data collection procedure was purely based on the principles of the Declaration of Helsinki. The Forty healthy subjects of age group between 26 to 41 voluntarily participated in the data collection and experimentation. All the experimental procedures were performed in signal processing lab at the school of electronics engineering. All of these healthy subjects were told that this study was performed for a research work related to sleep apnea detection. An hourly based experimentation was adopted for performing the data collection. At the beginning stage of the experiment, the participants were asked to sit in a calm and relaxed position. During the experiment, all of these participating subjects were asked to hold their breath for at least 10 seconds repeated 5 to 6 times in an hour, and provided with relaxation time The ECG signals of these subjects during the experiment were collected and exported to Matlab 2016a using an Arduino UNO micro controller board. The data acquisition part of the sensory system consists of ECG disposable electrodes and a Sparkfun Electronics AD8232 heart rate monitoring system. The entire test setup was interfaced to a laptop with a baud rate of 9600. The serial port values from this system were exported to Matlab 2016a. The entire test setup along with the experiment is represented in Fig. 4

Fig.4: *The ECG signal acquired with the test setup (a)The data acquisition system (b)The ECG signal acquired with the test setup.*

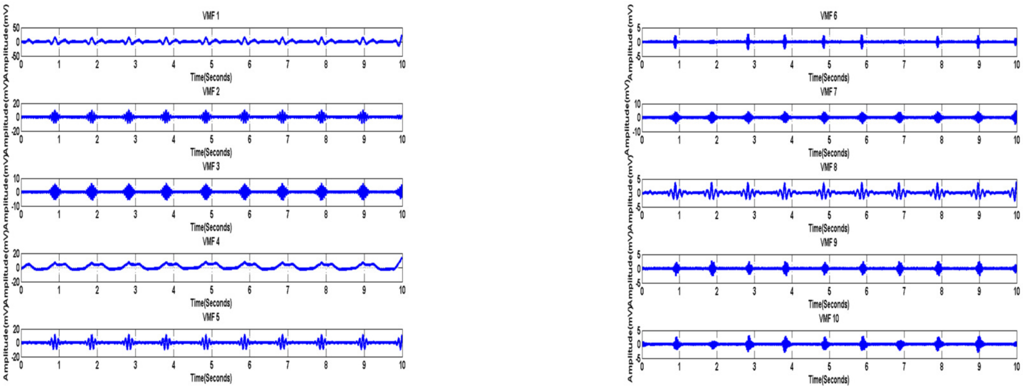

Fig.5: *The different decomposition modes obtained after performing VMD.*

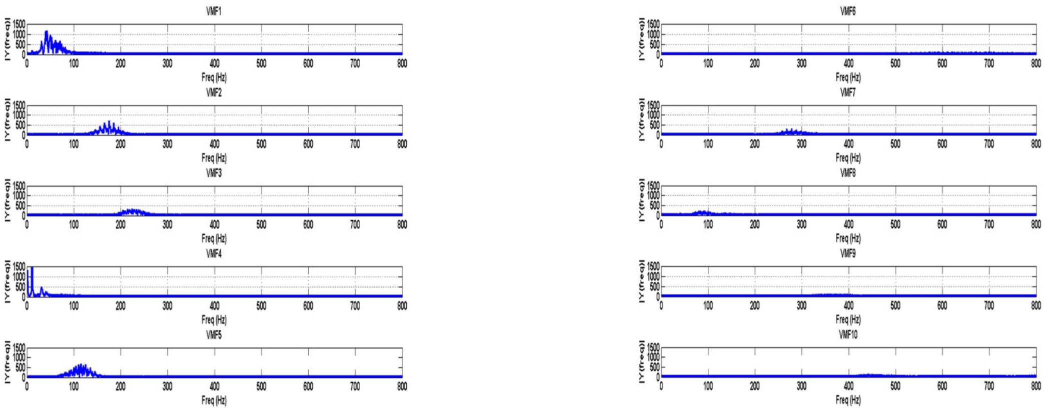

Fig.6: *The spectrum obtained after performing FFT on these decomposed levels.*

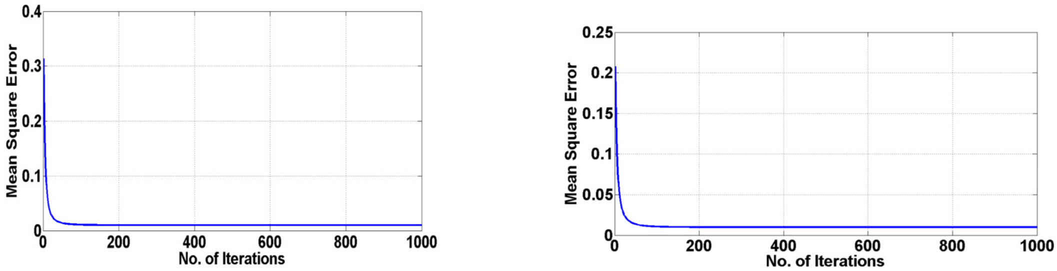

Fig.7: *: The mean square error plot obtained for different estimations. (a) The mean square error plot obtained for online estimation. (b)The mean square error plot obtained for offline estimation.*

Table 1: *Dimensions of the original data.*

Matrix	Size
Input Matrix	1×1280
VMF	10×1280
Conv 1	10×1285
Max-Pool 1	10×159
Conv 2	10×168
Max-Pool 2	10×20
Conv 3	10×29
Max-Pool 3	10×2

4.1.2 Offline Detection

The database created for the offline process was collected from the University College Database (UCD database) from Dublin. The UCD database can be accessed online from the website https://physionet.org. The UCD database contains the PSG ECG collection record of 21 male and 4 female subjects with a duration of 5.9-7.7 hr. These acquired signals are first decomposed into different decomposition levels using the VMD algorithm. These sets of decomposed levels are the input values to the CNN network. In order to find the noise level, a Fast Fourier Transform (FFT) was applied to all of these decomposed levels. The corresponding decomposed levels are represented in Fig. 5 and the spectrum of different decomposition levels is displayed in Fig. 6. Among these different levels, the level which falls in the same frequency spectrum of ECG signals (0-100Hz) was selected for constructing the kernel function. The mean of each of these selected levels was taken and kept as the kernel function. This kernel function matrix with a size of 1×10 was utilized for the convolution operation. The kernel function itself acts as the filter to the input signal. The main advantage of this function is that the noise level in the input signal can be easily separated. In our proposed work, the feature extraction was performed in terms of three convolution and max-pooling operations. In the classification part, 2 MLP layers were used. This architecture gives maximum efficiency while training the real-time input signals. Thus the mean square error was reduced to a minimum of 0.01. In our proposed work, there are two output maps. The number of hidden neurons was set to be 20. It gave the best performance in accordance with the related works [32] [33]. An iterative estimation of weight and biases was performed for reducing the mean square error. In our proposed work, the maximum number of iterations was set to be 1000 and a learning factor of 0.001 was also fixed for getting minimum mean square error. The dimensions of the original data set after each convolution and max-pooling are represented in Table 1. At the final step of the feature extraction phase, the dimension of the feature set was reduced to a minimum 1×2 matrix. Among the acquired data samples, 10 of the samples were taken for training the 1D CNN model and rest of them were utilized for the testing purpose. Fig. 7 represents the mean square error plot obtained for both online and offline estimations.

4.2 Cross Validation

The efficiency of a classifier is cross checked in terms of some statistical analysis called crossvalidation. The cross-validation technique is utilized to examine the classifier generalization capacity. The main aim of this cross-validation is to divide the entire set of data values into different sub-sets. These sets can be utilized for the training and testing of the classifier. In our proposed improved CNN approach, the cross validation was done in a leave-one out manner. This cross validation technique utilizes only a single feature value from the total feature samples as the validation sample and the remainder of the feature set is used as the training set. This process is repeated in such a way that each feature value in the total data sample is utilized once as the validation sample. After the leave one out cross validation, the miss-classification rate and mean square error rate are calculated. It was observed that the miss-classification and mean square error rates were reduced to 0.0154 and 0.01 respectively. These values show the efficiency of the CNN classifier over the other classifiers.

4.3 Performance Parameters

The performance analysis of our proposed system was carried out by utilizing different parameters. The main parameters used for this evaluation are represented by the following equations.

$$Sensitivity = \frac{T_p}{T_p + F_n} \tag{18}$$

$$Specificity = \frac{T_n}{T_n + F_p} \tag{19}$$

$$Accuracy = \frac{T_n + T_p}{T_n + T_p + F_n + F_p} \tag{20}$$

$$Precision = \frac{T_p}{F_p + T_p} \tag{21}$$

$$F1 - score = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \tag{22}$$

$$MCC = \frac{T_p \times T_n - F_p + F_n}{\sqrt{(T_p + F_n)(F_p + T_p)(F_p + T_n)(F_n + T_n)}} \tag{23}$$

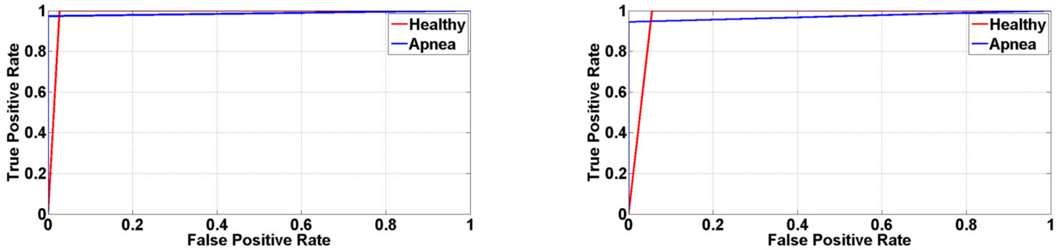

Fig.8: *The ROC curve obtained for different estimations.(a) The ROC curve obtained for online estimation. (b)The ROC curve obtained for offline estimation.*

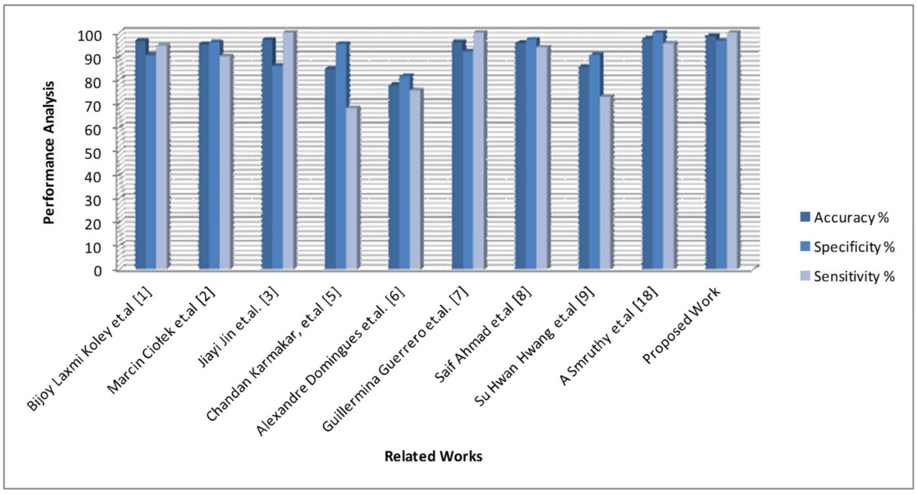

Fig.9: *The performance analysis of related works in terms of different feature extraction and classification techniques.*

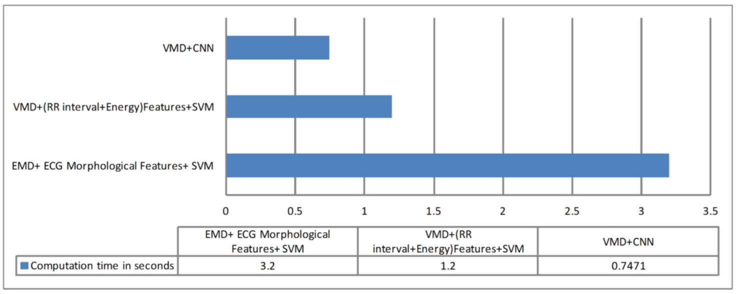

Fig.10: *The performance analysis chart of related works in terms of computational complexity.*

Table 2: *Parameters of PMSM.*

Performance Parameters	Online	Offline
Sensitivity	100	100
Specificity	96.5517	87.5
Accuracy	98	96
PPV	0.97	0.94
NPV	1	1
F1-score	1	1
Precision	1	1
MCC	0.9682	0.9091
AUC	0.9865	0.9722
Computation Time	0.7474 sec	0.6712 sec

In the above mentioned equations, T_p, T_n, F_p, and F_n covey the shorthand notations of True Positive, True Negative, False Positive, and False Negative measurements. The interrelationship between the target class and the response class can be evaluated by utilizing Mathew's Correlation Coefficient (MCC). The range of this value can be specified in between -1, 0, or 1 [34]. These values are related to the worst prediction, random guess, and perfect prediction respectively. The precision and sensitivity values are utilized for the prediction of F-1 score. This F-1 score value varies in the range of 0 and 1. The F-1 score value of 1 describes a system with the best recognition ability. The value 0 describes a system which is not efficient at all. The value of precision describes a proportion which is associated to the healthy patterns estimated and the original healthy patterns. Table 2 shows the various performance values measured in our proposed work.

The efficiency of our proposed work was validated by plotting the Receiver Operating Characteristic Curve (ROC). Fig. 8 represents the ROC curve obtained for different estimations with Area Under the Curve (AUC) values of 0.9865 and 0.9722. The performance of the entire system was compared with the techniques explored in the different feature extraction and classification fields based on the above mentioned performance values. The comparison chart is represented in Fig. 9. The comparison results represent the highly accurate performance of the proposed VMD + CNN technique. An accuracy of 98.4615% was obtained for our proposed work. The computational time and accuracy of our proposed work is high compared to the other related works. The detailed performance chart for the other related works is represented in Fig.10. It can be seen in the results that, the overall computational time of the VMD+CNN technique is about 0.7471 seconds. This is relatively lesser than the other techniques. The computational time and accuracy obtained for Empirical Mode Decomposition (EMD) along with the morphological features with the SVM classification were about 3.2 seconds and 93.33% respectively. The VMD + RR interval and energy features with SVM classification gave a computational time of 1.2 seconds. An accuracy of 97.50% was obtained for this technique. According to all these results it is clear state that the VMD + CNN technique is more accurate with a lower computational burden and reduced dimension. One of the shortcoming of our proposed study is that the classification of the apnea subjects was not studied in terms of mild, moderate, and severe cases. These classifications are collectively called the apnea severity index estimation. It is also important to consider the severity index of apnea for accurate predictions. We will address this limitation of our proposed work in future.

5. CONCLUSION

This paper shows an improved CNN approach for real-time apnea classification by utilizing the concept of variational mode decomposition. The results that the computational time and dimension of the features are reduced after utilizing the improved CNN algorithm. It is also apparent that the classification accuracy is higher after using the CNN model. An accuracy of 98.4375% was achieved after performing the proposed algorithm. As a future work, we will consider calculating the apnea severity index for more accurate prediction.

References

[1] B. L. Koley and D. Dey, "Real-Time Adaptive Apnea and Hypopnea Event Detection Methodology for Portable Sleep Apnea Monitoring Devices," in *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, pp. 3354-3363, Dec. 2013.

[2] M. Ciolek, M. Niedźwiecki, S. Sieklicki, J. Drozdowski and J. Siebert, "Automated detection of sleep apnea and hypopnea events based on robust airflow envelope tracking," *2014 22nd European Signal Processing Conference (EUSIPCO)*, pp. 2090-2094, 2014.

[3] J. Jin and E. Sánchez-Sinencio, "A Home Sleep Apnea Screening Device With Time-Domain Signal Processing and Autonomous Scoring Capability," in *IEEE Transactions on Biomedical Circuits and Systems*, vol. 9, no. 1, pp. 96-104, Feb. 2015.

[4] G. C. Gutiérrez-Tobal, D. Álvarez, F. del Campo and R. Hornero, "Utility of AdaBoost to Detect Sleep Apnea-Hypopnea Syndrome From Single-Channel Airflow," in *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 636-646, March 2016.

[5] C. Karmakar, A. Khandoker, T. Penzel, C. Schöbel and M. Palaniswami, "Detection of Respiratory Arousals Using Photoplethysmography (PPG) Signal in Sleep Apnea Patients," in *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 3, pp. 1065-1073, May 2014.

[6] A. Domingues, T. Paiva and J. M. Sanches, "Sleep and Wakefulness State Detection in Nocturnal Actigraphy Based on Movement Information," in *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 2, pp. 426-434, Feb. 2014.

[7] G. G. Mora, J. M. Kortelainen, E. R. P. Hernández, M. Tenhunen, A. M. Bianchi and M. O. Méndez, "Evaluation of Pressure Bed Sensor for Automatic SAHS Screening," in *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 7, pp. 1935-1943, July 2015.

[8] S. Ahmad, I. Batkin, O. Kelly, H. R. Dajani, M. Bolic and V. Groza, "Multiparameter Physiological Analysis in Obstructive Sleep Apnea Simulated With Mueller Maneuver," in *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 10, pp. 2751-2762, Oct. 2013.

[9] S. H. Hwang et al., "Unconstrained Sleep Apnea Monitoring Using Polyvinylidene Fluoride Film-Based Sensor," in *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 7, pp. 2125-2134, July 2014.

[10] T. S. Kumar and V. Kanhangad, "Gabor Filter-Based One-Dimensional Local Phase Descriptors for Obstructive Sleep Apnea Detection Using Single-Lead ECG," in *IEEE Sensors Letters*, vol. 2, no. 1, pp. 1-4, March 2018.

[11] A. Zarei and B. M. Asl, "Automatic Detection of Obstructive Sleep Apnea Using Wavelet Transform and Entropy-Based Features From Single-Lead ECG Signal," in *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1011-1021, May 2019.

[12] H. Korkalainen et al., "Accurate Deep Learning-Based Sleep Staging in a Clinical Population With Suspected Obstructive Sleep Apnea," in *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 7, pp. 2073-2081, July 2020.

[13] C. Sun, J. Fan, C. Chen, W. Li and W. Chen, "A Two-Stage Neural Network for Sleep Stage Classification Based on Feature Learning, Sequence Learning, and Data Augmentation," in IEEE Access, vol. 7, pp. 109386-109397, 2019.

[14] P. Y. Gumery et al., "An adaptive detector of genioglossus EMG reflex using Berkner transform for time latency measurement in OSA pathophysiological studies," in *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 8, pp. 1382-1389, Aug. 2005.

[15] G. C. Gutiérrez-Tobal, D. Álvarez, A. Crespo, F. del Campo and R. Hornero, "Evaluation of Machine-Learning Approaches to Estimate Sleep Apnea Severity From At-Home Oximetry Recordings," in *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, pp. 882-892, March 2019.

[16] V. P. Tran and A. A. Al-Jumaily, "A Novel Oxygen-Hemoglobin Model for Non-Contact Sleep Monitoring of Oxygen Saturation," in *IEEE Sensors Journal*, vol. 19, no. 24, pp. 12325-12332, 15 Dec.15, 2019.

[17] K. Dragomiretskiy and D.Zosso, "Variational mode decomposition," *IEEE transactions on signal processing*, vol. 62, No. 3,pp.531–544, 2013.

[18] K. Dragomiretskiy and D. Zosso, "Two dimensional variational mode decomposition," *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 197–208, 2015.

[19] J. Salamon and J. P. Bello, "Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification," in *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 279-283, March 2017.

[20] N. Kalchbrenner, E. Grefenstette and P. Blunsom, "A convolutional neural network for modelling sentences," *arXiv preprint arXiv:1404.2188*, 2014.

[21] Y. Chen, H. Jiang, C. Li, X. Jia and P. Ghamisi,"Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6232-6251, Oct. 2016.

[22] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen and C. I. Sánchez, "Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection in Color Fundus Images," in *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1273-1284, May 2016.

[23] S. Lekha and S. M., "Real-Time Non-Invasive Detection and Classification of Diabetes Using Modified Convolution Neural Network," in *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1630-1636, Sept. 2018.

[24] S. Lekha and M. Suchetha, "A Novel 1-D Convolution Neural Network With SVM Architecture for Real-Time Detection Applications," in *IEEE Sensors Journal*, vol. 18, no. 2, pp. 724-731, 15 Jan.15, 2018.

[25] I. De Falco, G. De Pietro, A. Della Cioppa, G. Sannino, U. Scafuri and E. Tarantino, "Evolution-based configuration optimization of a deep neural network for the classification of obstructive sleep apnea episodes," Future Generation Computer Systems, vol. 98, pp. 377–391, 2019.

[26] K. Chen, C. Zhang, J. Ma, G. Wang and J. Zhang, "Sleep staging from single-channel eeg with multi-scale feature and contextual information," *Sleep and Breathing*, vol. 23, No. 4, pp.1159–1167, 2019.

[27] F.-T. Wang, M.-H. Hsu, S.-C. Fang, L.- L. Chuang and H.-L. Chan, "The respiratory fluctuation index: A global metric of nasal airflow or thoracoabdominal wall movement time series to diagnose obstructive sleep apnea," *Biomedical Signal Processing and Control*, vol. 49, pp.250–262, 2019.

[28] C. S. Viswabhargav, R. Tripathy and U. R. Acharya, "Automated detection of sleep apnea using sparse residual entropy features with various dictionaries extracted from heart rate and edr signals," *Computers in biology and medicine*, vol. 108, pp.20–30, 2019.

[29] N. Banluesombatkul, T. Rakthanmanon and T. Wilaiprasitporn, "Single Channel ECG for Obstructive Sleep Apnea Severity Detection Using a Deep Learning Approach," *TENCON 2018 - 2018 IEEE Region 10 Conference*, pp. 2011-2016, 2018.

[30] T. Yang, L. Yu, Q.Jin, L. Wu and B. He, "Localization of Origins of Premature Ventricular Contraction by Means of Convolutional Neural Network from 12-lead ECG,"*IEEE transactions on biomedical engineering*, vol.65, No.7, pp.1662–1671, 2017.

[31] L. Yu, Y. Guo, Y. Wang, J. Yu and P. Chen, "Segmentation of Fetal Left Ventricle in Echocardiographic Sequences Based on Dynamic Convolutional Neural Networks," in *IEEE Transactions on Biomedical Engineering*, vol. 64,

no. 8, pp. 1886-1895, Aug. 2017.

[32] T. Ince, S. Kiranyaz, L. Eren, M. Askar and M. Gabbouj, "Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks," in *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067-7075, Nov. 2016.

[33] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE transactions on biomedical engineering*, vol.63, No.3, pp.664–675, 2015.

[34] V. Vijayakumar, M. Case, S. Shirinpour and B. He, "Quantifying and Characterizing Tonic Thermal Pain Across Subjects From EEG Data Using Random Forest Models," in *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 12, pp. 2988-2996, Dec. 2017.

Smruthy A received the B.Tech. degree in electronics and communication engineering from the Vidya Academy of Science and Technology, Calicut University, Thrissur, India, in 2011, and the M.Tech. degree from Karunya University, Coimbatore, India, in 2013. She is currently pursuing the Ph.D. degree in electronics and communication engineering with the Vellore Institute of Technology, Chennai Campus, Chennai, India.

Suchetha M received Ph.D. in Biomedical with the thesis title Empirical Mode Decomposition based denoising and classification techniques applied to Electrocardiogram signals. She is currently Deputy Director, Centre for Healthcare advancements, Innovation and Research in VIT University. Her areas of interest are wearable devices, Signal/ image processing techniques in biomedical, AI and Big data analytics for healthcare, developing non-invasive devices. She is the Principal Investigator for a funded project by ISRO and other funded projects. She has published more than 70 papers in journals, authored 4 books, and filed several patents. She is a life member of ISTE, Senior member of IEEE Engineering in Medicine and Biology Society (EMBS) and Faculty advisor of IEEE Robotics and Automation society.