Chapter 3

Removal of Artifacts from EEG Signals

3.1 Introduction

EEG records the electrical activity occurring within the brain and facilitates the detection of any abnormal activity that might lead to neurological disorders. Whenever an EEG recording is taken, it tends to get contaminated with various artifacts. These artifacts, also called noise, can be of various types. Both physiological and non-physiological artifacts can interfere with accurate readings during EEG tests. These can take various forms, such as ECG/CA, EMG/MA, EOG/OA, and background noise from the instruments used in the test [42–44, 47]. The patient might move their limbs or blink during the examination, or it could be their heartbeat rhythm, which might be found as an artifact in an EEG recording [34, 35, 41]. The emergence of portable systems that utilize SC-EEG has revolutionized real-time medical observation systems and BCI, especially in non-clinical and indoor settings. The artifacts mentioned above can easily become a part of the EEG signals being recorded in such settings [25].

Traditional techniques like adaptive filters can be used for AR, but they require an additional artifact reference channel, which adds to the hardware overhead [25] or needs initial calibration. Techniques such as wavelet decomposition [146], EMD [63], and VMD [43] do not necessitate the presence of either an artifact reference channel or an initial calibration process [25]. Various tasks involving pattern classification of EEG signals, including BCI, identifying seizures in epilepsy patients, detecting Alzheimer's, biometrics, and emotion recognition, require the elimination of the MoA as a prerequisite [49–51, 147–153]. Some works have acknowledged that compared to artifacts associated with eye movement and heart function, artifacts caused by muscle contractions are notably harder to eliminate. The primary challenge in addressing EMG artifacts arises because they exhibit a wide range of frequencies, are distributed across various locations in the scalp, and lack a consistent pattern or structure [56]. Another work proposed a methodology that integrates the signal's characteristics in both the time and frequency domains to remove OAs [60]. It utilizes a DL network and can be implemented with several other denoising methods based on DL techniques.

Based on the literature, the existing studies on AR from EEG signals are quite efficient. However, only a few studies have provided methodologies for EEG signals in epileptic patients. Apart from this, the variety of artifacts these techniques handle is also limited, as some cater to only individual artifact types. Further, it is almost always theoretically implied by most of these works that denoising will eventually improve the quality of signals and will become fit for classification or prediction purposes. Therefore, this research aims to develop a signal-denoising model to remove artifacts that may occur in EEG signals derived from epileptic patients. Initially, noisy data is created by injecting noise into the original EEG signal. After this, STFT is applied to the noisy EEG signal to obtain the signal's time-frequency representation. The transformed signal is then sent to the novel proposed denoising architecture, which is developed using a BLSTM, and this architecture is trained using the learning mechanism of a BSCN [154]. This architecture, called Bidirectional Stochastic LSTM (BS-LSTM), has the advantages of both the abovementioned techniques. The framework of a BLSTM enables learning of complex patterns in the past timesteps and the future time steps of the EEG signal. Along with this, the learning based on BSCN brings in the benefit of the non-iterative nature of training the classifier, making the training more efficient.

The main contribution of this research work is the proposed novel architecture that combines the framework of a BLSTM and the learning mechanism of a BSCN. This novel architecture is utilized to denoise EEG signals in epilepsy patients. The proposed methodology can handle multiple types of noises, like ECG, EMG, EOG, and PL, individually as well as in combination. Through experimentation, it is reported that the novel technique efficiently improves ES's classification and prediction accuracy. Further, the proposed technique has also been used on a sleep data set called CAP Sleep to prove the flexibility of domain adaptation.

3.2 Preliminaries

3.2.1 Short Term Fourier Transform

The STFT is a method of processing signals that includes analyzing small sections of the signal using the Fourier transform and moving the analysis window across the entire signal. By doing this, it is possible to analyze how the signal's frequency content changes over time, providing a time-frequency analysis. EEG signals are inherently unstable, so shorter analysis durations with STFT can improve preprocessing results. Compared to the Fourier transform, the STFT has superior properties of localizing in both time and frequency domains [155, 156].

3.2.2 Bidirectional Long Short Term Memory

A BLSTM network is commonly used for processing sequential data, such as time series data or signals with temporal dependencies. It uses two hidden states, one representing the forward and the other representing the backward sequences. These two hidden states are concatenated at each time step to form the final output, allowing it to capture dependencies in past and future contexts [157]. The BLSTM network has been widely used in ES classification and prediction using patients' EEG signals. The network can learn to capture temporal dependencies and patterns in the EEG signals, accurately classifying the signals into different seizure states. By analyzing the patterns and trends in the EEG signals leading up to seizure events, the network can provide an early warning system for epileptic patients, allowing them to take preventive measures. The BLSTM network's ability to process sequential data and capture temporal dependencies in the forward and backward directions in EEG signals effectively makes it well-suited for seizure classification and prediction tasks. Additionally, the network can be trained on large data sets of EEG signals from epileptic patients, enabling it to learn complex patterns and improve its performance over time. Overall, the BLSTM network has shown promise in classifying and predicting ESs using EEG signal data, offering potential benefits for epilepsy diagnosis, treatment, and management.

3.2.3 Bidirectrional Stochastic Configuration Network

BSCNs are derived from Stochastic Configuration Networks (SCNs), which are randomized NNs capable of conducting effective training without GPU. An SCN is generated incrementally by stochastic configuration algorithms. The parameters of the nodes in the hidden layer are randomly assigned using an initialization constraint, and the output weights are analytically

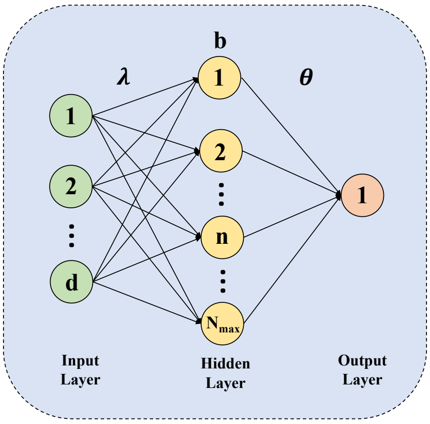


Figure 3.1: The network structure of the Bidirectional Stochastic Configuration Network model [7].

evaluated in either a constructive or selective manner. The proposed SCNs have been shown to have the universal approximation property and can be used for data regression and classification problems. SCNs are known to reduce the overhead of parameter tuning by automatically deciding the number of nodes to be added to the hidden layer, done by using an error threshold value [7]. Using the initialization constraint makes SCN more stable compared to other randomized algorithms, but this mechanism of model training is quite time-consuming.

To understand BSCN better, the mathematical formulation of an SCN is explained along with its training mechanism. In Figure 3.1, d is the dimension of the input data, λ refers to the input weights, b refers to the biases of the hidden layer, n refers to the number of nodes in the hidden layer of the current network, N_{max} refers to the maximum number of nodes that can be added to the network and θ refers to the output weights.

For a given data set, each sample can be represented as (X_i, Y_i), where $i = 1, 2, ...M$ and M is the number of samples in the training data set. The definition of an SCN is as follows:

$$g_n(X) = \sum_{j=1}^{n} \theta_j * a(\lambda_j \cdot X + b_j) \tag{3.1}$$

where $g_n(X)$ is the target function, $\lambda_j = [\lambda_{j1}, \lambda_{j2}, ...\lambda_{jd}]$, $\theta_j = [\theta_{j1}, \theta_{j2}, ...\theta_{jp}]$, and $a(\cdot)$ is the activation function. Also, p is the number of classes in a classification problem; this value is one for regression problems.

An error computed from the current network is given as:

$$r_n = g_{n-1} - g_n = [r_{n1}, r_{n2}, ..., r_{np}] \tag{3.2}$$

where g_{n-1} is the error derived from the network when the hidden layer had $n-1$ nodes. Assuming that r_n has not reached the acceptable error threshold ε, a new node is added to the SCN. The new SCN is represented as:

$$g_{n+1} = g_n + \theta_{n+1} a(\lambda_{n+1} \cdot X + b_{n+1}) \tag{3.3}$$

Here, λ_{n+1}, b_{n+1} and θ_{n+1} were assigned randomly, but their values need to adhere to the following condition, also known as the initialization constraint, represented as follows :

$$\sum_{k=1}^{p} (r_{nk}, a_{n+1})^2 \geq ||a_{n+1}||^2 \rho_{n+1} \tag{3.4}$$

$$a_{n+1} = a(\lambda_{n+1} \cdot X + b_{n+1}) \tag{3.5}$$

$$\rho_{n+1} = (1 - q - \gamma)||r_n||^2 \tag{3.6}$$

$$\gamma_{n+1} = \frac{1-q}{n+1}, 0 < q < 1 \tag{3.7}$$

The output weights were calculated as follows:

$$\theta_{n+1,k} = \frac{\langle r_{nk}, a_{n+1} \rangle}{||a_{n+1}||^2} \tag{3.8}$$

Here $k = 1, 2, \cdots, p$. After this, the error is calculated again for the network as $r_{n+1} = r_n - \theta_{n+1} a(\lambda_{n+1} \cdot X + b_{n+1})$. This process, from Eq. 2 to 5, is repeated until ε / N_{max} is reached.

Many applications utilize SCN and its variants, but they have one drawback. When a new node is added to the hidden layer, several candidates for the input parameters are created, and the best parameters must be selected according to certain preset conditions. This process of

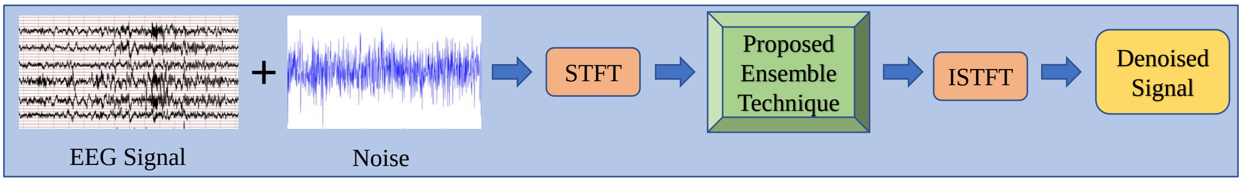


Figure 3.2: The proposed methodology to remove artifacts from EEG signals.

training the models takes time. To overcome such drawbacks, BSCN performs learning in two modes: i) forward and ii) backward. The BSCN uses the parameter initialization method used in SCN in the forward learning mode. In contrast, in the backward learning mode, the residual error feedback of the existing model is used for one-time-based parameter initialization. These types of learnings were executed one after the other until ε/N_{max} is reached. The semi-random learning technique of BSCN enhances its training ability while improving the quality of the nodes in the hidden layer. BSCN, therefore, has better generalization ability than SCN. It is faster with respect to training speed and has higher stability [154].

3.3 Proposed Methodology

Denoising is an essential preprocessing step that provides an artifact-free signal. This denoised signal can further classify and predict spike events like ES. Specific settings for EEG signal acquisition are more susceptible to artifacts, such as an ambulatory EEG recording. An extended recording period of typically 24 to 72 hours allows for detecting infrequent or intermittent abnormalities that may not be captured in a traditional, shorter-duration EEG recording. It records the brain's electrical activity using these monitoring systems even when the subject is mobile, making them practical for diagnosing neurological disorders [58, 158]. Various steps comprising the proposed method have been described in this section and are illustrated in Figure 3.2.

3.3.1 Creation of Noisy Data

The databases utilized for experimentation in this research work are mostly clean data sets. Thus, various noise types have been injected into the data using a fundamental formulation. The noises used were mainly biomedical signals, i.e., ECG, EOG, EMG, and PL. The reason for this is that in many clinical settings, various tests are performed along with EEG tests. The proposed work has also been tested for various combinations of these artifacts. The mathematical formulation for the creation of a noisy EEG signal is as follows:

$$Noisy_{EEG} = Original_{EEG} + \alpha \times Noise \tag{3.9}$$

Here, α is the contamination coefficient [63].

After the addition of noise, the data is transformed using STFT. It is commonly used in signal processing, including the analysis of EEG signals. It is an extensively used tool for analyzing the time-varying frequency content of signals, allowing researchers and clinicians to observe the fluctuations or variations in the brain's electrical or neural activity over time associated with different cognitive processes, brain states, or neurological conditions [159]. STFT allows for flexibility in window selection, which can impact the trade-off between time and frequency localization. Different window functions can be used to adjust the shape and size of the time window, enabling customization based on the characteristics of the EEG signal and the specific denoising requirements. A Hamming window has been used for the proposed work. This window has been selected as it improves the frequency representation of the signal by applying a taper to the ends [160–163].

3.3.2 Bidirectional Stochastic LSTM Denoising Method

The proposed method has the advantages of a BLSTM and a BSCN, a semi-random constructive algorithm. BSCN naturally inherits the universal approximation property and accelerates training efficiency. The BLSTM processes the frequency and time resolution information derived from the signal.

Let $n-1$ be the number of hidden nodes currently in the network, and a new n^{th} node is added. The next step is to identify the order of the node, which will further help in deciding the type of learning that will be performed. If the type of learning is represented as l, then identifying the node order can be defined as :

$$l = \begin{cases} n \in \{2t+1, t \in Z\}, & \text{order is odd} \\ n \in \{2t, t \in Z\}, & \text{order is even} \end{cases}$$

When the order is odd, $l = 0$, forward learning is performed. In this, the parameters are assigned in the same manner as that in an SCN. S_{max} pairs of (λ, b) are generated from a symmetric interval $[-\upsilon, \upsilon]$, according to the inequality in Eq. 4, and corresponding θ are generated from Eq. 5. The pair which gives a maximum reduction in r_n is selected. An updation of r_n is formulated as follows:

$$r_n(X) = r_{l-1}(X) - \theta_n a(\lambda_n \cdot X + b_n) \tag{3.10}$$

If ε/N_{max} is reached, the model training is terminated, or a new node is added to the network.

When the order is even, $l = 1$, and backward learning is performed. In this, the (λ, b) are generated according to the residual error feedback of the existing network and are calculated as follows:

$$G_n{}^r = r_{n-1} \cdot (\theta_{n-1})^{-1} \tag{3.11}$$

$$\lambda_n = a^{-1}(\mu(G_n{}^r)) \cdot X^{-1} \tag{3.12}$$

$$b_n = rmse(a^{-1}(\mu(G_n{}^r)) - \lambda_n \cdot X^{-1}) \tag{3.13}$$

Eq. 3.11 calculates the error feedback function sequence. θ of the new node is calculated as follows:

$$\theta_{n,k} = \frac{\langle r_{n-1}, G_n \rangle}{||G_n||^2}, k = 1, 2, ..., p \tag{3.14}$$

$$G_n = \mu^{-1}(a(\lambda_n \cdot X + b_n)) \tag{3.15}$$

After this, just as in forward learning, r is updated, and then if the termination conditions are met, the training is terminated.

The learning strategies used by BSCN are utilized to train the classifier. But, BSCNs are shallow networks and cannot be used for large data sets [154]. Utilizing BLSTM helps handle the complexity of modeling problems as the network can be improvised to increase the complexity of the model by removing any restrictions on adding hidden layers and the number of nodes in those layers. This flexibility makes the model efficient enough to handle the complexities of real-life problems. The task of denoising EEG signals needs a complex network. Thus, for implementing the BS-LSTM, the incremental nature of the BSCN was not considered; instead, the number of nodes present in the bidirectional layer was fixed. The training was performed for N_{max} iterations or until the ε is reached. The parameters of the existing network were updated according to both types of learning, i.e., forward and backward, and alternatively, until the termination condition was reached .The denoised signal is then transformed into the time domain by applying inverse STFT.

3.4 Result Analysis

In this section, the experimental setup and details are discussed. Apart from this, the experimental results of the above data sets and their detailed analysis have been presented. Also, for the exposition of implementation flexibility in domains other than epilepsy, a sleep data set was used. All the experiments have been developed according to the settings in [46]. Fast-ICA [164], MOFPA-WT [165], WDESN [66], VMD-DWT-WPT [70], RNN-LSTM, and BLSTM were the state-of-the-art techniques that have been used to compare the performance of the proposed novel BS-LSTM keeping Fast-ICA as the baseline. The aim is to develop a patient-generic AR system for EEG signals. For this purpose, all the patients in the data sets were considered to conduct this study.

3.4.1 Evaluation Based on Denoising Task

In this research, to perform denoising using the proposed BS-LSTM, noisy data is created according to Eq. 3.9. While creating noisy signals, an SNR of 10 dB was considered. Table 3.1 represents the artifacts and their Artifact Combinations (ACs). When the signal is injected with a Single Artifact (SA), one of AC1, AC2, AC3, or AC4 is used to create a noisy EEG signal. If the signal is injected with Artifacts in Pairs (AiP), then AC5, AC6, AC7, AC8, AC9, or AC10 were used to create noisy data. Similarly, when a Combination of Three Artifacts (CoTA) was injected into the EEG signal, AC11, AC12, AC13, or AC14 were used. In cases where all four artifacts were used to create the noisy data, AC15 is used. Only the additive inclusion of artifacts in the EEG signal has been considered for simplification [146]; although, some works suggest the removal of multiplicative artifact interference as well [166]. The value of α in Eq. 3.9 changes depending on the type of artifact or combination, i.e., SA, AiP, or CoTA, affecting

Table 3.1: Representation of artifacts and their combinations in the experimental results.

Symbol	Group	Artifacts	Symbol	Group	Artifacts
SA	AC1	ECG	AiP	AC9	EMG, PL
	AC2	EMG		AC10	EOG, PL
	AC3	EOG	CoTA	AC11	ECG, EMG, EOG
	AC4	PL		AC12	ECG, EOG, PL
AiP	AC5	ECG, EMG		AC13	ECG, EMG, PL
	AC6	ECG, EOG		AC14	EMG, EOG, PL
	AC7	ECG, PL		AC15	ECG, EOG, EMG, PL
	AC8	EMG, EOG			

the signal. Also, in cases of AiP and CoTA, if ECG is affecting the signal, then it is used as $0.5\times$ ECG [167] to create a noisy EEG signal.

Implementing the proposed novel BS-LSTM required specific parameters, such as MSE, which was the choice of the error function to calculate r at each iteration of learning performed. The number of units in the architecture was 64, dropout was taken as 30.0%, the number of epochs was taken as 100, and the learning rate taken was 0.001.

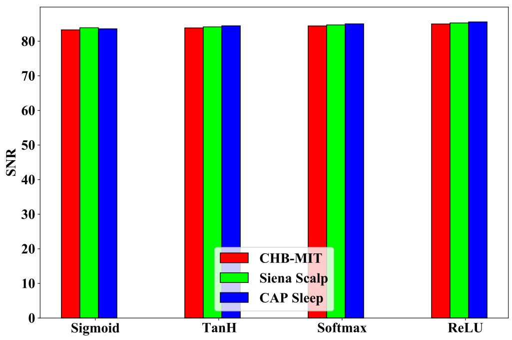
The performance metrics used to assess the performance of various methods and the proposed technique to perform denoising on EEG signals are SNR, Percentage Root Mean Square Difference (PRD), and MSE. These metrics are defined as follows :

$$SNR(dB) = 10 \cdot \log_{10} \frac{Var(D)}{Var(C)} \tag{3.16}$$

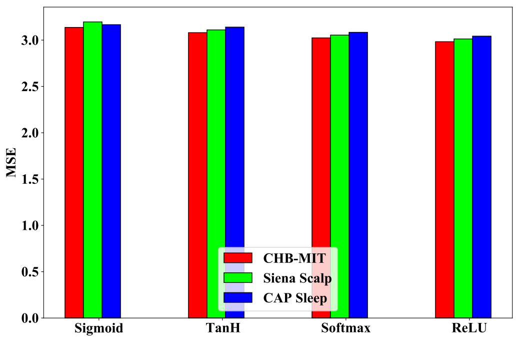
$$PRD(\%) = \left[\sqrt{\frac{\sum_{i=1}^{N} [d_i - c_i]^2}{\sum_{i=1}^{N} {c_i}^2}} \right] \times 100 \tag{3.17}$$

$$MSE = \frac{1}{N} \sum_{i=1}^{N} [d_i - c_i]^2 \tag{3.18}$$

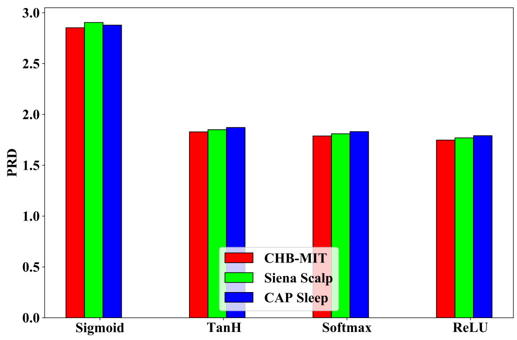
Here D is the denoised signal, C is the contaminated signal, N is the number of data points in the signal, d is an element of the signal D, and c is an element of the signal C [53].



(a)



(b)



(c)

Figure 3.3: The performance of various activation functions concerning (a) SNR, (b) MSE, and (c) PRD on DB1, DB2, and DB5 when used with the proposed BS-LSTM. These experiments were performed to remove ECG noise.

3.4.1.1 Result Analysis For Selection of the Activation Function

The inequality in Eq. 3.4 dictates the initialization of model parameters. The activation function is an essential governing factor in this inequality. Experiments were performed to assess the choice of activation function with sigmoid, tanh, softmax, and ReLU. Figure 3.3 illustrates the experimental results of utilizing various activation functions with the proposed BS-LSTM network to denoise the EEG signal of DB1, DB2, and DB5 when it gets contaminated with ECG noise.

The results depict that ReLU is the best choice for the activation function as it gives improved results compared to sigmoid, tanh, and softmax across SNR, MSE, and PRD. Thus, further experimentation was done using ReLU. MSE was used for r calculation, N_{max} was taken as 64, and $\varepsilon = 0.001$. Also, for simplification purposes, the artifacts and their combinations are here on represented as given in Table 3.1.

3.4.1.2 Result Analysis For DB1

The results in Table 3.2 show that for DB1, the proposed BS-LSTM performs the best. The proposed technique shows an improvement of approximately 5 dB in terms of SNR. This value is the maximum for AC2, i.e., 4.99 dB. However, the maximum SNR achieved is after the elimination of AC1 using the proposed BS-LSTM denoising technique. When multiple artifacts were combined, it can be observed that the proposed technique improves upon the baseline in terms of SNR by 6.00% to 7.10%. When the signal has AiP, the maximum improvement of 6.86% over the baseline was observed for BS-LSTM to remove AC5. BS-LSTM attains the highest SNR of 62.26 dB for the removal of AC9. In the case of CoTA, the highest SNR of 48.65 dB was attained by the proposed technique to remove the AC14, giving the maximum improvement of 6.17% on the baseline. Further, it was observed that the performance of BLSTM was enhanced by performing the learning based on BSCN. It can be observed that BS-LSTM shows an improvement of 0.31% to 4.12% over BLSTM. The maximum improvement of 7.10% over baseline is seen when the technique removes AC15 with an SNR of 25.418 dB. The slightest improvement of 0.62% over baseline is seen when the EEG signals were affected by AC12. However, the SNR falls substantially when AC15 affects the data. For the proposed technique, SNR was reduced by 69.10% when AC15 affected the signal, which shows that these artifacts were not entirely eliminated. Regarding SNR, the proposed BS-LSTM gives maximum improvement when used for EMG noise compared to BLSTM.

MSE has also improved in the range of 0.657 to 1.159, with the most significant improve-

Table 3.2: performance of various techniques when used to perform denoising of EEG signals of patients in DB1.

	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM
	SNR							MSE							PRD						
AC1	81.1815	81.9717	82.2512	82.1469	82.9286	83.9792	85.6644	4.2341	4.0725	3.7511	4.1557	3.8855	3.6899	2.0502	5.6930	5.0525	5.3172	4.6929	4.8983	4.1261	2.5823
AC2	79.5719	80.4304	81.2098	81.3887	82.4683	82.2221	84.5653	3.5807	3.6340	3.9158	3.3878	3.7043	3.8176	2.1912	5.3757	5.2952	4.8637	5.1827	4.9160	5.0558	4.3598
AC3	73.5356	73.8433	74.1597	74.8636	75.9244	75.8056	78.1279	3.3506	3.3080	2.8904	3.0384	2.4351	3.1762	2.9210	5.0242	5.1844	4.5000	4.9209	4.8141	4.9243	4.4417
AC4	76.0084	76.4961	77.0953	78.1424	78.3011	79.1567	80.5583	3.2579	3.4353	3.0797	3.6712	2.9769	2.6948	2.5144	5.3153	4.9588	4.8273	4.8099	4.8078	4.6891	4.4958
AC5	51.3632	51.9055	52.0665	52.4712	54.0040	54.5039	54.8885	13.5930	15.6500	9.1603	8.9446	8.5579	10.8841	5.6010	9.6021	9.2469	8.2258	8.8412	8.6195	7.8282	6.7573
AC6	50.1559	50.6222	50.6565	51.1538	52.6990	53.2164	53.5797	13.3629	15.3239	8.1349	8.5952	7.2886	10.2427	6.3308	9.2506	9.1360	7.8620	8.5793	8.5176	7.6967	6.8392
AC7	50.6505	51.2924	51.2436	51.6843	53.3548	53.7025	54.0550	13.2701	15.4512	8.3242	9.2280	7.8304	9.7614	5.9242	9.5416	8.9104	8.1893	8.4684	8.5113	7.4616	6.8934
AC8	57.9522	58.6733	58.6687	59.0427	60.7620	61.5631	61.7805	14.8266	16.9217	10.1752	9.9052	9.0502	12.2154	7.4969	11.7798	11.9050	10.0672	11.4156	10.9844	10.6895	9.9078
AC9	58.4467	59.2604	59.3389	59.5732	61.4178	62.0492	62.2558	14.7338	17.0489	10.3644	10.5380	9.5920	11.7340	7.0903	12.0709	11.6794	10.3945	11.3047	10.9782	10.4543	9.9620
AC10	57.2395	57.8504	58.0556	58.2558	60.1128	60.7617	60.9470	14.5037	16.7229	9.3391	10.1885	8.3227	11.0926	7.8202	11.7194	11.5685	10.0307	11.0428	10.8762	10.3228	10.0439
AC11	38.7397	39.0519	39.2990	39.4651	39.9714	40.0035	41.1051	16.9436	18.9579	12.0508	11.9830	10.9929	14.0603	8.5220	14.6263	14.4313	12.7257	13.7620	13.4336	12.7525	11.1990
AC12	38.0270	38.2651	38.4761	38.8159	39.1380	39.3904	40.3037	16.6208	18.7592	11.2146	12.2664	10.2655	12.9376	8.8453	14.5658	14.0948	12.6892	13.3892	13.3254	12.3858	11.3350
AC13	39.2342	39.5825	39.8861	40.1209	40.4467	40.6737	41.5912	16.8508	19.0852	12.2400	12.6158	11.5347	13.5790	8.1154	14.9174	14.2056	13.0531	13.6511	13.4273	12.5173	11.2531
AC14	45.8232	46.1540	46.4930	46.8789	47.3388	47.4369	48.6503	18.0844	20.3569	13.2549	13.5763	12.0271	14.9102	10.0113	17.0951	16.8637	14.8944	16.2255	15.7922	15.3786	14.4037
AC15	24.7123	25.2844	25.0615	25.3427	25.8666	25.4187	26.4676	26.8681	23.4386	22.4699	25.1123	23.5617	23.2817	21.5620	26.5343	28.3295	25.7529	24.6878	22.9192	22.9139	22.7719

Table 3.3: performance of various techniques when used to perform denoising of EEG signals of patients in DB2.

	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM
	SNR							MSE							PRD						
AC1	81.2050	82.0364	81.8548	82.6629	83.6932	84.0576	86.7844	4.6165	4.5511	3.8930	4.8249	4.0629	4.1725	2.9220	6.1298	4.9678	5.0102	5.3325	4.9156	4.3034	2.3899
AC2	79.7077	80.2712	80.8809	81.9427	82.5235	82.7599	84.8519	4.2886	3.9283	4.5341	4.6628	4.5385	4.0916	3.4729	5.3522	5.8111	4.9461	5.1125	4.7270	4.7272	4.6161
AC3	73.6228	74.3549	75.0563	75.8076	75.7051	76.4457	79.2989	3.9228	3.8352	4.1250	3.6841	3.6800	3.4545	3.9560	5.1024	4.4047	4.4918	4.5400	4.4993	4.7478	3.8968
AC4	75.8288	76.5018	77.6717	78.0203	78.5450	79.0652	81.7922	4.1929	3.8890	4.1730	3.4305	3.6828	3.2528	3.1096	4.6348	4.6970	5.3254	4.5822	5.1876	5.1155	4.2662
AC5	51.8196	52.9949	53.8560	54.3542	54.8219	53.4752	53.8589	7.5354	7.6675	7.8819	8.4450	7.8581	7.2858	5.5184	9.6468	9.3467	8.0437	8.9787	8.2849	7.4073	6.5163
AC6	50.6026	51.8116	52.6911	53.1271	53.4582	52.2123	52.7483	7.1697	7.5745	7.4727	7.4663	6.9995	6.6488	6.0014	9.3969	7.9403	7.5894	8.4062	8.0573	7.4279	5.7970
AC7	51.0438	52.2410	53.2142	53.5697	54.0272	52.7362	53.2470	7.4398	7.6283	7.5208	7.2127	7.0024	6.4470	5.1550	8.9293	8.2325	8.4230	8.4484	8.7456	7.7956	6.1664
AC8	58.4236	59.6622	60.6818	61.2494	61.5936	60.3586	61.0402	9.1500	9.2272	10.0603	9.7166	9.5066	8.6541	8.0133	11.6842	11.2675	10.0304	10.8524	10.3265	10.0034	9.2182
AC9	58.8648	60.0916	61.2049	61.6919	62.1625	60.8825	61.5389	9.4201	9.2810	10.1084	9.4630	9.5095	8.4523	7.1669	11.2167	11.5597	10.8640	10.8946	11.0147	10.3711	9.5875
AC10	57.6478	58.9083	60.0400	60.4649	60.7988	59.6190	60.4283	9.0544	9.1880	9.6992	8.4843	8.6509	7.8153	7.6500	10.9668	10.1534	10.4097	10.3221	10.7871	10.3916	8.8682
AC11	38.7866	39.1289	39.3729	39.8163	40.0150	40.2469	41.5086	11.4583	11.5028	12.0068	12.1291	11.5380	10.7404	9.4744	14.7491	13.7514	12.5355	13.5187	12.7842	12.1551	10.4131
AC12	38.0108	38.3750	38.7311	39.0319	39.2203	39.5080	40.8967	11.3626	11.4635	11.6458	10.8968	10.6824	9.9015	9.1110	14.0317	12.6373	12.9148	12.9884	13.2449	12.5433	10.0632
AC13	39.2278	39.5582	39.8960	40.2589	40.5840	40.7708	42.0073	11.7283	11.5565	12.0549	11.8755	11.5409	10.5386	8.6279	14.2816	14.0437	13.3692	13.5609	13.4725	12.5228	10.7825
AC14	45.8319	46.2256	46.7218	47.1541	47.3557	47.6542	49.1886	13.3429	13.1162	14.2333	13.1471	13.1894	11.9069	11.1229	16.3191	15.9645	15.3558	15.4346	15.5140	15.1189	13.4843
AC15	23.8980	23.9597	25.4412	27.5581	27.9499	26.9304	29.4472	26.9876	24.0000	27.8165	28.0595	27.8651	19.1021	18.7575	26.6482	26.9554	25.9578	26.2938	24.5610	22.3485	22.2724

ment being for AC1. When AC15 was considered for creating noise data, the MSE increased more than five times when the signal was contaminated by just one type of artifact. Further, if the EEG signal is contaminated with just one type of artifact, the proposed technique gives the most significant improvement of 51.58% over baseline while removing AC1. However, BS-LSTM gave a minor improvement of 12.82% while removing AC3, which was the overall slightest improvement in MSE in DB1. When the signal has AiP, the best performance of BS-LSTM was for AC5. The performance fell for AC10, increasing the MSE to 7.82. However, this was an improvement of 46.08% over baseline, whereas maximum improvement was seen in removing AC5. When CoTA affected the signal, AC13 could be removed easily compared to AC14, which caused the MSE to increase. BS-LSTM was able to show improvement in the range of 49.70% to 44.64%. When AC15 affected the signal, it gave a substantial rise in MSE. The integration of BLSTM and BSCN improved the performance of BLSTM in the range of 7.38% to 48.54%

Similarly, for PRD, the improvement ranged from 1.238% to 54.640%. The minimum PRD was achieved in the case of BS-LSTM for removing AC1; it was a boost of 54.64% over the baseline. Nevertheless, the PRD increased with the introduction of various combinations of artifacts. For AiP, the most significant boost over baseline was seen when BS-LSTM was used to remove AC5, i.e., 29.63%. However, when applied to remove AC8, it could achieve only a 15.89% improvement over the baseline. When CoTA affected the signal, AC11 gave the best results when denoising was performed using the proposed BS-LSTM. For DB1, it was observed that the removal of AC1 was easier than other noises, which showed poor results on all performance metrics. It is to be noted that the improvement in BLSTM performance doubles when the BSCN learning mechanism is deployed to the BLSTM. For MSE and PRD, the maximum improvement was observed for AC1 when BSCN was implemented with BLSTM.

3.4.1.3 Result Analysis For DB2

Similarly, in Table 3.3, the improvement is observed in SNR, MSE, and PRD for DB2. SNR improves in the range of 5.144 dB to 5.963 dB. The removal of AC1 using BS-LSTM gives the maximum SNR of 86.78 dB. The lowest SNR achieved for contamination using SA was for AC3 when denoising was performed using BS-LSTM. AC4 noise removal resulted in the most significant improvement in SNR of 7.86% over baseline. When the EEG signal is contaminated by AiP, the proposed technique showed a maximum boost of 4.82% over baseline when it was used to remove AC5. With the addition of more artifacts, the SNR fell, but the proposed technique still performed better than the baseline and the state-of-the-art techniques. When CoTA affected the signal, the proposed technique performed best for AC14. For AC15, SNR was reduced by 66.06% compared to the SA cases.

In terms of MSE, removing AC1 gives the minimum MSE of 2.922, resulting in a maximum improvement of 36.70% over the baseline achieved by the proposed technique. However, adding artifacts led to an increase in MSE, and the maximum MSE was observed to be 18.75 when denoising was performed for AC15 using BS-LSTM. The removal of AC3 was difficult as the MSE attained was 3.95 when the proposed technique was used for denoising. When AiP affected the signal, the removal of AC5 using BS-LSTM gave the MSE as 5.52, an improvement of 26.76% over the baseline. This value of MSE was closely followed by the performance of BS-LSTM for removing AC13 among the cases of CoTA-affected signal. The combination of AC14 gave the MSE of 11.12, the highest amongst such cases. BS-LSTM also showed improvement over BLSTM in terms of MSE in the range of 2.11% to 29.97%.

In terms of PRD, the proposed technique achieves a maximum improvement of 61.01% over the baseline for AC1, giving a minimum PRD of 2.39%. However, the maximum PRD was achieved for AC2, i.e., 4.62% among the cases of the SA. For cases of AiP, the proposed method's best performance was given by removing AC6, giving the PRD of 5.79%. This case also gave the best improvement over the baseline of 38.31%. The worst performance was for AC9, giving PRD of 9.59% amongst such cases when denoising was performed using BS-LSTM, which was 14.52% better than the baseline. For cases of CoTA, the minimum PRD of 10.06% was observed for removing AC12 using BS-LSTM. Amongst such cases, a maximum improvement of 29.39% over baseline is observed when BS-LSTM removed AC11. BS-LSTM also shows an improvement of 44.46% over BLSTM in terms of PRD. From the addition of SA to AC15, PRD increased almost ten times.

3.4.1.4 Result Analysis For DB3

For DB3, the experimental results are given in Table 3.4. It is observed that the maximum SNR is achieved when denoising is performed using BS-LSTM, i.e., 77.85 dB for AC2, and this value was the highest among all the ACs. This case also gave the maximum improvement of 19.67% over the baseline. Among cases with SA-affected signals, removing AC4 gave a minimum improvement of 15.05%

Similarly, for MSE, the removal of AC1 using BS-LSTM gives a minimum MSE of 3.07, which in the case of AC2 becomes 3.52, the highest among such cases. Also, for AC4, the uplift from the baseline is 36.93% when denoising is being performed using the proposed work. This value is the highest among all ACs. For cases of AiP, AC7 is eliminated efficiently using BS-LSTM, giving an MSE of 5.47. This case also gives the most considerable boost in performance over baseline, i.e., 38.42%. The MSE increases substantially in the case of AC8 elimination using BS-LSTM; it is reported to be 7.78. When CoTA affected the signal, the minimum MSE attained was for AC13, i.e., 14.44, and the maximum was for AC14, i.e., 16.20. But, for the elimination of such ACs, the improvement of BS-LSTM over baseline ranged from 23.64%-

Table 3.4: performance of various techniques when used to perform denoising of EEG signals of patients in DB3.

	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM
	SNR							MSE							PRD						
AC1	64.9331	71.6421	73.1349	73.893	73.0284	75.4267	75.8922	4.6685	4.7349	3.8001	3.66	3.5086	3.1583	3.0749	4.537	4.5025	4.2496	3.7884	3.2172	3.0854	2.5018
AC2	65.0594	71.9108	73.7532	74.4798	74.7554	76.5633	77.8566	5.4882	5.1552	4.583	4.0644	3.922	3.5963	3.5241	5.2334	5.4309	4.5365	4.3553	3.0966	2.901	0.6247
AC3	64.9652	72.6849	73.6533	73.7396	75.3966	76.7859	76.7696	5.0181	4.8549	4.3663	4.4544	3.6733	3.5957	3.5164	4.0975	3.1523	2.3926	2.9336	2.2959	2.0611	0.9992
AC4	66.2416	71.7253	73.8004	75.0597	75.0057	77.2817	76.2167	5.5689	5.1282	4.1045	3.923	3.5791	3.5236	3.5125	4.2812	4.118	3.5932	2.7045	1.9403	1.7945	1.4788
AC5	48.2387	49.4967	49.0642	50.5352	51.8941	52.1339	52.8845	8.8117	8.2021	7.441	6.5142	6.0342	6.017	6.0158	10.7766	9.7936	9.795	8.4888	7.8042	6.6151	4.2135
AC6	48.2199	49.0614	49.6515	50.5797	51.9686	52.0223	52.6671	8.3417	7.9018	7.2242	7.0466	6.1057	5.9359	5.6259	9.6407	7.515	7.6511	7.0671	6.7688	6.01	4.588
AC7	48.4752	49.0736	49.4596	50.6788	51.9441	52.2499	52.5565	8.8925	8.1751	6.781	6.6966	6.0891	5.932	5.4762	9.6612	8.6439	8.8517	6.8381	6.5021	5.5291	5.1928
AC8	54.7385	56.5092	56.8695	58.3497	59.5511	59.5841	60.6492	11.4624	10.7228	9.9072	9.1386	8.1111	8.0924	7.7778	12.7493	10.6946	9.9191	9.8138	7.9712	7.3683	3.9618
AC9	54.9938	56.5214	56.6775	58.4488	59.8323	59.5059	60.5386	12.0132	10.9961	9.4639	8.7887	8.0758	7.9614	7.774	12.7698	11.8235	11.1197	9.5848	7.7045	6.5511	4.9029
AC10	54.9749	56.5187	56.8324	58.4933	59.6342	59.667	60.3212	11.5431	10.6957	9.321	9.2472	8.1474	7.6941	7.5703	11.6339	9.545	8.9758	8.1631	6.6691	5.946	5.2774
AC11	38.1731	41.7291	42.7713	42.9098	42.631	44.0297	44.1992	20.0526	17.8967	17.9021	17.3994	16.1812	15.349	14.4444	17.5048	16.392	15.6297	15.8353	13.8767	13.3268	8.1042
AC12	38.4095	41.692	42.8872	42.9598	42.6404	43.8713	44.1734	19.3926	17.9828	17.8696	17.4543	16.3636	14.3607	14.8081	16.3893	15.2423	14.6864	14.1845	12.4545	12.0247	9.4199
AC13	38.4284	41.5372	42.6431	42.8316	43.0526	44.0887	44.1289	19.6093	18.4529	18.17	17.3828	15.8313	15.1993	14.4406	17.5253	17.5209	16.8303	15.6063	13.0602	13.0596	9.0454
AC14	44.9281	48.9099	50.4697	50.5216	50.0882	51.8534	51.9434	22.0755	21.1035	20.6906	19.441	18.4557	17.2934	16.2026	19.498	18.4219	16.9544	16.9312	13.2271	13.8128	8.7937
AC15	27.728	28.1344	28.807	28.9253	29.0237	29.0843	29.3161	28.1896	28.0233	27.9552	27.7737	27.923	27.6022	25.8845	30.4715	30.2397	29.9668	28.3749	28.456	27.9573	25.8689

Table 3.5: performance of various techniques when used to perform denoising of EEG signals of patients in DB4.

	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM
	SNR							MSE							PRD						
AC1	68.1251	71.7507	72.9207	79.0020	79.5635	81.0136	83.0287	4.7101	3.4550	3.7564	3.5128	3.9980	3.6618	3.6786	4.9373	5.6246	5.0170	5.0462	4.0849	4.9727	3.9501
AC2	68.9728	72.5932	73.7217	79.3402	80.2717	81.0311	83.2362	4.3937	4.4446	4.1042	4.0902	3.8879	3.7891	3.1161	3.7863	4.7900	3.1480	4.2644	4.9624	3.5644	2.5728
AC3	75.9629	79.8711	81.0531	86.6056	86.6754	87.9083	90.9483	4.1333	3.8925	3.1140	3.0856	3.5106	2.9381	2.8947	4.0348	5.4592	3.8048	4.3921	5.3397	5.1231	2.2371
AC4	73.9688	75.8169	77.5587	84.2843	83.6211	86.8052	87.7292	4.6575	4.8913	4.0320	3.9693	3.8827	3.2814	3.2781	3.2196	3.7841	6.3657	5.0640	4.0129	4.4838	2.3993
AC5	48.4828	49.7187	49.8166	52.3346	52.3460	52.2746	53.1223	12.5628	12.1480	12.5296	10.9844	11.8236	12.1518	11.7002	9.1874	10.0118	8.0561	10.0095	7.4483	9.6149	8.7518
AC6	49.8808	51.1743	51.2829	53.6267	53.6500	53.8771	54.5754	12.5221	12.0959	9.9224	10.4495	10.3428	10.6859	10.6871	9.4359	10.6810	8.7129	10.1371	10.2152	9.7755	7.4246
AC7	49.4820	50.3634	50.5840	53.0158	53.2332	53.4294	54.1112	12.7886	13.5551	12.1511	11.8175	11.5011	11.5354	10.8351	8.6207	9.0058	11.2738	10.8090	8.8885	9.1363	7.5867
AC8	56.8628	58.5178	58.7351	61.7247	61.7549	62.2214	62.5433	14.0488	13.3427	13.5724	12.5350	12.5529	11.8475	13.0683	10.7535	12.6587	9.3524	11.8785	10.7456	12.2516	9.0139
AC9	56.4640	57.7070	58.0362	61.1138	61.5343	61.5776	62.0790	14.4746	15.4033	14.1034	13.7698	12.0979	11.6851	12.4337	9.9383	10.9836	11.9133	12.5504	9.4188	11.6124	9.1761
AC10	57.8620	59.1626	59.5025	62.3946	62.9097	63.1201	63.5321	14.3734	15.0176	12.5040	13.0344	12.1647	12.0081	12.9606	10.1868	11.6528	12.5701	12.6780	12.1857	11.7730	7.8488
AC11	35.7997	37.6679	38.2470	41.0894	41.3458	41.8892	43.1398	17.5547	17.2982	15.7650	15.6054	16.8761	15.8450	15.0020	16.0187	19.6988	16.0450	18.3599	16.2985	18.2626	13.3804
AC12	36.7989	38.3127	39.0144	42.0782	42.0157	43.0441	44.0384	17.8185	17.7449	15.6929	15.4845	16.8709	15.3374	15.1640	15.4520	18.6929	19.2627	19.1594	17.7386	17.7840	12.2153
AC13	35.4008	36.8571	37.5481	40.6251	40.7349	41.6686	42.4960	18.0789	18.2970	16.6830	16.4890	17.2483	16.1883	15.3855	15.2035	18.0237	18.6059	19.0318	14.9717	17.6233	13.5425
AC14	43.7809	45.6562	46.4667	50.0460	50.1137	51.1489	52.3828	19.8572	20.4620	17.9188	17.8182	18.7599	17.2955	16.4408	16.7696	20.6706	19.9022	20.9008	18.2689	20.2601	13.8046
AC15	27.6256	27.4912	28.9007	28.7378	29.6273	29.0200	29.3105	31.1463	29.3728	26.6568	27.0455	25.7090	24.4038	23.8807	29.1560	28.6849	26.8989	26.4404	26.0866	24.4256	23.9214

27.97%. When AC15 affects the signal, the MSE increases by 88.12% to 25.88.

Regarding PRD, minimum PRD is attained when BS-LSTM is used to remove AC2, i.e., 0.62%; this value also results in a maximum improvement of 88.06% over the baseline among SA-affected signal cases. AC1 was harder to eliminate even when denoising was performed using the proposed technique, as it gave the PRD of 2.50%, which was a 44.85% improvement over the baseline. When AiP affected the signal, the minimum PRD attained was 3.96% for AC8 when it was denoised using BS-LSTM. This case also gave the maximum boost over the baseline, i.e., 68.92%. Among these cases, AC10 has the maximum PRD of 5.28%. For cases of CoTA, AC11 is eliminated using BS-LSTM, resulting in a minimum PRD of 8.10% among such cases. But, the maximum improvement BS-LSTM gave over the baseline was for AC14, i.e., 54.90%. This value was minimum for AC12, i.e., 42.52%. When AC15 affected the signal, the PRD increased by 90.33%, giving a PRD of 25.86%. It is to be observed that BLSTM performance is also enhanced with the addition of BSCN, giving a boost in the range of 6.08% to 78.47%.

3.4.1.5 Result Analysis For DB4

Table 3.5 for DB4 indicates that the SNR for the elimination of AC3 using BS-LSTM was the highest, i.e., 90.95 dB. For AC1 elimination, SNR achieved was 83.02 dB, which also gave the highest boost of 21.87% over the baseline when BS-LSTM was used for denoising. In the cases of AiP, AC10 gave the maximum SNR of 63.53 dB when elimination was done using BS-LSTM among such cases. However, the elimination of AC8 using BS-LSTM gave the maximum improvement over the baseline, i.e., 9.98%. When CoTA affected the signal, the maximum SNR of 52.38 dB was achieved for AC14 when the denoising was performed using BS-LSTM. However, this gave the minimum improvement of 19.65% over the baseline among such cases. When AC15 affected the signal, the SNR dropped by 67.77% from the SNR attained from SA, i.e., 29.31 dB. BLSTM also improved in terms of efficiency when BSCN is associated with it. It was observed that BS-LSTM was 3.46% more efficient as compared to BLSTM.

In terms of MSE, the minimum MSE was achieved for AC2, i.e., 2.89, when the denoising was performed using BS-LSTM, which was the least among all the ACs. This also showed the most significant boost of 29.97% over baseline when denoising was performed using BS-LSTM. However, in cases where SA affected the signal, maximum MSE was attained for AC1, i.e., 3.68, which gave the minimum improvement of 21.90% over baseline among such cases. For cases of AiP, the minimum MSE was attained for AC6, i.e., 10.68 when denoising was performed using the proposed technique. However, maximum MSE was achieved for AC8, i.e., 13.07. The improvement over the baseline achieved was in the range of 6.86% to 15.27% when denoising was performed using BS-LSTM. For cases of CoTA, AC11 was removed most efficiently as it gave the MSE as 15.00. However, this case gave the minimum improvement over the baseline of 14.54%. Among these cases, AC14 gave the maximum MSE of 16.44, but this case also had the most significant boost over the baseline of 17.20%. When AC15 affected the signal, the MSE increased 6.5 times from the SA-affected signal. The performance of BLSTM also improved with the inclusion of the BSCN learning technique in the range of 0.10% to 17.76%

Similarly, in terms of PRD, the best performance was given by BS-LSTM for eliminating AC3; PRD attained was 2.23%, which was the minimum across all ACs. This case also gave the most significant improvement over baseline, i.e., 44.55%. In some cases, such performance could not be achieved, e.g., in the case of elimination of AC1 using the proposed technique, only a 19.99% boost over baseline could be achieved, giving PRD of 3.95%. For cases of AiP, AC6 was eliminated using BS-LSTM, which gave the best performance with a PRD of 7.42%. For some cases like AC9, the PRD was high, i.e., 9.17%, even when denoising was performed using BS-LSTM. For cases of CoTA, the proposed technique removed the combination of AC12 most efficiently, giving a PRD of 12.21%. This case also gave the most considerable improvement over the baseline of 20.95% when the denoising was performed using BS-LSTM among all the ACs. However, the combination of AC14 was eliminated, giving a PRD of 13.80%, which was the maximum in such cases. PRD increased substantially when AC15 contaminated the signal while denoising was performed using the proposed technique. It attained the overall maximum PRD of 23.72% amongst all the ACs. The BLSTM's performance was also enhanced with the inclusion of BSCN, giving a boost of 56.33%, maximum for AC3.

3.4.1.6 Result Analysis For DB5

DB5 is a sleep data set, and experiments were performed on this data to prove the domain flexibility of the proposed denoising methodology. BS-LSTM can be used for any EEG data set apart from epilepsy EEG data sets. As can be observed from Table 3.6 for DB5, the SNR was above 74.11 dB for all SA cases on all the techniques, including the proposed novel BS-LSTM. The removal of AC1 gave the best results, with SNR reaching 85.61 for the proposed method. For cases of SA-affected signal, denoising of AC3 signal gave the minimum SNR of 78.83 dB using the proposed BS-LSTM. The removal of AC4 using BS-LSTM gave the maximum boost of 6.79% over the baseline. Among the cases of AiP, AC9 was eliminated quite efficiently as it gave the highest SNR of 63.74 dB, and AC6 could not be eliminated that efficiently as it gave an SNR of 54.55 dB. In terms of improvement over baseline, the maximum improvement of 8.58% was observed in the case of AC5 removal, which was also the overall maximum boost given by the proposed technique across all ACs. When CoTA affected the signal, the proposed technique could remove AC14 most efficiently, giving an SNR of 49.08 dB. However, for AC12, the SNR was reduced to 41.29 dB when denoising was performed using the proposed BS-SLTM. When AC15 affected the signal, SNR was reduced by 68.49% when denoising was performed using the proposed technique on the SA-affected signal. However, this case showed an improvement of 5.96% over the baseline.

Similarly, in terms of MSE, too, AC3 gave the lowest MSE of 2.83, which is the best MSE achieved overall among all the ACs, followed closely by 2.85 reported for AC4 when denoising is performed using the proposed BS-LSTM. However, a maximum improvement over baseline was observed to be 19.23% for AC2 elimination among the SAs, using the proposed BS-LSTM. When AiP affected the signal, it was found that AC6 gave the least MSE of 8.055. This case also significantly improved performance over the baseline of 17.13% among such cases. However, for AC10, the MSE increased to 9.75 when the denoising was performed using BS-LSTM. In cases of CoTA, AC11 was eliminated efficiently as it gave the minimum MSE of 8.70, whereas AC14 was challenging to remove, which resulted in an MSE of 17.85. For AC15, the MSE was the overall highest among all ACs, i.e., 22.88. However, this gave the maximum improvement over the baseline of 20.24%. BS-LSTM also showcased the most significant boost over BLSTM for AC15, i.e., 19.61%.

Regarding PRD, the proposed BS-LSTM proved superior, giving the best result for AC4, giving PRD of 2.46%. Also, this case had the most significant improvement of 36.98% over the baseline, overall, for all the ACs. The PRD value was highest for AC1, i.e., 3.45% among the individual artifacts affecting the EEG signal. For AiP, AC7 was eliminated efficiently, giving a PRD of 7.45% when the proposed technique is used for denoising. This case had the most significant improvement of 28.07% over the baseline. For cases of CoTA, AC11 gave the minimum PRD of 8.46%, which was the most considerable improvement among such cases, i.e., 23.82%. Nevertheless, for the case of AC14, the PRD increased substantially, i.e., 17.24%.

Table 3.6: Performance of various techniques when used to perform denoising of EEG signals of patients in DB5.

	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM	Fast-ICA	MOFPA-WT	WD-ESN	VMD-DWT-WPT	RNN-LSTM	BLSTM	Proposed BS-LSTM
	SNR							MSE							PRD						
AC1	80.5857	81.9474	82.4576	83.0955	83.2621	83.4079	85.6100	3.6959	3.9901	4.3603	3.6562	3.8155	3.6578	3.4015	4.3603	4.0577	3.8155	3.6959	3.6578	3.6562	3.4599
AC2	80.2488	80.3972	80.7266	81.5748	81.8324	83.1669	84.9885	4.2687	3.7910	4.1383	3.5925	3.3976	3.9255	3.4474	4.9367	4.1383	3.9255	3.7910	3.5925	3.4474	3.3976
AC3	74.1194	74.6142	74.8273	75.7495	76.1998	76.5039	78.8290	2.9359	2.9825	3.5048	2.8226	2.6602	3.3114	2.8343	3.5048	3.3114	2.9825	2.9359	2.8343	2.8226	2.6602
AC4	76.2567	76.4696	76.8412	77.6051	78.5138	78.7857	81.4377	3.1507	3.4949	3.9108	3.4652	3.3495	3.1867	2.8590	3.9108	3.6065	3.4652	3.3495	3.1867	3.1507	2.8590
AC5	51.7907	53.3136	53.4661	53.0530	54.9526	54.7669	56.2331	9.7022	8.6974	9.3012	8.1121	10.3061	8.8608	8.3190	10.3061	9.7022	9.2896	8.8608	8.6974	8.3190	8.1121
AC6	50.8239	51.3753	51.5732	51.6647	53.9760	53.6886	54.5521	9.7213	8.0180	8.0004	8.3377	10.1617	7.9842	8.0560	10.1617	9.7213	8.3377	8.0560	8.0180	8.0004	7.9842
AC7	51.4304	52.1634	52.2883	52.4049	54.2413	54.5820	54.5747	9.7067	7.4522	8.5819	8.4384	10.3611	8.8413	8.4090	10.3611	9.7067	8.8413	8.5819	8.4384	8.4090	7.4522
AC8	59.1609	59.5172	60.7881	59.3107	61.9619	61.5820	63.0870	10.7015	9.2402	10.3409	9.4570	10.8194	9.9004	9.5379	10.8194	10.7015	10.4946	9.9004	9.5379	9.4570	9.2402
AC9	59.7288	60.0549	60.3724	60.4054	62.5029	62.3606	63.7374	10.8507	9.3439	9.6123	9.8404	10.9122	10.3794	9.4199	10.9122	10.8507	10.3794	9.9510	9.6123	9.4199	9.3439
AC10	57.8464	59.1873	58.8131	58.8258	61.3151	61.1003	61.9222	11.1571	11.4631	9.7314	10.4438	9.5257	10.4629	9.7578	11.4631	11.1571	10.4629	10.0780	9.7578	9.7314	9.5257
AC11	38.6258	39.2718	39.3374	40.0845	40.7124	40.8854	41.8811	10.2216	11.1128	9.3622	8.6928	8.4653	9.1920	8.7089	11.1128	10.2216	9.5302	9.1920	8.7089	8.6928	8.4653
AC12	38.6746	38.7452	39.0628	39.5007	39.2565	40.2599	41.2961	17.8052	17.3300	18.3300	16.7561	16.0504	15.1901	15.7106	18.3300	17.8052	17.3300	16.5094	16.0278	15.7106	15.1901
AC13	39.3799	40.0068	40.5883	40.1885	40.2238	41.1690	41.8458	18.1801	17.8297	17.4089	16.1756	15.7324	14.9691	15.2209	19.5511	19.1964	18.7394	17.4782	16.7720	15.2209	14.9691
AC14	46.1498	46.2323	46.6260	47.1480	47.8347	47.9341	49.0834	21.0051	19.5083	20.1681	18.4229	18.1995	17.2415	17.8513	21.9292	20.1681	19.5083	18.3750	18.1151	17.8513	17.2415
AC15	25.4529	25.4430	26.0663	26.3146	26.0344	26.8200	26.9721	28.6934	29.1398	28.1414	23.5171	30.6947	28.4685	22.8841	30.6947	29.0584	28.6934	28.4685	28.1414	23.5171	22.8935

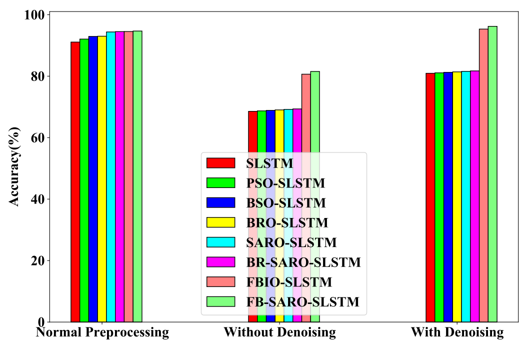
Further, when AC15 affected the signal, the PRD reached 22.89%, which was 25.41% over the baseline when denoising was performed using BS-LSTM. In terms of improving BS-LSTM over BLSTM, it was observed that the proposed method improved in the range of 0.56% to 3.36% in SNR, 4.88 to 19.61% in MSE, and 0.20 to 5.37% in terms of PRD.

It should be noted that the addition of BSCN learning impacts the performance of BLSTM and provides improved performance in all types of artifacts across all data sets. The reduced values of SNR, MSE, and PRD due to the combination of all four artifacts across all the denoising techniques for all the data sets show that these denoising techniques did not wholly eradicate all artifacts when all four types were simultaneously present.

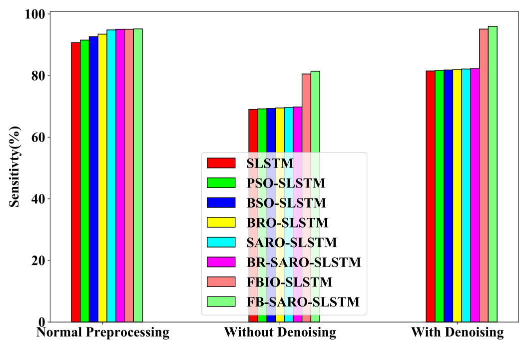
3.4.2 Evaluation Based on Classification and Prediction Tasks

After performing denoising of the EEG signal and evaluating various techniques on the denoising task, the experimentation was taken a step further. Experiments were performed to compare the classification and prediction performance of various state-of-the-art classifiers. These experiments were performed to evaluate the quality of the EEG signals after the required preprocessing or denoising had been performed on them. The classifiers used were Stacked LSTM (SLSTM), PSO based SLSTM, Brainstorm Optimization (BSO) based SLSTM, Battle-Royale Optimization (BRO) based SLSTM, Search and Rescue Optimization (SARO) based SLSTM, Forensic Based Investigation Optimization (FBIO) based SLSTM, and FB-SARO based SLSTM (see [46, 168] for experiment details).

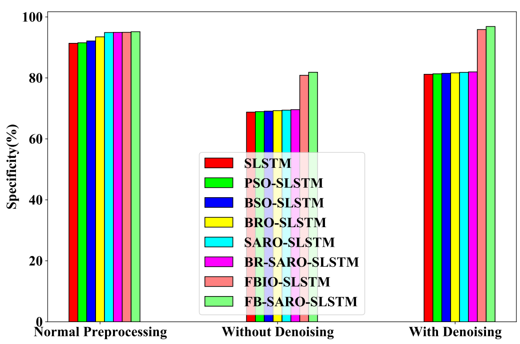
From each of these files, 275 features were extracted, and classification was performed using this feature set to compare the accuracies after the injection of noise into the signal and after the noise was removed from the signal. This gives us the significance of the proposed technique. The performance of the classifiers on the noisy and denoised signal is measured in



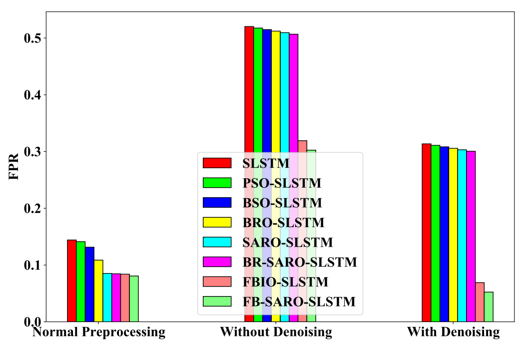
(a)



(b)



(c)



(d)

Figure 3.4: The effect of denoising performed using Bidirectional Stochastic LSTM. During the classification of Epileptic Seizures in DB1, the improvement can be seen with respect to (a) Accuracy (b) Sensitivity (c) Specificity. Also, while predicting the occurrence of Epileptic Seizures, improvement in terms of False Prediction Rate can be observed in (d).

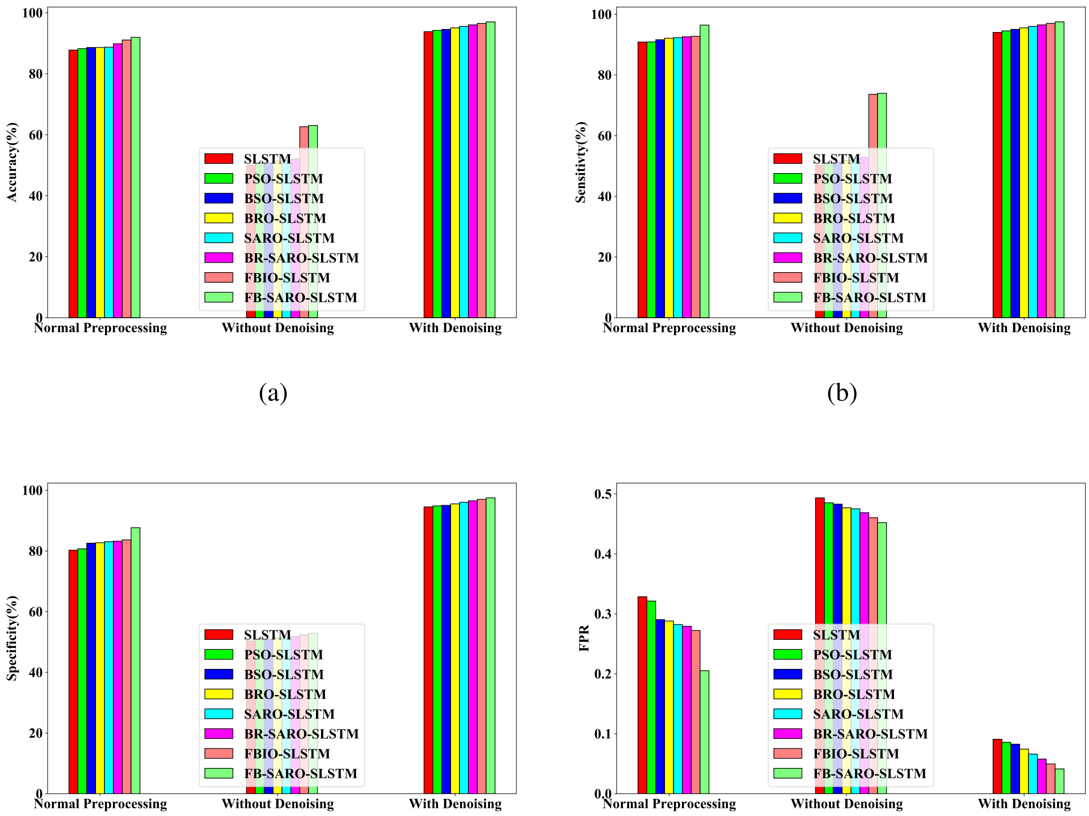
terms of Accuracy, Sensitivity, and Specificity. For DB1 and DB2, FPR has also been reported, which is defined as:

$$FPR = \frac{FalsePositives}{FalsePositives + TrueNegatives} \tag{3.19}$$

The performance was compared for three scenarios.

- **Normal Preprocessing:** This involves preprocessing the EEG signal before sending it to the classifier. It uses a Finite Impulse Response (FIR) band-pass filter. The filter has a frequency range of 0.25–25.0 Hz [168, 169].

- **Without Denoising:** In this scenario, the noisy data, created in Section 3.3, is used for



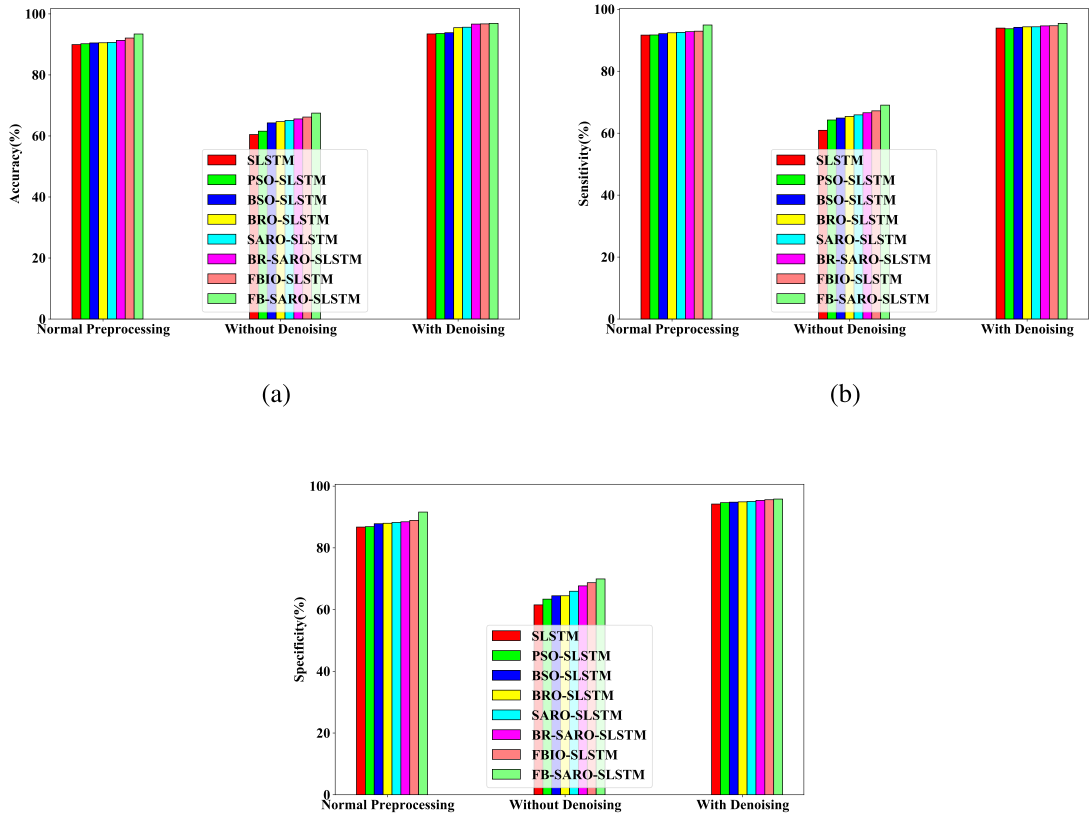
(a) (b) (c) (d)

Figure 3.5: The effect of denoising performed using Bidirectional Stochastic LSTM. During the classification of Epileptic Seizures (ESs) in DB2, the improvement can be seen with respect to (a) Accuracy (b) Sensitivity (c) Specificity. Also, while predicting the occurrence of ESs, improvement in terms of False Prediction Rate can be observed in (d).

the classification and prediction of ES occurrence. These input EEG signals were not preprocessed or denoised.

- **With Denoising:** This scenario involves the proposed denoising methodology using BS-LSTM as the preprocessing step to clean noisy or contaminated EEG signals. After denoising has been performed, the signal is sent for performing the classification and prediction of ES.

Figure 3.4 indicates that when the noisy data was classified, the accuracy reduced to 68.55% from 91.10%, which was achieved using normal preprocessing. Whereas, after per-

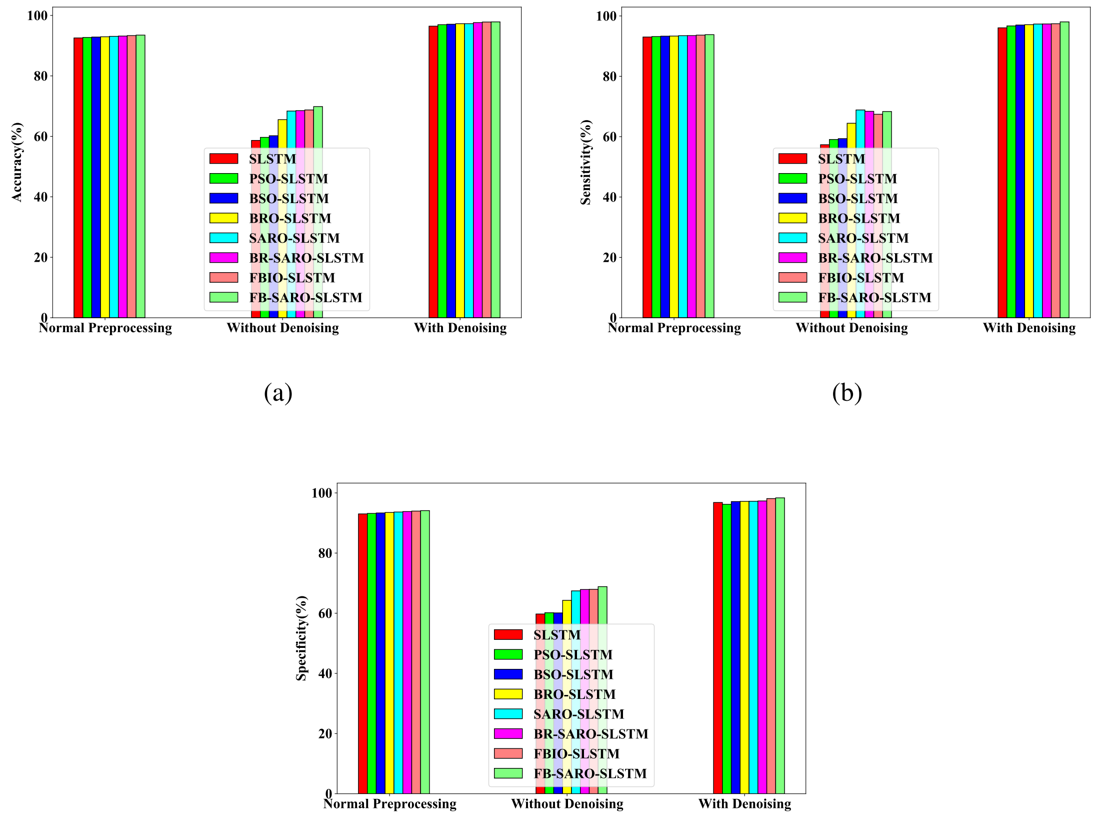


(a) (b)

(c)

Figure 3.6: The effect of denoising performed using Bidirectional Stochastic LSTM. During the classification of Epileptic Seizures in DB3, the improvement can be seen with respect to (a) Accuracy (b) Sensitivity (c) Specificity.

forming denoising using the proposed BS-LSTM, it gave an accuracy of 80.93%. The best accuracy has been achieved using FB-SARO-SLSTM using normal preprocessing at 94.69%; without the denoising, the accuracy goes down to 81.53%, and after the denoising has been performed, the accuracy increases to 96.21%. Similarly, for sensitivity, classification with noisy data reduced the sensitivity to the range of 69.02% to 81.40%. However, after performing denoising on the EEG signals, the sensitivity achieved was in the range of 81.50% to 95.98%. When the band-pass filter was used for preprocessing, the sensitivity achieved was in the range of 90.71% to 95.17%. In the case of specificity, 17.7% to 18.5% improvement is seen when data denoising is done using the proposed novel BS-LSTM-based methodology. After the clas-



(a) (b)

(c)

Figure 3.7: The effect of denoising performed using Bidirectional Stochastic LSTM. During the classification of Epileptic Seizures in DB4, the improvement can be seen with respect to (a) Accuracy (b) Sensitivity (c) Specificity.

sification task was experimented with, the prediction task was performed on the EEG signals to perform seizure occurrence. FPR reported is in the range of 0.14/hr to 0.08/hr when data preprocessing is done without the proposed technique, whereas when proposed denoising is used, FPR lies in the range of 0.31/hr to 0.05/hr.

When classification and prediction of ES were made for the patients of DB2, similar trends were observed across all the performance metrics, as shown in Figure 3.5. In terms of accuracy, the noisy data caused the accuracy to become as low as 50.09%. However, once the denoising was performed on the signals, the accuracy increased by 53.88% to 87.32%, giving a maximum accuracy of 97.029% given by FB-SARO-SLSTM. This boost is more than 75.24% over the accuracy achieved using band-pass filter preprocessing. For sensitivity and specificity, the proposed denoising technique made significant improvements. Sensitivity increased in the range 31.82% to 86.64%, and specificity increased in the range 84.46% to 87.6%. While predicting the ESs, the FPR from noisy data was observed to be 0.45/hr, but after performing denoising using BS-SLTM, the FPR was reduced to 0.04/hr.

Figs. 3.6(a) - 3.6(c) indicate that accuracy attained using a band-pass filter for preprocessing lies in the range 89.95% to 93.41%; when classification is performed on the noisy data, the accuracy falls to 60.46% when SLSTM was used. When the proposed denoising technique is used as part of the preprocessing, it is observed that accuracy improved in the range of 93.43% to 96.90%; FB-SARO-SLSTM gives the best performance. Similarly, for sensitivity, the classifiers performed poorly when noisy data were used to classify seizures from non-seizures. The sensitivity was observed to be 60.93% when SLSTM was used for classification; this was the minimum value obtained. The signal is passed through a band-pass filter to improve the performance of the classification, but it improves by a narrow margin, giving a sensitivity of 91.67% for SLSTM. However, once BS-LSTM is utilized to perform the preprocessing of the signal, sensitivity increased to 95.46% for FB-SARO-SLSTM, followed closely by FBIO-SLSTM, which gave a sensitivity of 94.69%. A similar trend can be observed for specificity as well. The noisy data caused the specificity to get reduced to 61.52% when the classification was performed using SLSTM. When the classification of EEG signals involved normal processing using a band-pass filter, the specificity attained is in the range of 86.74% to 91.60%. However, an improvement can be observed in the overall performance of the classification done by the SLSTM-based techniques, as the specificity attained was in the range of 94.20% to 95.81%. FB-SARO-SLSTM performed the best on all performance metrics.

These experiments were performed on DB4 as well, and the results are illustrated in Figs.

3.7(a) - 3.7(c). The resulting accuracy after classification was performed on noisy data using SLSTM is 58.55%. This value could only reach 69.85% when the classification was performed using FB-SARO-SLSTM. When preprocessing is performed on the signal using a band-pass filter, the accuracy increases in the range of 92.56% to 93.49%. However, if the proposed denoising technique is used in the preprocessing step, the accuracy increases substantially to 97.87%, which is the highest for the FB-SARO-SLSTM classifier. Similarly, for sensitivity, the noises in the signal cause it to fall to 57.33% when classification is performed using SLSTM. For other techniques like PSO-SLSTM or BSO-SLSTM, the sensitivity only improves a little due to the artifacts present in the signal. It reaches 68.33% for classification performed by FB-SARO-SLSTM. After the signal goes through preprocessing using a band-pass filter, the sensitivity improves and is observed to be 93.02%-93.80%. However, when denoising using BS-LSTM is done as a part of the preprocessing step, it can be seen that sensitivity increases significantly to 98.04% for classification performed by FB-SARO-SLSTM. A similar trend can be observed in specificity as the classification done on noisy signal causes it to reduce to 59.75%. When normal preprocessing is performed on the signal, the specificity increases by 26.86% to 35.77%. However, this range increases to 30.0%-62.02%, the best performance given by FB-SARO-SLSTM, giving a specificity of 98.35%.

For DB5, the experiments were performed with noisy data, and results were compared after denoising was performed using the novel BS-LSTM, as shown in Figure 3.8. Accuracy reduced to 60.59% but, after denoising, increased to 94.55%. There was an improvement in accuracy in the range of 34.044% to 37.63%. The best accuracy achieved was for FB-SARO-SLSTM, closely followed by FBIO-SLSTM. Regarding sensitivity, the improvements were seen in the range of 34.045% to 37.63%. Similarly, for specificity, the improvements were in the range of 36.70% to 39.87%. FPR was derived for the prediction task, and the lowest value was

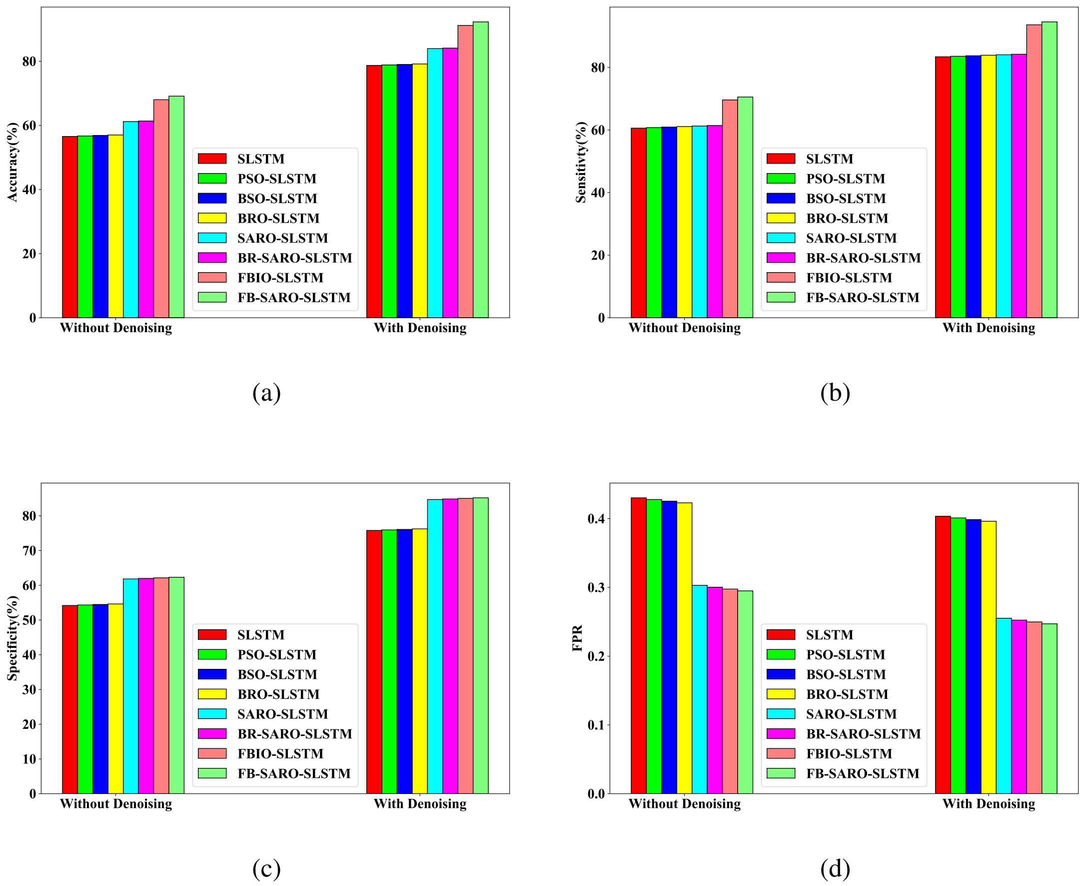


Figure 3.8: The effect of denoising performed using Bidirectional Stochastic LSTM. During the classification of Epileptic Seizures (ESs) in DB5, the improvement can be seen with respect to (a) Accuracy (b) Sensitivity (c) Specificity. Also, while predicting the occurrence of ESs, improvement in terms of False Prediction Rate can be observed in (d).

achieved using FB-SARO-SLSTM, i.e., 0.25/hr, which was a 19.36% improvement on the FPR derived using noisy data.

It is observed that the quality of the EEG signal degrades due to contamination from artifacts to the extent that it becomes unfit to be used for any classification or prediction task. Through experimentation, this research proves that the denoising being performed by the proposed novel BS-LSTM on the EEG signals enhances the quality of the input signal. These EEG signals, therefore, increase the efficiency of the classification task or the prediction task at hand. The presence of noise in the data can negatively affect the performance of DL models, making it harder for the model to identify patterns and relationships in the data accurately. The proposed denoising technique is designed to reduce the effect of noise on the data by removing or reducing the impact of the noise in the input. Therefore, when the proposed denoising technique is applied, the input data becomes cleaner, and the models can more accurately identify the patterns and relationships in the data, resulting in improved performance.

3.5 Summary

EEG signals frequently experience several sorts of artifacts, which makes it challenging to extract useful information from them. This impairs the ability to do end tasks like ES classification and prediction. This work suggests a novel BS-LSTM architecture combining the BLSTM framework with the BSCN learning mechanism. The EEG data of epileptic patients were cleaned up using this innovative design. The proposed method works well with five publicly accessible data sets and is effective. BS-LSTM can effectively eliminate ECG, EOG, EMG, PL, and their combinations. When learning was carried out using the technique utilized in BSCN, it was observed that the performance of BLSTM improved. The proposed approach provides consistent improvement for all the data sets. Additionally, tests have demonstrated that the proposed denoising approach improves the quality of the EEG data by eliminating artifacts and increasing the effectiveness of classification or prediction tasks that use these signals. The underlying brain activity may be more accurately analyzed and interpreted when noise is removed from the EEG readings. This is because denoising using the proposed BS-LSTM resulted in more precise and dependable feature extraction while maintaining the temporal and spatial properties of the EEG data. This characteristic of the proposed method makes it particularly useful for research that intends to identify biomarkers, gauge the effectiveness of treatments, or look at minuscule brain abnormalities.

In this chapter, it is seen how EEG signals can be cleaned and made artifact-free. This makes the signal ready for classification and prediction tasks to be performed on them to create efficient seizure management systems. Therefore, the next chapter introduces a new hybrid optimization technique to improve seizure detection performance by identifying the optimal LSTM architecture. The hybrid feature set effectively captures the non-linear characteristics of EEG signals, enhancing seizure classification.