IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 10, 2020, accepted April 19, 2020, date of publication April 22, 2020, date of current version May 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2989442

Classification of Focal and Non-Focal Epileptic Patients Using Single Channel EEG and Long Short-Term Memory Learning System

LUAY FRAIWAN[1,2], (Senior Member, IEEE), AND MOHANAD ALKHODARI[1]
[1]Department of Electrical and Computer Engineering, Abu Dhabi University, Abu Dhabi, United Arab Emirates
[2]Department of Biomedical Engineering, Jordan University of Science and Technology, Irbid 22110, Jordan

Corresponding author: Luay Fraiwan (fraiwan@just.edu.jo)

This work was supported by the Abu Dhabi University's Office of Research and Sponsored Programs.

ABSTRACT The process of inspecting electroencephalography (EEG) signals of patients with epilepsy to distinguish between focal and non-focal seizure source is a crucial step prior to surgical interference. In this paper, a deep learning approach using a long short-term memory (LSTM) algorithm is investigated for the purpose of automatic discrimination between focal and non-focal epileptic EEG signals. The study is carried out by acquiring 7500 pairs of x and y EEG channels signals from the publicly available Bern-Barcelona EEG database. The manual classification of each signal type was visually done by two board-certified electroencephalographers and neurologists. Initially, every channel signals are pre-processed using z-score normalization and Savitzky-Golay filtering. The signals are used as inputs to a pre-defined Bi-directional LSTM algorithm for the training process. The classification is performed using a k-fold cross-validation following 4-, 6-, and 10-fold schemes. At the end, the performance of the algorithm is evaluated using several metrics with a complete summary table of the recent state-of-art studies in the field. The developed algorithm achieved an overall Cohen's kappa κ, accuracy, sensitivity, and specificity values of 99.20%, 99.60%, 99.55%. and 99.65%, respectively, using x channels and 10-fold cross-validation scheme. The study pave the ways toward implementing deep learning algorithms for the purpose of EEG signals identification in a clinical environment to overcome human errors resulting from visually inspection.

INDEX TERMS Classification, electroencephalography (EEG), epilepsy, focal, long-short-term memory (LSTM), non-focal, training.

I. INTRODUCTION

Epilepsy has become a major neurological disorder of the brain that is characterized by the occurrence of repeated seizures. Normally, the brain sends electrical impulses to the whole body throughout neurons and neurotransmitters. In a case of an epileptic seizure, these electrical waves are disrupted resulting in an imbalanced reactions from the body [1]. According to the World Health Organization (WHO) [2], more than 50 million people around the world are suffering from epileptic seizures, where as 80% among them are living in low- and middle-income countries. Despite the development of anti-epileptic medications, 33% of epileptic patients do not respond to the medication and are expected to suffer from further unpredictable seizures in future.

Epilepsy is of two major types; focal and non-focal (generalized). In focal epilepsy, only a specific part of the brain is affected, i.e. one hemisphere of the brain. On the other hand, non-focal epilepsy affects multiple areas within the brain even though they were not affected directly by the seizures [1], [3]. The majority of epileptic patients (60%) become pharmacoresistant, that is not responding to medications. Therefore, they require surgical interference to treat the seizures [4]. As a result, the precise localization of seizure areas is important prior to the surgery to reduce the risks accompanied with invasive interference with the brain.

Electroencephalogram (EEG) is considered the gold standards in epilepsy diagnosis for its ability to detect the brain electrical activity as well as the presence of epileptic seizures

The associate editor coordinating the review of this manuscript and approving it for publication was Siddhartha Bhattacharyya.

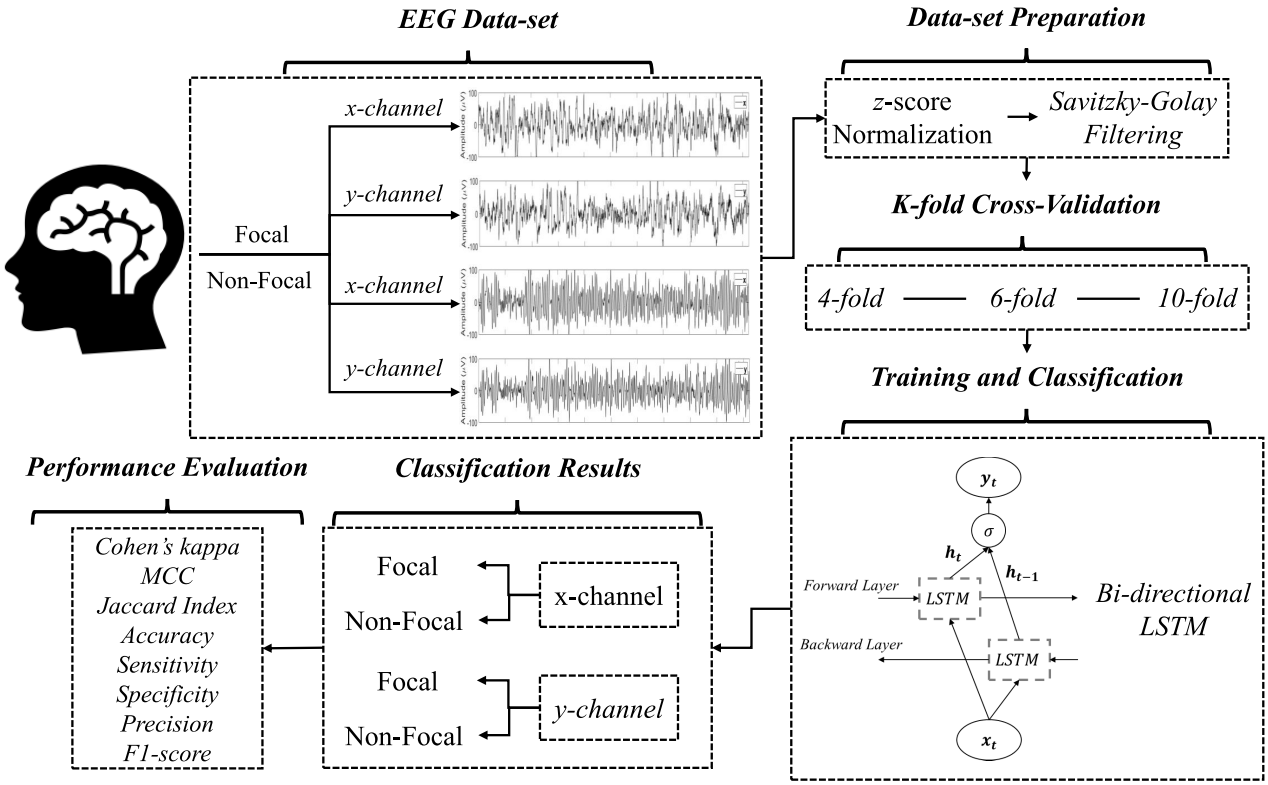


FIGURE 1. The complete procedure followed in the study.

and abnormalities [5]. EEG allows for obtaining focal and non-focal signals, thus, discriminating both epilepsy types and identifying the region that the seizure is originating from. In addition, EEG allows for the identification of both the ictal and inter-ictal seizure activities. The ictal EEG represent a continuous wave of spikes. On the other hand, inter-ictal EEG is represented by the presence of temporary sharp waves. Sometimes, clinicians use long-term intercranial EEG recording to detect deeper signals and localize the source of the seizure within the brain [3], [5]. Therefore, visual identification of focal and non-focal epileptic EEG signals is considered time consuming for medical doctors. In addition, multiple experts may have different views for the patient EEG patterns. Thus, human error in diagnosing epileptic seizure source is of a high chance of occurring [6]. Thus, there is a major need for an automated classification approach that is able of distinguishing between the characteristics of such signals and assisting clinicians in their decision making process prior to surgical intervention.

Due to the difficulty in visually differentiating between EEG signals in both epilepsy types, machine learning algorithms have become of a high importance in detecting differences and classifying such signals. Among these algorithms are the Least-Squares Support Vector Machine (LS-SVM) [7], [8], Adaptive Neuro-Fuzzy Inference Systems (ANFIS) [9], K-Nearest Neighbor (KNN) [10], and Bayesian Linear Discriminant Analysis (BLDA) [11] learning systems. However, these algorithms require manual feature extraction including the time- and frequency-domains features. In addition, features extracted in entropy and wavelet transforms are also utilized. Therefore, the use of a deep learning approach could ease the classification process, as it does not require manual feature engineering that usually requires continuous tuning due to the variable nature of EEG signals. Among these algorithms are the Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM), which has been used very recently in few studies for other EEG classification applications [12]–[14].

A. OUR CONTRIBUTION

In this paper, a deep learning approach following an LSTM algorithm is explored for the purpose of focal and non-focal epileptic EEG signals classification (Figure 1). The obtained EEG signals corresponding to both epilepsy types are acquired from an online database and processed for two channels to be described later in the paper; the x- and y-channels. Initially, the data-set is pre-processed for all signals by passing them to z-score normalization step followed by digital filtering. The training and classification process was applied using both channels through a pre-defined Bi-directional LSTM algorithm. For each channel, the LSTM classifier was trained following a k-fold cross-validation using 4-, 6-, and 10-fold schemes. At the end, the overall performance is evaluated using several performance metrics including the accuracy, sensitivity, specificity, precision,

F1-score, Cohen's Kappa (κ), Matthews Correlation Coefficient (MCC), and Jaccard Index (JI). Each step is briefly explained in the upcoming sections.

B. PAPER ORGANIZATION

The paper is organized as follows. In Section II, the LSTM network architecture is described briefly with general background information. For Section III, the materials and methods used in the study are presented including the data-set selected, the pre-processing steps, the training and classification schemes, and the evaluation metrics definitions. For section IV, the results are provided along with a brief discussion on the observations. The paper is concluded with future work in Section V.

II. LONG SHORT-TERM MEMORY

In Recurrent Neural Networks (RNNs), chain-like structures are used to capture temporal sequences between the data. However, this cause problems when training data using back-propagating processes such as the exploding/vanishing gradient problems [15]. Therefore, LSTM networks were first introduced back in 1997 by Hochreiter and Schmidhuber [16] to store the long-term dependencies between data points. LSTM has been used for several applications including speech recognition, image detection, and language modeling [17], [18].

The architecture of the LSTM network includes memory blocks, which are the input (i), output (o), and forget (f) gates, and a cell that manages the flow of information between the three surrounding gates. The input and output gates control the activation of the input and output information flow, and are described as follows,

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (1)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (2)$$

In addition, the forget gate controls the memory needed to be kept within the network, and is given as,

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (3)$$

where W_{x*} is the input to gate weight, W_{h*} is the hidden to hidden weight, W_{c*} is the peephole weight, b_* is a bias vector, and h_t is the hidden layer output and is given by,

$$h_t = o_t tanh(c_t) \quad (4)$$

where c_t represents the input cell which can be defined as,

$$c_t = tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (5)$$

Moreover, the output cell C_t is described as,

$$C_t = f_t C_{t-1} + i_t c_t \quad (6)$$

The input/output cells are connected to the gates by feedback sources with a Constant Error Carousel (CEC) that activates on each input entry to allow gradients to flow unchanged. The activation function is the sigmoid $\sigma()$ bounded by (0,1).

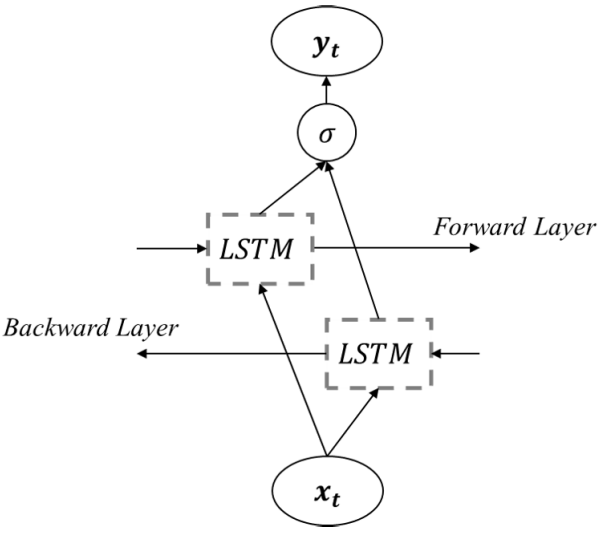


FIGURE 2. The Bi-directional LSTM shortcut architecture showing both the forward and backward directions.

The previous equations describe the LSTM standard model in the forward chain using a hidden layer h_{t-1}. To illustrate the backward chain used in the Bi-directional LSTM (BDLSTM) networks, hidden layer h_{t+1} is utilized. This results in having the overall out as,

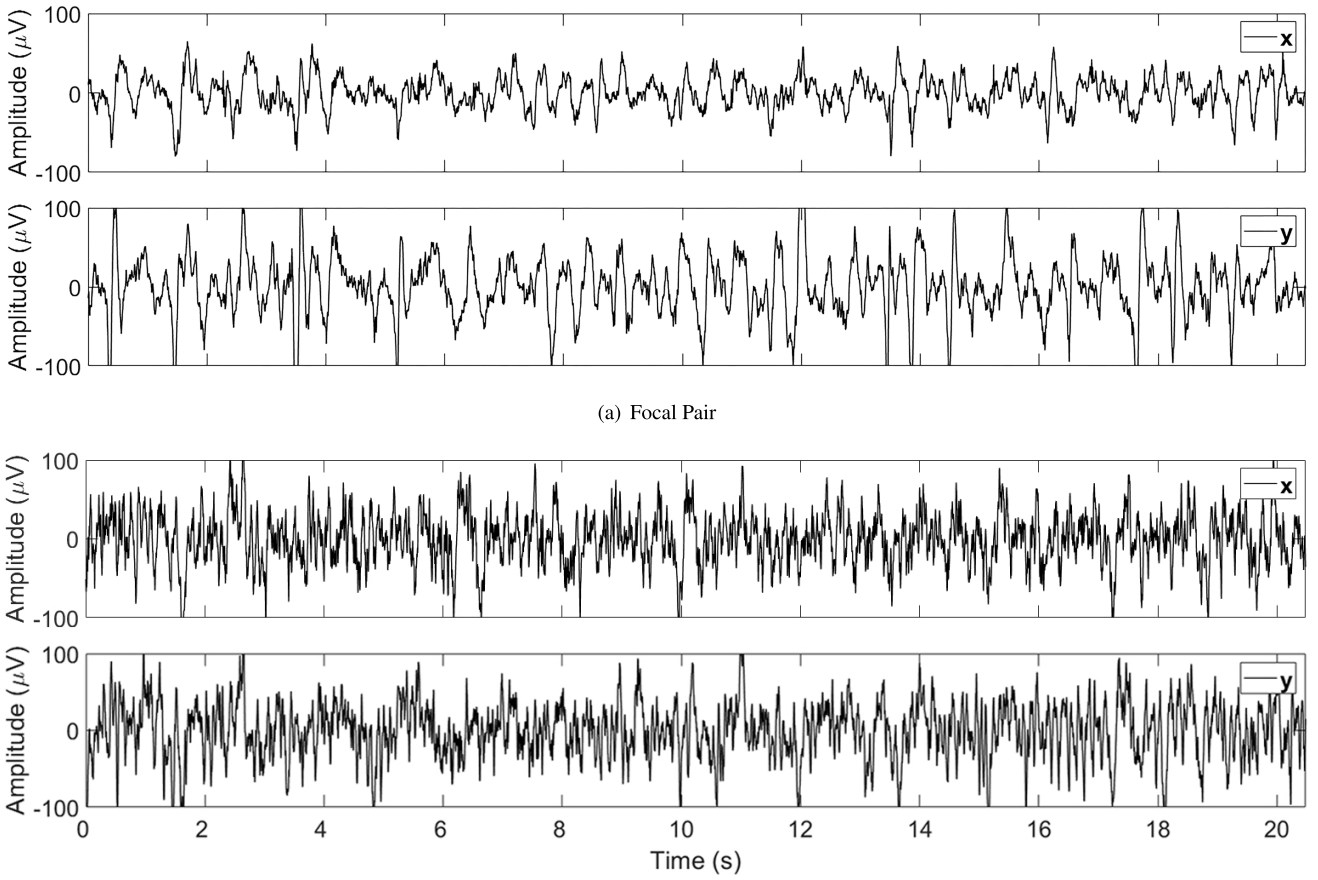
$$y_t = W_{\overrightarrow{h}y}\overrightarrow{h^N} + W_{\overleftarrow{h}y}\overleftarrow{h^N} + b_y \quad (7)$$

where for all N levels of stack, $\overrightarrow{h^N}$ and $\overleftarrow{h^N}$ are the hidden layers output in the forward and backward directions, respectively. A shortcut view of the BDLSTM is illustrated in Fig. 2, where each LSTM block contains all the gates described earlier.

III. MATERIALS AND METHODS

A. DATA-SET

The data-set is obtained from Bern-Barcelona EEG Database; which is an intracranial EEG study of epileptic patients done at the Department of Neurology of the University of Bern, Bern, Switzerland [19]. Five patients with longstanding pharmacoresistant temporal lobe epilepsy were included in the study. All patients were enrolled for epilepsy surgery and EEG signal acquisition was part of the diagnostic procedure prior to the surgery. Three patients had a complete seizure freedom, two patients had auras with no post surgery seizures, and All five patients had successful surgical outcomes. For every EEG recording, multi-channel EEG signals were acquired using AD-TECH (Racine, WI,USA) intracranial strip and depth electrodes. The reference electrode was extracranial located between 10/20 Fz and Pz positions. All EEG signals were sampled at 512 or 1024 Hz based on the number of channels used (more or less than 64 channels). Each EEG signal was recorded for about 41.66 hours duration for the focal and non-focal seizure activity. Based on the intracranial recordings, brain seizures could be localized for all patients.



(a) Focal Pair

(b) Non-Focal Pair

FIGURE 3. Sample pairs (x and y channels) from the EEG recordings data-set: (a) Focal pair, (c) Non-Focal pair.

Focal EEG signals were recorded using channels that detect the first ictal signal and were judged by visual inspection of two board-certified electroencephalographers and neurologists. The remaining EEG channels were used to record the non-focal signals. The EEG recordings were randomly divided into 3750 pairs of signals name as x and y. For an x signal, one focal EEG channel was selected from any patient at random while for the corresponding y signal, one of this channel's neighboring focal channels was selected. The random selection of focal channels was done without replacement using a uniform random number generator. The same selection procedure was applied for the non-focal data-set. Prior to storing the EEG signals pairs, they were band-pass filtered (0.5-150 Hz) using a fourth-order Butterworth filter. In addition, they were visually inspected to insure that no artifacts took place within the recordings. It worth mentioning that no clinical selection criteria was applied such as the absence of presence of epileptiform activity. A sample of each of the focal and non-focal EEG recordings pairs from the x and y channels are shown in Figure 3.

B. DATASET PREPARATION

In this study, all signals pass through an algorithm to enhance their features prior to running through the training and classification process. The steps are,

- Z-score Normalization
- Savitzky-Golay Filtering

1) Z-SCORE NORMALIZATION

Z-score normalization is a common signal processing technique used to ensure a balanced view for the data. Usually, large trends in the signals dominates the smaller trends, thus, increasing the dynamic amplitude range of the signal. The normalization process yields signals of a mean (μ) of zero and standard deviation (σ) of 1 by forcing all features to follow a normal Gaussian distribution based on the following equation,

$$x = \frac{x - \mu}{\sigma} \tag{8}$$

It is found in literature that signal normalization enhance the overall classification process [20], [21].

2) SAVITZKY-GOLAY FILTERING

Savitzky-Golay (SG) filter is a finite impulse response filter commonly used to enhance the overall precision of the data [22]. Through convolution mechanism, the input signal is smoothed with a higher Signal-to-Noise (SNR) ratio. For each signal, a set of polynomial coefficients are obtained using a least-square method. These coefficients are

convoluted at each data point of the input signal to produce an enhanced smother signal [23]. The coefficients are usually obtained by fitting a low order polynomial to the input data following this equation,

$$Y_j = \sum_{i=-(m-1)/2}^{i=(m-1)/2} C_i y_{j+1} \frac{m+1}{2} \tag{9}$$

where Y_j is the smoothed output signal from the input signal nx_j, y_j at every $j = 1, 2, \ldots,$ n points, x is the independent variable, y is the dependent variable, and C_i is the set of m polynomial coefficients.

C. TRAINING AND CLASSIFICATION

The training and classification process was performed using both the x and y EEG channels and following a k-fold cross-validation scheme. The process was applied using 4-folds, 6-folds, and 10-folds to ensure several splits of the data when performing the training process.

The LSTM network architecture selected was the BDLSTM. In addition, five layers were used in the training network including input sequence, fully connected, softmax, and classification layers. The sequence length of the data was 10240 representing the number of elements. Furthermore, a fully connected layer of two classes was used to provide the output sequence to the softmax and classification layers.

D. EVALUATION METRICS

The evaluation metrics selected in this study include the accuracy, sensitivity, specificity, precision, and F1-score, and are defined as follows,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{11}$$

$$Specificity = \frac{TN}{TN + FP} \tag{12}$$

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

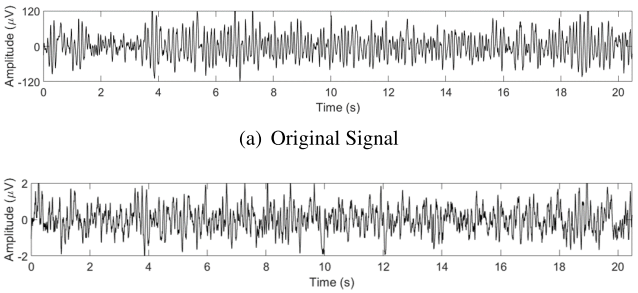
$$F1 - score = \frac{2TP}{2TP + FP + FN} \tag{14}$$

where true positives (TP) represents the total number of signals being classified correctly, true negatives (TN) represents the total number of the other signals being classified correctly as other signals, false positives (FP) represents the total number of signals classified incorrectly, and false negatives (FN) represents the total number of other signals classified incorrectly as other signals.

Furthermore, to evaluate the agreement between the observed and the expert classifications, Cohen's Kappa (κ) coefficient [24] is used following this equation,

$$\kappa = \frac{P_0 - P_c}{1 - P_c} \tag{15}$$

where P_0 the observed agreements and P_c represent the agreements expected by experts.



(a) Original Signal

(b) After Pre-processing

FIGURE 4. Sample pairs (x and y channels) from the EEG recordings data-set: (a) Focal pair, (c) Non-Focal pair.

To elaborate more on the observations of the LSTM classifier, the Matthews Correlation Coefficient (MCC) is a measure introduced by Matthews [25] in 1975 to measure the quality of binary classification, and it is calculated based on the following equation,

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{16}$$

In addition, the Jaccard Index (JI) [26] is included among the evaluation metrics and is often introduced to measure the similarities between two datasets or two classification observations. It is formulated as,

$$JI = \frac{TP}{TP + FP + FN} \tag{17}$$

IV. RESULTS AND DISCUSSION

A. CURRENT OBSERVATIONS

The proposed method was implement using MATLAB, where each of the 7500 EEG records were pre-processed as mentioned earlier. A sample from the non-focal EEG signals is shown in Fig. 4 before and after the pre-processing steps (normalization and SG filtering). The SG digital filter uses a window of length 25, averaging each 25 samples with a polynomial order of 3. The selected of these parameters was done after several trial and error tests.

To evaluate features between the two signal groups (focal and non-focal), Pearson correlation coefficients were calculated between two random recordings using the x and y EEG channel data and were observed as 0.01 and 0.04, respectively. These signals are used as an input for the BDLSTM network, where 10 cells are used to extract corresponding signals features [27]. Each time step is considered a feature to be utilized in the LSTM training network. Each cell includes a forward and a backward input/output stream as described earlier in Section II. At each time steps, having a 10240 time steps, the outputs of both streams are element-wise multiplied to obtain corresponding 10-dimensional representation, which are concatenated and fed to a 2-dimensional fully connected layer followed by a softmax activation function and a classification layer for predictions. The process of this feature extraction/learning process is illustrated briefly in Fig. 5.

TABLE 1. The overall performance of the developed algorithm in classifying FEEG and NFEEG signals using x-channels.

Scenario	Cohen's Kappa (κ)	MCC	JI	Accuracy	Classes	Sensitivity	Specificity	Precision	F1-score
4-fold	93.97%	93.97%	94.15%	96.99%	***fEEG***	96.99%	96.99%	96.99%	96.99%
					NfEEG	98.73%	99.14%	99.15%	98.94%
6-fold	94.29%	94.29%	94.45%	97.15%	***fEEG***	97.22%	97.07%	97.07%	97.14%
					NfEEG	97.07%	97.22%	97.23%	97.15%
10-fold	99.20%	99.20%	99.20%	99.60%	***fEEG***	99.55%	99.65%	99.65%	99.60%
					NfEEG	99.65%	99.55%	99.55%	99.60%

TABLE 2. The overall performance of the developed algorithm in classifying FEEG and NFEEG signals using y-channels.

Scenario	Cohen's Kappa (κ)	MCC	JI	Accuracy	Classes	Sensitivity	Specificity	Precision	F1-score
4-fold	97.47%	97.47%	97.62%	98.73%	***fEEG***	98.49%	98.98%	98.99%	98.74%
					NfEEG	98.98%	98.49%	98.48%	98.73%
6-fold	98.43%	98.43%	98.63%	99.21%	***fEEG***	99.07%	99.36%	99.36%	99.21%
					NfEEG	99.36%	99.07%	99.07%	99.21%
10-fold	98.80%	98.80%	98.80%	99.40%	***fEEG***	99.49%	99.31%	99.31%	99.40%
					NfEEG	99.31%	99.49%	99.49%	99.40%

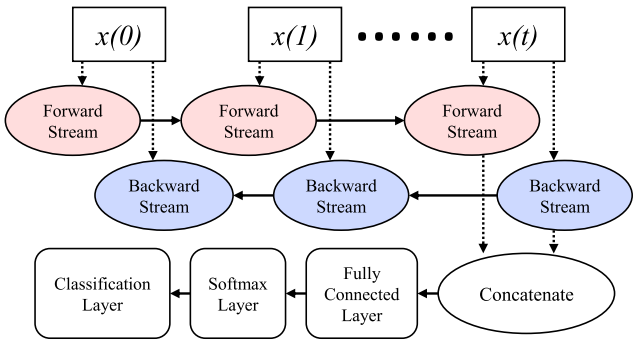


FIGURE 5. Training and classification network layers showing the forward and backward streams of the BDLSTM.

For the LSTM network settings, the Adaptive Moment Estimation (ADAM) solver was used as the LSTM optimizer for its ability to provide adaptive learning rates [28]. The solver parameters were 0.001, 0.9, and 0.999 for the learning rate (α), β_1, and β_2, respectively. The training network is selected to have a total number of 30 epochs with a mini-batch size of 50. The selection of these parameters was done after several trials.

Table 1 shows the overall performance of the developed algorithm in classifying the two EEG signal categories using the x-channels. The best k-fold scenario was the 10-fold cross-validation, as it covers more data in the training process. In addition, the Cohen's kappa (κ) value reached 99.20% with an accuracy of 99.60% in classification for both classes. All other metrics were higher than 99.00%, which reflects high sensitivity and precision. On the other hand, Table 2 shows the overall performance using the y-channels. A Cohen's kappa (κ) of 98.80% was observed with an accuracy of 99.40%. The values of the MCC and JI were close to the Cohen's kappa (κ) values following the two channel signals in 10-fold cross validation. This suggests strong similarity between the original expert classes and predicted classes using the LSTM classifier.

TABLE 3. The Confusion matrix of the original classes and the LSTM predicted classes using the x-channel signals.

		Original Classes		Total
		Focal	*Non-Focal*	
LSTM Classes	*Focal*	**3737**	17	3750
	Non-Focal	13	**3733**	3750
Total		3750	3750	**7500**

TABLE 4. The Confusion matrix of the original classes and the LSTM predicted classes using the y-channel signals.

		Original Classes		Total
		Focal	*Non-Focal*	
LSTM Classes	*Focal*	**3724**	19	3750
	Non-Focal	26	**3731**	3750
Total		3750	3750	**7500**

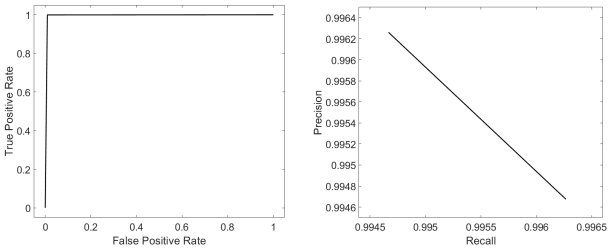
To elaborate more on the observations, the confusion matrix of the classifier when using the x-channels (highest κ value) is shown in Table 3. The TPs of each class is shown in the diagonal boxes. The LSTM algorithm successfully classified 3737 signal as focal and 3733 signal as non-focal. On the other hand, the algorithm wrongly classified 17 signals as non-focal and 13 as focal. For the confusion matrix of the y-channels shown in Table 4, the observations is close to the classification process of the x-channels, however, the number of correctly classified classes is less. Both tables provided comparable results which suggest the possibility of

TABLE 5. Summary table of the recent studies in classifying focal and non-focal EEG signals.

Study	Year	Extracted Features	Method	Overall Performance			Additional Information
				Accuracy	*Sensitivity*	*Specificity*	
Deivasigamani et al. [9]	2016	Dual Tree Complex Wavelet Transform (DT-CWT) features	ANFIS	99.00%	98.00%	100.00%	EEG signals were decomposed using the DT-CWT
Singh et al. [29]	2017	Discrete Fourier Transform (DFT) features	LS-SVM	89.52%	-	-	AUC = 97.89%
Sriraam et al. [30]	2017	Time, frequency, information theory, and statistically based features	Optimized SVM	92.15%	94.56%	89.74%	AUC = 94.83%
Gupta et al. [31]	2017	Cross entropy, log energy entropy, and SURE entropy based features	LS-SVM	94.41%	93.25%	95.57%	-
Sharma et al. [32]	2018	Bivariate Empirical Mode Decomposition (BEMD) based features	LS-SVM	84.01%	-	-	EEG signals were decomposed using the BEMD
Sharma et al. [33]	2019	Frequency domain (Bispectrum) based features	SVM	96.20%	95.46%	96.96%	Data reduction using Locality Sensitive Discriminant Analysis (LSDA) Positive predictive value (PPV) = 97.01% Negative predictive value (NPV) = 95.39% AUC = 96.00%
This Study	**2020**	**Bi-directional LSTM based features**	**BDLSTM**	**99.60%**	**99.65%**	**99.55%**	**AUC = 99.24%**

using both groups in the classification process. In addition, the high value of correct predictions of the LSTM classifier shows efficient outcomes in the discrimination between focal and non-focal EEG signals.

In addition, the Receiver Operating Characteristic (ROC) curve and the corresponding Precision-Recall plot are shown in Figure 6 (a) and (b), respectively, for the x-channel classification process. The algorithm achieved an Area Under the Curve (AUC) of 99.24%, a Positive Predictive Value (PPV) of 99.47%, and a Negative Predictive Value (NPV) of 99.63%. The Precision-Recall plot shows that the algorithm returned relevant results more than irrelevant ones with a high precision and recall of >99% at different thresholds.



(a) ROC Curve (b) Precision-Recall

FIGURE 6. The ROC curve and Precision-Recall plot for the classifier output using the x-channels and 10-fold cross-validation scheme: (a) ROC curve, (b) Precision-Recall.

B. STATE-OF-ART STUDIES

To investigate the observations found in this study relative to other studies, several researchers implemented machine learning algorithms for the purpose of focal and non-focal EEG classification. Table 5 provides a brief summary of the recent research works in this area. All researches covered in the table utilized the famous Bern-Barcelona database [19] described in section III. The table shows different machine algorithms including Adaptive Neuro Fuzzy Inference System (ANFIS), Least-Squares Support Vector Machine (LS-SVM), Optimized SVM, and regular SVM. In addition, a couple of these research works required further decomposition of the EEG signals using Dual Tree Complex Wavelet Transform (DT-CWT) and Bivariate Empirical Mode Decomposition (BEMD). The proposed algorithms did not require features extraction or signals decomposition step prior to the training and classification process. The current study showed close to literature observations with a slightly higher values in the averaged overall performance. The overall accuracy of the algorithms have reached 99.60% with high sensitivity of 99.65% and specificity of 99.55% using the x-channels and 10-fold cross-validation scheme. It worth noting that the studies covered in the summary table did not use deep learning algorithms and required manual feature extraction from the EEG signals.

V. CONCLUSION

In this paper, a study is conducted to investigate the use of an LSTM classifier to distinguish between focal and non-focal EGG signals of patients with epilepsy. The study showed high levels of accuracy of 99.60% in the classification process using a 10-fold cross-validation scheme. The higher accuracy was obtained when using x-channels and showed a high agreement with experts classification of 99.20%. Both the x and y EEG channels provided comparable results and suggest the use of both channels to discriminate focal and non-focal EEG signals. The study suggest LSTM as a potential deep learning algorithm in clinical EEG signals identification. Future work includes improvements on the architecture of the network with further testings on different epilepsy data-sets.

REFERENCES

[1] M. M. Rahman, M. I. H. Bhuiyan, and A. B. Das, "Classification of focal and non-focal EEG signals in VMD-DWT domain using ensemble stacking," *Biomed. Signal Process. Control*, vol. 50, pp. 72–82, Apr. 2019.
[2] World Health Organization. (2018). *Epilepsy*. [Online]. Available: http://www.who.int/mentalhealth/neurology/epilepsy/en
[3] A. F. Hussein, N. Arunkumar, C. Gomes, A. K. Alzubaidi, Q. A. Habash, L. Santamaria-Granados, J. F. Mendoza-Moreno, and G. Ramirez-Gonzalez, "Focal and non-focal epilepsy localization: A review," *IEEE Access*, vol. 6, pp. 49306–49324, 2018.

[4] S. Pati and A. V. Alexopoulos, "Pharmacoresistant epilepsy: From pathogenesis to current and emerging therapies," *Cleveland Clinic J. Med.*, vol. 77, no. 7, pp. 457–467, Jul. 2010.
[5] T. Ono, J. Wagenaar, F. S. Giorgi, P. Fabera, R. Hanaya, J. Jefferys, J. T. Moyer, L. C. Harte-Hargrove, and A. S. Galanopoulou, "A companion to the preclinical common data elements and case report forms for rodent EEG studies. A report of the TASK3 EEG working group of the ILAE/AES joint translational task force," *Epilepsia Open*, vol. 3, pp. 90–103, Nov. 2018.
[6] S. Raghu and N. Sriraam, "Classification of focal and non-focal EEG signals using neighborhood component analysis and machine learning algorithms," *Expert Syst. Appl.*, vol. 113, pp. 18–32, Dec. 2018.
[7] O. K. Fasil, R. Rajesh, and T. M. Thasleema, "Influence of differential features in focal and non-focal EEG signal classification," in *Proc. IEEE Region Humanitarian Technol. Conf. (R-HTC)*, Dec. 2017, pp. 646–649.
[8] A. Bhattacharyya, M. Sharma, R. B. Pachori, P. Sircar, and U. R. Acharya, "A novel approach for automated detection of focal EEG signals using empirical wavelet transform," *Neural Comput. Appl.*, vol. 29, no. 8, pp. 47–57, Apr. 2018.
[9] S. Deivasigamani, C. Senthilpari, and W. H. Yong, "Classification of focal and nonfocal EEG signals using ANFIS classifier for epilepsy detection," *Int. J. Imag. Syst. Technol.*, vol. 26, no. 4, pp. 277–283, Dec. 2016.
[10] S. Chatterjee, S. Pratiher, and R. Bose, "Multifractal detrended fluctuation analysis based novel feature extraction technique for automated detection of focal and non-focal electroencephalogram signals," *IET Sci., Meas. Technol.*, vol. 11, no. 8, pp. 1014–1021, Nov. 2017.
[11] W. Zhou, Y. Liu, Q. Yuan, and X. Li, "Epileptic seizure detection using lacunarity and Bayesian linear discriminant analysis in intracranial EEG," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3375–3381, Dec. 2013.
[12] R. Hussein, H. Palangi, R. K. Ward, and Z. J. Wang, "Optimized deep neural network architecture for robust detection of epileptic seizures using EEG signals," *Clin. Neurophysiol.*, vol. 130, no. 1, pp. 25–37, Jan. 2019.
[13] N. Michielli, U. R. Acharya, and F. Molinari, "Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals," *Comput. Biol. Med.*, vol. 106, pp. 71–81, Mar. 2019.
[14] P. Nagabushanam, S. T. George, and S. Radha, "EEG signal classification using LSTM and improved neural network algorithms," *Soft Comput.*, 2019, doi: 10.1007/s00500-019-04515-0.
[15] Z. Cui, R. Ke, Z. Pu, and Y. Wang, "Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction," 2018, *arXiv:1801.02143*. [Online]. Available: http://arxiv.org/abs/1801.02143
[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[17] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds. Red Hook, NY, USA: Curran Associates, 2014, pp. 3104–3112.
[18] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," 2015, *arXiv:1503.04069*. [Online]. Available: http://arxiv.org/abs/1503.04069
[19] R. G. Andrzejak, K. Schindler, and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 86, no. 4, Oct. 2012, Art. no. 046206.
[20] X. Zhang, L. Yao, D. Zhang, X. Wang, Q. Z. Sheng, and T. Gu, "Multi-person brain activity recognition via comprehensive EEG signal analysis," in *Proc. 14th EAI Int. Conf. Mobile Ubiquitous Syst. Comput. Netw. Services*, 2017, pp. 28–37.
[21] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," 2019, *arXiv:1901.05498*. [Online]. Available: http://arxiv.org/abs/1901.05498
[22] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares Procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.
[23] R. Schafer, "What is a Savitzky–Golay filter," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111–117, Jul. 2011.
[24] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, Apr. 1960.
[25] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochimica Biophysica Acta (BBA)-Protein Struct.*, vol. 405, no. 2, pp. 442–451, Oct. 1975.
[26] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Reading, MA, USA: Addison-Wesley, May 2005. [Online]. Available: http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0321321367
[27] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Netw.*, vol. 18, nos. 5–6, pp. 602–610, Jul. 2005.
[28] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[29] P. Singh and R. B. Pachori, "Classification of focal and nonfocal EEG signals using features derived from Fourier-based rhythms," *J. Mech. Med. Biol.*, vol. 17, no. 7, Nov. 2017, Art. no. 1740002.
[30] N. Sriraam and S. Raghu, "Classification of focal and non focal epileptic seizures using multi-features and SVM classifier," *J. Med. Syst.*, vol. 41, no. 10, p. 160, Oct. 2017.
[31] V. Gupta, T. Priya, A. K. Yadav, R. B. Pachori, and U. R. Acharya, "Automated detection of focal EEG signals using features extracted from flexible analytic wavelet transform," *Pattern Recognit. Lett.*, vol. 94, pp. 180–188, Jul. 2017.
[32] R. Sharma and R. B. Pachori, "Automated classification of focal and non-focal EEG signals based on bivariate empirical mode decomposition," in *Biomedical Signal and Image Processing in Patient Care*. Hershey, PA, USA: IGI Global, 2018, pp. 13–33.
[33] R. Sharma, P. Sircar, and R. B. Pachori, "A new technique for classification of focal and nonfocal eeg signals using higher-order spectra," *J. Mech. Med. Biol.*, vol. 19, no. 1, Feb. 2019, Art. no. 1940010.

LUAY FRAIWAN (Senior Member, IEEE) received the bachelor's degree in electrical engineering from the Jordan University of Science and Technology (JUST), in 1993, the master's degree in medical imaging from the University of Technology of Compiègne, France, and the second master's and Ph.D. degrees in biomedical engineering from The University of Akron, OH, USA, in 2002 and 2005 respectively. He is currently working as a Professor of biomedical engineering at Abu Dhabi University (ADU), on leave from JUST. Prior to joining ADU, he was an Associate Professor of biomedical engineering with JUST. He was the Department Chair, from 2011 to 2013. He has a solid research experience in the field of biomedical signal processing and sleep analysis.

MOHANAD ALKHODARI received the bachelor's degree (Hons.) in electrical engineering from Abu Dhabi University (ADU), United Arab Emirates, in 2017, the master's degree in biomedical engineering from the American University of Sharjah (AUS), United Arab Emirates, in 2019. In summer 2017, he joined the Microsystems Lab (iMicro) at the Masdar Institute, United Arab Emirates, working on ultrasound systems design as a Research Assistant. While pursuing his master's degree at AUS, he worked as a Teaching and Research Assistant at the College of Engineering under a Full Assistantship granted by the university to distinguished students. He was a Representative of the Electrical and Biomedical Engineering Master's programs for the IEEE Engineering in Medicine and Biology Society (EMBS) Student Chapter at AUS. His current research interests are biomedical signal processing, medical imaging, and deep/machine learning.