Deep Classification of Epileptic Signals

David Ahmedt-Aristizabal[1], Clinton Fookes[1], Kien Nguyen[1], Sridha Sridharan[1]

***Abstract*— Electrophysiological observation plays a major role in epilepsy evaluation. However, human interpretation of brain signals is subjective and prone to misdiagnosis. Automating this process, especially seizure detection relying on scalp-based Electroencephalography (EEG) and intracranial EEG, has been the focus of research over recent decades. Nevertheless, its numerous challenges have inhibited a definitive solution. Inspired by recent advances in deep learning, here we describe a new classification approach for EEG time series based on Recurrent Neural Networks (RNNs) via the use of Long-Short Term Memory (LSTM) networks. The proposed deep network effectively learns and models discriminative temporal patterns from EEG sequential data. Especially, the features are automatically discovered from the raw EEG data without any pre-processing step, eliminating humans from laborious feature design task. Our light-weight system has a low computational complexity and reduced memory requirement for large training datasets. On a public dataset, a multi-fold cross-validation scheme of the proposed architecture exhibited an average validation accuracy of 95.54% and an average AUC of 0.9582 of the ROC curve among all sets defined in the experiment. This work reinforces the benefits of deep learning to be further attended in clinical applications and neuroscientific research.**

I. INTRODUCTION

Epilepsy is a neurological disorder characterised by frequent and unpredictable seizures. Prior to epilepsy diagnosis, patients are usually monitored using a broad range of information from neuroimaging and electrophysiological methods [1]. Electroencephalography (EEG) has long been considered a gold standard for the diagnosis of seizures. The goal of the epilepsy evaluation is to delineate the brain network affected. However, this network could comprise other networks, which are involved in originating interictal epileptiform discharges and producing the first clinical manifestation of a seizure [2]. Misjudgement of the location of these networks causes ineffective clinical decisions. Despite recent advances in developing automated seizure detection devices [3], none of them is universally accepted because the performance in clinical scenarios has not been satisfactory. Significant work is still needed to reach expert-level evaluation, especially in understanding the epileptiform activities [4], [5], and by generalizing representations that are invariant to inter- and intra-subject differences. The performance of the traditional detection approaches relies heavily on expert knowledge to design the signal features employed and regularly include frequency-based features such as the wavelet transform, and energy analysis [6]. However, there is no warranty that these hand-crafted features are optimal for the chosen task, especially in the complex scenario of brain electrical activity. A major question to be asked is whether the feature engineering can be conducted automatically to discover the optimal features directly from the data, without the need for human-expert knowledge, and domain knowledge.

In this paper we conduct investigations to ascertain if the recent advances in deep learning could be the answer to this question. Deep learning is a subset of the machine learning techniques which emulates structures and operations of a human brain through a hierarchical multiple-layer signal representation coupled with advanced training algorithms [7]. The major advantage of deep learning in comparison with traditional machine learning is that the spatial, spectral and temporal feature representation is automatically learned from the training data, not by human assumption, leading to natural and effective signal representation and superior performance [7]. Deep learning has revolutionised many computer vision and medical applications, including the classification of brain signals [3], [8], [9]. Additionally, these deep architectures have been proposed for the tasks of seizure detection during the processing of EEG recordings for epilepsy diagnosis [10], [11], [12], [13]. Despite their benefits, there are two major limitations of these existing approaches: 1) Current methods either pre-process the raw data into some other forms before being fed into a deep learning architectures such as the Convolutional Neural Network (CNN); and 2) they use very deep and complex networks which have millions of parameters to be trained [14] and require very large training datasets, which are usually not available in the clinical scenarios.

In order to address these limitations, we investigate the plausibility of using deep learning architectures that are capable of both abstracting high-order features with limited training data and classifying them according to the physiological brain state and achieve state-of-the-art performance. We propose a light-weight Long Short Term Memory (LSTM) network that retains the benefits of deep models. Our system achieve high performance with fast run-time and reduced need for large datasets. Unlike currently used machine learning methods for EEG, the proposed network processes the raw data directly, without any transformation to the original EEG recordings and exploit the temporal patterns through the use of LSTMs. By automatically exploiting and discovering features from the temporal data, the proposed network can extract robust and reliable patterns to classify epileptic signals. The remainder of this paper is organised as follows: Section II presents the dataset and the methodology, Section III illustrates and discusses the results. Finally, Section IV reviews and concludes the paper.

[1] David Ahmedt-Aristizabal, Prof. Clinton Fookes, Dr. Kien Nguyen and Prof. Sridha Sridharan are with Speech, Audio, Image and Video Technologies (SAIVT) research program, School of Electrical Engineering & Computer Science, Queensland University of Technology, Australia.
Contact: david.aristizabal@hdr.qut.edu.au

978-1-5386-3646-6/18/$31.00 ©2018 IEEE

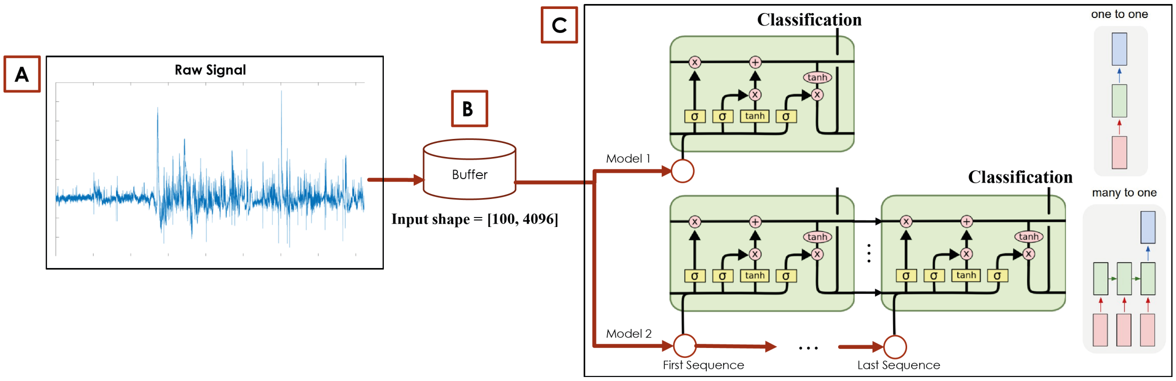

Fig. 1. The proposed deep framework to classify brain electrical activity. **A.** The raw samples for each type of brain-state are concatenated without a pre-processing. The value of amplitude of the signal is considered as a single representation of the segment size. **B.** The temporal evolution of the signal is analysed using the complete length of the signal, which indicates a total number of 4,096 segments and 100 samples for each type. **C.** The feature sequence is fed to a Long-Short-Term-Memory (LSTM) structure to exploit the temporal relation between segments and to predict brain-states signals. Two LSTM models were considered to analyse the EEG activity.

II. MATERIALS AND METHODS

A. Dataset

The experimental data we have used to validate our system is from the publicly available dataset from the Department of Epileptology, University of Bonn [15]. The dataset includes five sets (denoted from A to E) with a total of 100 EEG samples for each set. Each sample is a single channel EEG recorded at 173.6 Hz with 23.6 seconds of duration. Thus, the sample length of each sample is 4,096. Set A and B were recorded using scalp EEG from five healthy volunteers (healthy state) with eyes open and closed respectively. Set C, D and E, from five epileptic patients prior to surgery diagnosed with Temporal Lobe Epilepsy, were recorded using depth electrodes implanted symmetrically into the hippocampal formation. Set C and D were during seizure-free intervals, where set D was recorded from the epileptogenic zone (Inter-Ictal state or between seizures) and set C from the opposite brain hemisphere. Finally, set E described the recordings of the epileptogenic zone during an epileptic seizure (Ictal state).

B. The proposed system

The aim of this research is to compare properties of brain electrical activity from different recording regions and from different pathological brain states, *i.e.* classify healthy, inter-ictal and ictal EEG signals. To achieve this, we propose a deep framework which receives the raw EEG signals and extracts temporal features using an end-to-end training scheme based on a recurrent deep learning model known as Long Short-Term Memory (LSTM) architecture [16]. Unlike conventional signal-processing techniques where the features are hand-crafted and the signals are pre-processed, our method automatically learns the inherent characteristics of seizure data. The block diagram of the proposed deep learning system is displayed in Fig. 1.

An LSTM is a deep learning network which has evolved from the well known Recurrent Neural Network (RNN), especially suited for sequential data such as EEG signals since their neurons contain connections (weights) not only between the successive layers but also to themselves, which are used to memorise information from previous inputs. With a special gated mechanism, LSTM networks are capable of learning long-term dependencies present in sequential data and could predict time series when there are long time lags of unknown size between important events [17]. RNNs are called recurrent because they perform the same task for every element of a sequence, with the output being depended on the previous computations. Another way to think about RNNs is that they have a memory which captures information about what has been calculated so far. However, RNNs perform poorly when dealing with long sequences due to its frequently encountered drawback in gradient vanishing and exploding. LSTMs seek to address this issue by using a gated mechanism. Three gates, *i.e.* forget, input and output gates, are used to control the flow of information. The amount of information that is let through each gate is controlled by a point-wise multiplication and sigmoid function. For a system with input x_t, an output y_t and a hidden state h_t, a conventional RNN is constructed by defining the transition function and the output function as,

$$h_t = \phi_b(W^T h_{t-1} + U^T x_t), y_t = \phi_o(V^T h_t), \tag{1}$$

where W, U and V are the transition, input and output matrices respectively and ϕ_b and ϕ_o are element-wise nonlinear functions. Sigmoid or a hyperbolic tangent function are common examples of nonlinear functions. When the forget and input gates have determined how much information of the previous cell state C_{t-1} and the new cell state candidate $\hat{C}_t$ should be let through, the dynamic equations to represent the LSTM is given as,

$$\begin{aligned} \hat{C}_t &= tanh(W^T(r_t * h_{t-1}) + U^T x_t) \\ z_t &= \sigma_b(W_z^T h_{t-1} + U_z^T x_t + V_z^T C_{t-1}) \\ C_t &= f_t * C_{t-1} + i_t * \hat{C}_t \\ h_t &= o_t * \phi_b(C_t) \end{aligned} \tag{2}$$

where $z = \{i, f, o, r\}$, representing the gating functions: input gate, the forget gate, the output gate and the internal gate, and is the Sigmoid function. The trainable model parameters are: $\{W, W_z, U, U_z, V_z\}$.

The number of LSTM layers is one significant hyper-parameter to consider in the LSTM network. The model one-to-one indicates that from one single layer, the model estimates one single output. On the other hand, the model many-to-one, refers to multiple stack LSTMs that infer one output. The architectures are selected according to the performance of the model for each pair-set. Table I displays the specific LSTM models adopted in the classification process. We obtained the best performance with a network configured with one single layer with 64 hidden units (Model 1) and with 2 hidden layers of 128 and 64 hidden units, respectively (Model 2). For each model, we perform classification using a soft-max layer. We tested more complex architectures but the performance gain is not significant. More complicated architectures have more capability to model complicated signals, but practical clinical implementation would be affected; hence light-weight architectures with one or two layers could yield very accurate results in the experimental data. Therefore, our models are lightweight, with on the order of less than 17,000 trainable parameters in the case of Model 1. Once the model has been trained, the temporal features that lead to one or another prediction depending on brain state are extracted. These features illustrate the specific structures that should exist in a signal to trigger a specific classification.

TABLE I
LSTM ARCHITECTURES

	Model 1: One to One		**Model 2: Many to One**	
Layer Type	Output	Parameters	Output	Parameters
Input	(4097,1)		(4097,1)	
$LSTM_1$	(4097,1)	16,896	(4097,128)	66,560
$Dropout_1$			(4097,128)	
$LSTM_2$			(64)	49,408
$Dropout_2$			(64)	
$Dense_1$	(-1)	65	(1)	65
Total		**16,961**		**116,033**

III. EXPERIMENTS

A. Experimental setup

The proposed network is employed to classify six pairs of EEG recordings. These pairs are illustrated in Table II. For instance, the classification between set A and E refers to the verification of healthy volunteers with eyes open and ictal EEG signals. The complete temporal sequence for each set has an input shape of $[100, 4096]$. This illustrates 100 samples, each of them with 4096 segments.

We adopted a k-fold cross-validation to verify the generalization and robustness of the proposed architecture. The k-fold cross-validation [18] allow us to confirm the reliability of the model to predict data that have not been seen during training. For this evaluation, the samples of each set are randomly split into 70% for training, 20% for validation and 10% for testing. The difference between the validation and test samples is that the last one is not seen during the training phase. The validation and test accuracy of the framework is computed as the average performance of each fold (10-folds in this experiment). The performance of the classification task can also be expressed using sensitivity, specificity, precision and the area under the curve (AUC) values.

Training of the LSTM networks is carried out by optimizing the binary cross entropy loss function. The model is optimised with the ADAM optimizer with a learning factor of 10^{-3}, and decay rate of first and second moments as 0.9 and 0.999, respectively. Batch size set to 4 and dropout with a probability of 0.35, for Model 2, are considered to reduce the overfitting in deep neural networks when dealing with a small training data. We perform the model training using 20 epochs and use the default initialization parameters from Keras [19] for initializing the weights of the LSTM hidden units.

B. Experimental results

The multi-fold cross-validation average performance is displayed in Table II. The deep framework was capable of achieving an average of 95.54% in the validation accuracy and an average area under the curve of 0.9582 between all the sets pairs. The validation accuracy and error of the training process are shown in Fig. 3 and Fig. 4, respectively. This demonstrates that the learned features showed clear differences in dynamical properties of brain electrical activity from different physiological brain states.

We can see that the proposed framework achieves a significantly high accuracy of classification with the proposed light-weight deep learning architecture, which has a low computational cost (*e.g.* 4.5 sec average of training time and 200MB of RAM on a 2.6GHz CPU for each set-pair). As illustrated the Table II, this performance outperformed the results of a deep learning model based on the same data reported in [13], where 90% of the data was used for training in comparison with only 70% of the data in our training. Additionally, our results have reached similar high performance compared to state-of-the-art algorithms based on powerful feature extraction techniques and robust classifiers [20], which rely on specific expert knowledge and manual extraction of data. In the experiments, the highest accuracy is obtained with the pair Set A-E, while the lowest is Set A-D. This result was expected because the dynamical properties of the signals from the epileptogenic zone between seizures are more similar to healthy EEG segments than to ictal signals.

IV. CONCLUSIONS

We have investigated the benefits of a recurrent deep learning framework to classify EEG segments from epileptic signals. We adopt LSTM networks to extract temporal patterns in the frame sequences. Experimental results, confirms that our computationally efficient models can achieve a very high degree of accuracy. The proposed approach demonstrates the capability of recurrent models to learn a general representation of a seizure event directly and automatically from the raw data. The fast run-time coupled with extremely sparse use of computing resources makes our model desirable for real-time use, for such low computation devices as wearable sensors, which could enhance the diagnosis and treatment planning for patients that experience epilepsy. In

TABLE II

MULTI-FOLD CROSS-VALIDATION PERFORMANCE (10-TIME AVERAGE)

Sets	Type Model	Validation Accuracy (%)	Test Accuracy (%)	Test Sensitivity (%)	Test Specificity (%)	Test Precision (%)	AUC	Validation Accuracy (%) [13]
A and E	1	**99.50**	**97.00**	96.00	98.00	98.09	**0.9820**	**95.50**
B and E	1	94.75	92.50	91.00	94.00	94.27	0.9850	92.50
C and E	1	97.25	92.00	95.00	89.00	90.06	0.9650	91.67
D and E	1	96.50	91.00	95.00	87.00	89.06	0.9510	93.34
A and D	2	**90.25**	**82.00**	82.00	82.00	84.78	**0.9030**	**86.42**
B and D	2	95.00	93.00	92.00	93.00	93.00	0.9630	N.A
Average		**95.54**	**91.25**	91.83	90.50	91.50	0.9582	

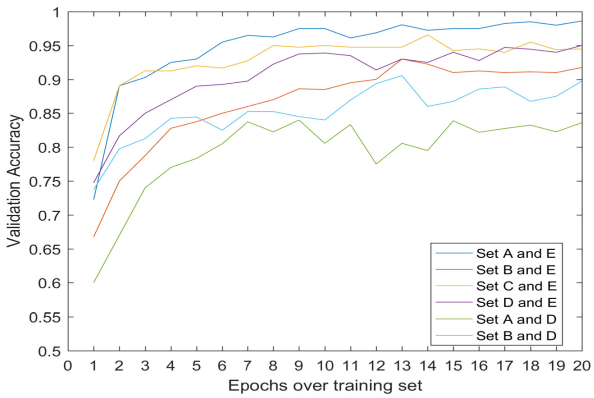

Fig. 2. Validation accuracy performance of all sets. (Best in color).

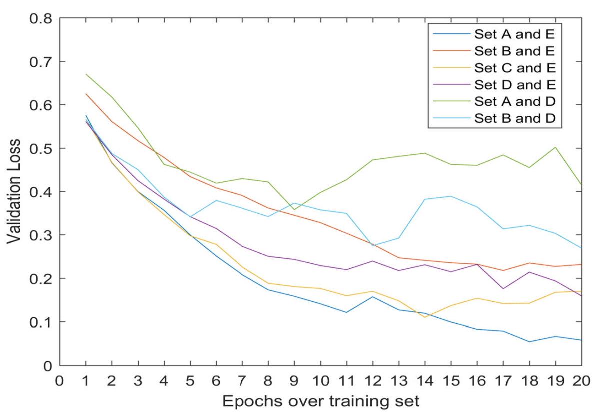

Fig. 3. Validation error performance of all sets. (Best in color).

future work, we will investigate the use the temporal features extracted from the LSTM architecture to predict seizure events before it happens, thereby able to alert the patients and clinical staff to be aware of the event about to happen. We plan also to further analyse complex signals from invasive monitoring such as stereo EEG recordings.

V. ACKNOWLEDGEMENT

The authors acknowledge the support of the QUT-HPC for providing the computational resources for this research.

REFERENCES

[1] F. Rosenow and H. Lüders, "Presurgical evaluation of epilepsy," *Brain*, vol. 124, no. 9, pp. 1683–1700, 2001.

[2] P. Chauvel and A. McGonigal, "Emergence of semiology in epileptic seizures," *Epilepsy & Behavior*, vol. 38, pp. 94–103, 2014.

[3] D. Ahmedt-Aristizabal, C. Fookes, S. Dionisio, K. Nguyen, J. P. S. Cunha, and S. Sridharan, "Automated analysis of seizure semiology and brain electrical activity in presurgery evaluation of epilepsy: A focused survey," *Epilepsia*, vol. 58, no. 11, pp. 1817–1831, 2017.

[4] F. Wendling, P. Benquet, F. Bartolomei, and V. Jirsa, "Computational models of epileptiform activity," *Journal of neuroscience methods*, vol. 260, pp. 233–251, 2016.

[5] A. Antoniades, L. Spyrou, C. C. Took, and S. Sanei, "Deep learning for epileptic intracranial eeg data," in *Machine Learning for Signal Processing (MLSP), 2016 IEEE 26th International Workshop on*. IEEE, 2016, pp. 1–6.

[6] S. Ramgopal, S. Thome-Souza, M. Jackson, N. E. Kadish, I. S. Fernández, J. Klehm, W. Bosl, C. Reinsberger, S. Schachter, and T. Loddenkemper, "Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy," *Epilepsy & behavior*, vol. 37, pp. 291–307, 2014.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[8] L. Bozhkov, "Overview of deep learning architectures for classifying brain signals," *KNOWLEDG EE SOCIET YY*, p. 54.

[9] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[10] P. Thodoroff, J. Pineau, and A. Lim, "Learning robust features using deep learning for automatic seizure detection," in *Machine Learning for Healthcare Conference*, 2016, pp. 178–190.

[11] L. Vidyaratne, A. Glandon, M. Alam, and K. M. Iftekharuddin, "Deep recurrent neural network for seizure detection," in *Neural Networks (IJCNN), 2016 International Joint Conference on*. IEEE, 2016, pp. 1202–1207.

[12] A. Page, C. Shea, and T. Mohsenin, "Wearable seizure detection using convolutional neural networks with transfer learning," in *Circuits and Systems (ISCAS), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 1086–1089.

[13] Q. Lin, S.-q. Ye, X.-m. Huang, S.-y. Li, M.-z. Zhang, Y. Xue, and W.-S. Chen, "Classification of epileptic eeg signals with stacked sparse autoencoder based on deep learning," in *International Conference on Intelligent Computing*. Springer, 2016, pp. 802–810.

[14] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals," *Computers in biology and medicine*, 2017.

[15] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6, p. 061907, 2001.

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[17] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "Lstm: A search space odyssey," *IEEE transactions on neural networks and learning systems*, 2017.

[18] R. Kohavi *et al.*, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Ijcai*, vol. 14, no. 2. Stanford, CA, 1995, pp. 1137–1145.

[19] F. Chollet *et al.*, "Keras (2015)," 2017.

[20] Y. Li, X.-D. Wang, M.-L. Luo, K. Li, X.-F. Yang, and Q. Guo, "Epileptic seizure classification of eegs using time–frequency analysis based multiscale radial basis functions," *IEEE journal of biomedical and health informatics*, vol. 22, no. 2, pp. 386–397, 2018.