The 9th IEEE International Conference on E-Health and Bioengineering - EHB 2021
Grigore T. Popa University of Medicine and Pharmacy, Web Conference, Romania, November 18-19, 2021

Transfer Learning of EEG for Analysis of Interictal Epileptiform Discharges

Cheong Took[1], S. Alty[1], D. Martin-Lopez[2], A. Valentin[3], G. Alarcon[3], S. Sanei[4]
[1]Royal Holloway, University of London TW20 0EX UK, {clive.cheongtook, steve.alty}@rhul.ac.uk
[2]Kingston Hospital London SE5 9RS UK, davidneurofisio@gmail.com
[3]King's College Hospital London WC2R 2LS UK, antonio.valentin@kcl.ac.uk, galarcon@aol.com
[4]Nottingham Trent University NG11 8NS UK, saeid.sanei@ntu.ac.uk

***Abstract*—Analysis of EEG requires years of clinical training and mentorship. To alleviate the human cost involved in EEG analysis, we propose a general LSTM-Autoencoder-CNN for EEG (GLACE) framework, which is adequately general to facilitate the use of transfer learning in smart healthcare. Traditionally in transfer learning, only the last few layers of the neural network are changed and adapted to the new task. Instead, we focus on the adaptation of the first layers to each new task. We exploit the inter-trial couplings in our proposed deep learning approach called GLACE. The efficacy of GLACE was assessed against a real-world clinical problem, i.e. the detection of interictal epileptiform discharges; GLACE circumvents the need for the neurophysiologist to spend hours on EEG analysis. Simulations show that the adaptation of the first layers of the trained model leads to an accuracy improvement of 12%.**

Keywords — EEG, interictal epileptiform discharge (IED), asymmetric autoencoders, long short-term memory, transfer learning

I. Introduction

Electroencephalography (EEG) is a gold standard in healthcare for the diagnosis of epilepsy, sleep disorders, disorders of consciousness, to name a few. EEG typically involves electrodes capturing the brain electrical activities. The 10-20 international EEG system can have as many as 256 electrodes, which means the number of electrodes depends on the EEG application. The number of electrodes used varies for each EEG acquisition system, e.g. the OpenBCI Ganglion caters for 4 electrodes, the intracranial EEG typically has 6-13 electrodes, and higher resolution EEG system such as the 10-5 system uses over 300 electrodes [1].

These varying number of electrodes pose a problem for transfer learning of EEG. In fact, transfer learning usually involves *freezing* the first layers of a trained neural network and adapting its last layers to the problem at hand. In other words, it is almost a prerequisite that the number of input neurons (EEG electrodes) required to learn from the original task is the same as that of the new task. If this is not the case, then the common approach is to reduce or increase the input dimension of the trained neural network to match the number of input neurons required for a new task. However, this common approach has problems. On one hand, a decrease in input neurons means less data and may cause the deep network to effectively run out of data. On the other hand, an increase in input neurons may require additional layers to achieve reasonable accuracy, leading to lengthy trial-and-error experimentation before the optimal network is found. To circumvent these problems, we propose a deep neural network (DNN) based on the transfer learning from scalp EEG to intracranial EEG for epilepsy. The availability of scalp EEG means that in most cases it is the only accessible way for health monitoring in epilepsy.

Epilepsy is a chronic brain disorder that can affect people at any age [2]. It causes recurrent and erratic alterations in brain functionality due to abnormal excessive or synchronous neuronal activity in the brain, called epileptic seizure. The seizure arises owing to the dysfunction of the electrophysiological system of the brain and uncontrolled electrical discharges in a group of neurons in the cerebral cortex [3]. It is associated with abnormal patterns of cortical activation, known as epileptiform brain activity, which can be captured by the EEG. Nonetheless, scalp EEG suffers from low sensitivity in capturing epileptiform discharges and, consequently, around 30% to 40% of patients considered for epilepsy surgery require invasive intracranial EEG recording [4]. As a result, increasing the sensitivity of sEEG for epilepsy diagnosis and management as a low cost noninvasive approach becomes very important. Moreover, findings from sEEG are very useful in presurgical assessment to decide if and where to implant iEEG electrodes. Therefore, developing an effective method for identification of interictal epileptiform discharges (IEDs) from over the scalp can directly alleviate the need for intracranial EEG.

II. EEG Deep Learning

A. Dataset and problem statement

The details of the dataset used to demonstrate the efficacy of our proposed general LSTM-Autoencoder-CNN for EEG (GLACE) algorithm are provided in this section. The dataset is comprised of two parts: 20 channels of scalp EEG and the 12 simultaneously recorded channels of intracranial EEG sampled at 200 Hz. There are N segments for each pair of EEG, called `trials'. Fig. 5 provides a colour map to illustrate the inter-trial

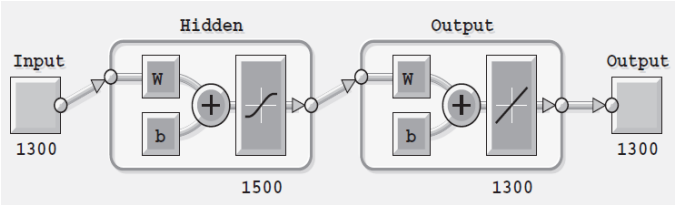


Fig. 1. Autoencoder 1: 1300 input neurons correspond to the 20 channels of scalp EEG (20 channels × 65 samples).

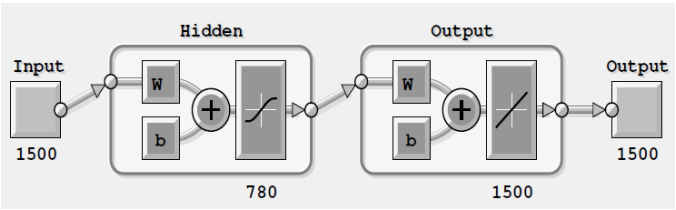


Fig. 2. Autoencoder 2: 780 neurons correspond to the 12channels of intracranial EEG (12 channels × 65 samples)

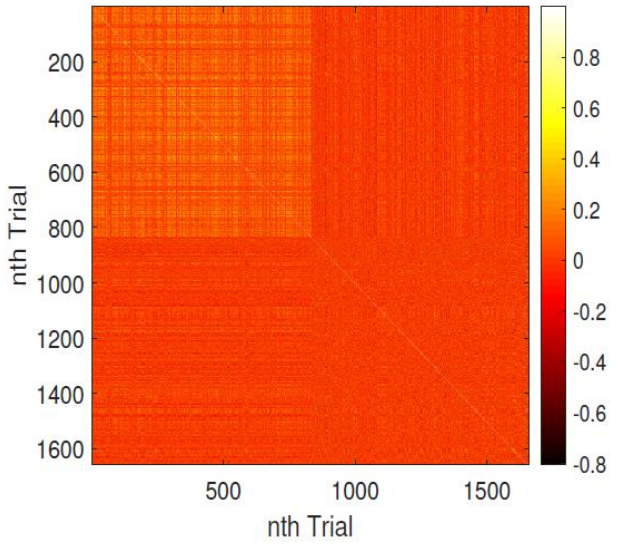


Fig. 5. Colour map showing the correlation between different trials, where the most correlated ones are amongst the first ~ 800 trials.

correlation for the dataset – which is exploited in our work. More details on this dataset can be found in [5]-[12].

Remark: Notice the strong degree of correlation amongst the first ~800 trials in Fig. 5. Both these short and long term temporal dependencies can be exploited using long short-term memory (LSTM). As expected, the maximum correlation is along the diagonalline due to the nth trial autocorrelation.

The machine learning problem is two fold: 1) Estimation of interictal epileptiform discharge (IED) enhanced scalp EEG; 2) Classification of the EEG neural signature into IED and non-IED waveforms.

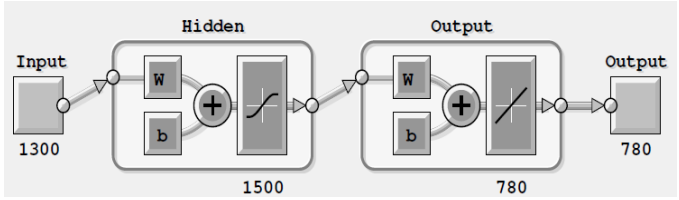


Fig. 3. Encoder 1 + Encoder 2 = Asymmetric Autoencoder, where the number of input neurons ≠ output neurons.

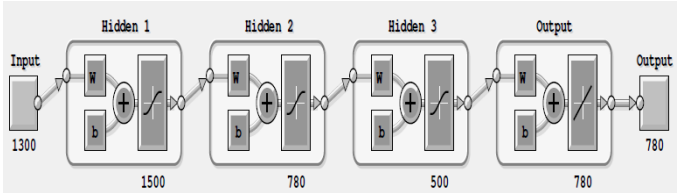


Fig. 4. Final stacked autoencoders to map scalp to intracranial EEG

B. Autoencoders to map scalp onto intracranial EEG

It is well established that autoencoders can be exploited to pre-train DNNs such as the one considered in Fig. 4. In this direction, the asymmetric autoencoders are employed, whereby the number of input neurons is not equal to the number of output neurons. We map the input from one domain to another, i.e. from scalp EEG to the corresponding intracranial sources. To achieve this, two autoencoders depicted in Fig. 1-2 are first trained, and their weights are used to initialise the asymmetric autoencoder depicted in Fig. 3. Each trial lasts 325 ms, which corresponds to the 65 samples. The 65-sample segments are adequately long to cover the occurrence of a typical IED waveform.

Remark: In the asymmetric autoencoder, the hidden layer is greater than the input so that the inputs can be broken down into smaller ones; this break down of inputs facilitates the learning of the output layer, provided adequate training is performed. For the symmetric autoencoder, it is well established that the hidden layer must be smaller than the input layer, so that the output layer can `ignore' the noise by taking the dimension reduction approach.

C. LSTM for inter-trial learning

We take advantage of the inter-trial dependencies illustrated in Fig. 3 by using long short-term memory NNs. For completeness, this section provides a brief description of LSTM. LSTM consists of many layers comprised of cells. The long term memory is implicitly captured by the cell state whose output $\mathbf{c}(t)$ is given by

$$\mathbf{c}(t) = \mathbf{f}(t) * \mathbf{c}(t-1) + \mathbf{i}(t) * \mathbf{g}(t) \qquad (1)$$

where the current cell state $\mathbf{c}(t)$ depends on the past cell state $\mathbf{c}(t-1)$ and $*$ denotes elementwise multiplication. The newly processed information $\mathbf{g}(t)$ is computed as

$$\mathbf{g}(t) = \tanh(\mathbf{x}(t),\mathbf{h}(t-1)) \qquad (2)$$

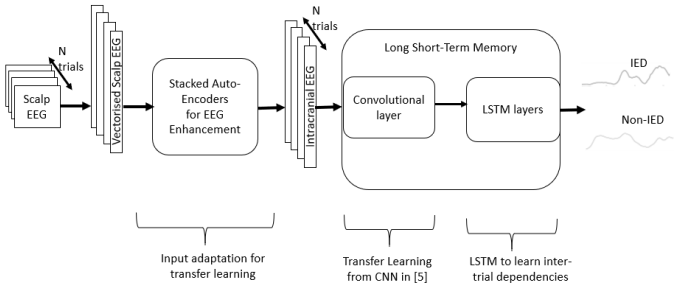


Fig. 6: Overall architecture: From EEG mapping to IED classification.

h(t-1) is the hidden output from the previous cell. In fact, the hidden state **h**(t) implements the short-term memory as:

$$\mathbf{h}(t) = \mathbf{o}(t) * \tanh(\mathbf{c}(t)) \quad (3)$$

Gates control the two outputs of a cell via sigmoidal functions with range [0,1] such as **o**(t) in (3), and **f**(t), **i**(t) in (1). More details on LSTM can be found in [14]. Convolutional neural networks (CNNs) can extract the temporal information in EEG [15]. Zhu incorporated convolutional layers in LSTM for feature extraction [16]; we adopt this approach in our proposed GLACE to extract one feature vector per trial.

D. Overall architecture

The overall deep learning architecture is shown in Fig. 6. More specifically, the main layers comprised of

1) one layer of asymmetric autoencoder to map scalp to intracranial EEG;

2) one layer of symmetric autoencoder to consolidate estimation of intracranial EEG;

3) one layer of convolutional layer to extract temporal local features within each trial;

4) one layer of LSTM to exploit inter-trial correlation.

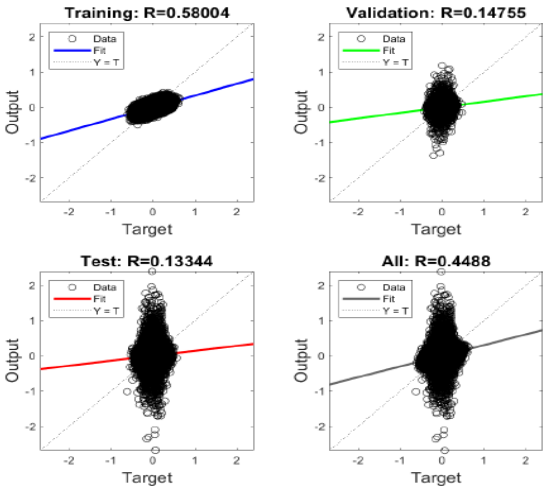
This deep learning architecture is concluded with a fully connected layer to optimise the LSTM estimate followed by one layer of softmax for classification.

III. SIMULATIONS

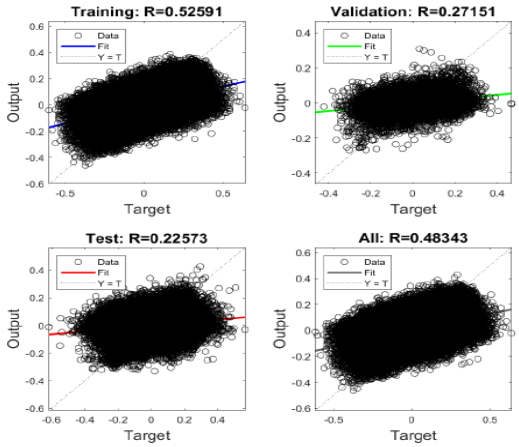
To verify the proposed GLACE, comprehensive simulations were conducted using the EEG dataset considered in [5]-[6]. The data was divided into training data (84%), validation data (1%), and testing data (15%). The experiments were to evaluate (i) the performance of proposed stacked autoencoders detailed in Section IIB and (ii) the proposed LSTM-based GLACE method on the classification of IED. Our previous work [6] was used to assess the efficacy of our proposed GLACE method as a benchmark algorithm. Other benchmark results can be found in [6].

A. Experiment 1: Estimation of Intracranial EEG

The regression plots in Fig. 7 compare the proposed GLACE with the method in [6]. The more aligned the data is on the x =y axis, the better the intracranial EEG estimates.



(a) 2-stage trained autoencoders considered in [6].



(b) Layer-by-layer trained autoencoders in GLACE.

Fig. 7. Regression plots for mapping scalp to intracranial EEG.

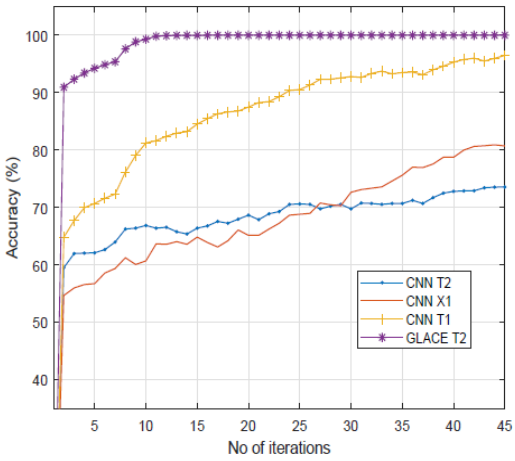


Fig. 8. Learning curves for the training data: CNN applied to scalp EEG.

TABLE I. CLASSIFICATION ACCURACY ON TEST DATA. X1, T1, AND T2 REFER RESPECTIVELY TO SCALP, INTRACRANIAL, AND ESTIMATED INTRACANIAL EEG.

Method	CNN X1	CNN T1	CNN T2	GLACE T2
Accuracy	61.04%	93.17%	74.44%	100%

Similarly, the closer to unity the value of R-statistics (Goodness-of-fit) is, the better the estimates. The performance measure R on the test data has almost doubled using GLACE (R = 0.23) compared to the convolutional neural network (R = 0.13) in [6]. Similarly, the best validation performance in terms of mean-squared-error was 0.028 for [6] compared to 0.014 for GLACE.

B. Experiment 2: Classification of interictal epileptiform discharges

First, we illustrate the advantage when using intracranial EEG (T1) over scalp EEG (X1). Second, we show the benefit of exploiting the correlation between different trials by GLACE. The learning curves for training the convolutional neural network considered in [5] for scalp (X1), intracranial EEG (T1) and estimated intracranial EEG from GLACE (T2) are shown in Fig. 8. Moreover, the learning curve of our LSTM-based GLACE is also plotted in the same figure. Table I summarises the results for the test data.

Remark: Due to transfer learning, all methods start at an accuracy of at least 55% in Fig. 8 whether learning from scalp or intracranial EEG. Clearly, it is more straightforward for CNNs to learn from intracranial EEG than scalp EEG, due to the quicker convergence of 'CNN T1' and 'CNN T2'. Fig. 8 shows that the training accuracy of scalp EEG is better than the estimated intracranial T2, however, the test accuracy shows the superiority of learning from intracranial EEG (whether estimated or ground truth) in Table I. The existing inter-trial correlation boosted the performance of our proposed GLACE in both training in Fig. 8 and testing phase in Table I. On the other hand, CNN did reasonably well, but was still outperformed by GLACE.

IV. CONCLUSIONS

A unified deep learning framework has been proposed. LSTM, autoencoders, and convolutional layers are merged leading to a general mapping approach for linking two EEG modality recordings. In particular, an asymmetric autoencoder layer has been exploited to correlate the input layer from the old task to the new task in transfer learning. The convolutional layers extracts the temporal information from each trial, whereas the LSTM layers learn the inter-dependencies between different trials as well.

REFERENCES

[1] S. Sanei, *Adaptive Processing of Brain Signals*. John Wiley, 2013.

[2] R. S. Fisher, W. V. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. E. Jr, "Epileptic seizures and epilepsy: definitions proposed by the international league against epilepsy (ILAE) and the international bureau for epilepsy (IBE)," *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005.

[3] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri,"Automated EEG analysis of epilepsy: a review," *Knowledge-Based Systems*, vol. 45, pp. 147–165, 2013.

[4] S. Kovac, V. N. Vakharia, C. Scott, and B. Diehl, "Invasive epilepsy surgery evaluation," *Seizure*, vol. 44, pp. 125–136, 2017.

[5] A. Antoniades, L. Spyrou, D. Martin-Lopez, A. Valentin, G. Alarcon, S. Sanei, and C. Cheong Took, "Detection of interictal discharges with convolutional neural networks using discrete ordered multichannel intracranial EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2285–2294, 2017.

[6] ——, "Deep neural architectures for mapping scalp to intracranial EEG," *International journal of neural systems*, vol. 28, no. 8, p. 1850009, 2018.

[7] A. Antoniades, L. Spyrou, C. C. Took, and S. Sanei, "Deep learning for epileptic EEG data," *Proceedings of the IEEE Machine Learning for Signal Processing Workshop, Italy*, 2016.

[8] L. Spyrou, D. M. Lopez, G. Alarcon, A. Valentin, and S. Sanei, "Detection of intracranial signatures of interictal epileptiform discharges from concurrent scalp EEG," *International Journal of Neural Systems*, vol. 26, no. 4, 2016.

[9] L. Spyrou, S. Kouchaki, and S. Sanei, "Multiview classification and dimensionality reduction of EEG data through tensor factorisation," *Elsevier Journal of Signal Processing Systems*, pp. 1–12, 2016.

[10] L. Spyrou and S. Sanei, "Coupled dictionary learning for multimodal data: an application to concurrent intracranial and scalp EEG," *Proceedings of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Shanghai China*, 2016.

[11] B. Abdi-Sargezeh, A. Valentin, G. Alarcon, and S. Sanei, "Incorporating uncertainty in data labeling into automatic detection of interictal epileptiform discharges from concurrent scalp EEG via multi-way analysis," *International Journal of Neural Systems*, pp. 1–15, 2021.

[12] ——, "Detection of brain interictal epileptiform discharges from intracranial EEG by exploiting their morphology in the tensor structure," *Proceedings of European Signal Processing Conference*, 2021.

[13] V. Le Quoc, "A Tutorial on Deep Learning, Part 2: Autoencoders, Convolutional Neural Networks and Recurrent Neural Networks," *Lecture Notes*, 2015.

[14] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222 – 2232, 2017.

[15] S. Sakhavi, C. Guan, and S. Yan, "Learning Temporal Information for Brain-Computer Interface Using Convolutional Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5619–5629, 2018.

[16] G. Zhu, L. Zhang, L. Yang, L. Mei, S. Shah, M. Bennamoun, and P. Shen, "Redundancy and Attention in Convolutional LSTM for Gesture Recognition," *IEEE Transactions on Neural Network and Learning Systems*, pp. 1 – 13, June 2019.