**Deep Architectures for Spatio-Temporal Modeling:  
Automated Seizure Detection in Scalp EEGs**

Meysam Golmohammadi, Saeedeh Ziyabari, Vinit Shah, Iyad Obeid, Joseph Picone

Neural Engineering Data Consortium, Temple University, Philadelphia, Pennsylvania, USA  
{meysam, saeedeh, vinitshah, obeid, picone}@temple.edu

*Abstract*—Automated seizure detection using clinical electroencephalograms is a challenging machine learning problem because the multichannel signal often has an extremely low signal to noise ratio. Events of interest such as seizures are easily confused with signal artifacts (e.g., eye movements) or benign variants (e.g., slowing). Commercially available systems suffer from unacceptably high false alarm rates. Deep learning algorithms that employ high dimensional models have not previously been effective due to the lack of big data resources. In this paper, we use the TUH EEG Seizure Corpus to evaluate a variety of hybrid deep structures including Convolutional Neural Networks and Long Short-Term Memory Networks. We introduce a novel recurrent convolutional architecture that delivers 30% sensitivity at 7 false alarms per 24 hours. We have also evaluated our system on a held-out evaluation set based on the Duke University Seizure Corpus and demonstrate that performance trends are similar to the TUH EEG Seizure Corpus. This is a significant finding because the Duke corpus was collected with different instrumentation and at different hospitals. Our work shows that deep learning architectures that integrate spatial and temporal contexts are critical to achieving state of the art performance and will enable a new generation of clinically-acceptable technology.

Keywords—deep learning, EEG, seizure, long short-term memory, convolutional neural networks

# Introduction

Electroencephalograms (EEGs) are used in a wide range of clinical settings to record electrical activity along the scalp. Scalp EEGs are the primary means by which physicians diagnose brain-related illnesses such as epilepsy and seizures [1]. However, manual analysis of EEG signals requires a highly-trained board-certified neurologist and is a process that is known to have relatively low inter-rater agreement (IRA) . It is a time-consuming and expensive process since the volume of the data far exceeds the available resources for detailed interpretation in real time. Automated analysis can improve the quality of patient care by reducing manual error, thereby reducing the time to diagnosis and latency. In this paper, we focus on the specific problem of seizure detection, although the work presented here is also applicable to other EEG problems such as signal event detection and abnormal detection .

Like most machine learning problems of this nature, many algorithms have been applied, including time–frequency analysis methods and nonlinear techniques . Machine learning has made tremendous progress over the past three decades due to rapid advances in low-cost highly-parallel computational infrastructure, powerful machine learning algorithms, and, most importantly, big data. Although contemporary approaches for automatic interpretation of EEGs have employed more modern machine learning approaches such as neural networks and deep learning [7], state of the art machine learning algorithms that typically employ high dimensional models have not previously been utilized successfully in EEG analysis because there has been a lack of big data resources. Recently, however, a significant new data resource known as the TUH EEG Seizure Corpus (TUSZ) has become available for EEG interpretation , thereby creating a unique opportunity to advance the technology.

Despite much research progress, commercially available automated EEG analysis systems are impractical due to high false alarm rates . Servicing false alarms in critical care settings places too much of a cognitive burden on caregivers, and hence, the outputs from these systems are ignored . This creates quality of healthcare issues as well as cost and efficiency challenges.

The goal of this work is to demonstrate that advanced deep learning approaches that have been successful in tasks like image and speech recognition, where ample amounts of annotated training data are available, can be applied to EEG interpretation. To achieve this goal, we evaluated several popular approaches on a standard seizure detection task. We propose a novel deep learning architecture that reduces the false alarm rate while maintaining sensitivity and specificity. We demonstrate that the performance of this system is now approaching the threshold for clinical acceptance.

# Exploiting Spatio-Temporal Context

Spatial and temporal context are required for accurate disambiguation of seizures from artifacts . shows our generic architecture for processing EEG signals to acquire this context. The multichannel signal is sampled at 250 Hz using 16 bits of resolution, converted to a feature-based representation, processed through a sequential modeler, and then postprocessed using a variety of statistical models that impose constraints based on subject matter expertise. Several architectures that implement Gaussian Mixture Models (GMMs), hidden Markov model (HMMs) and deep learning (DL) have been evaluated.

Feature extraction, which is not the primary focus of this paper, typically relies on time frequency representations of the signal. Although it is possible to replace conventional feature extraction with deep learning-based approaches operating directly on sampled data, here we focus only on traditional cepstral-based features . The use of more advanced discriminative features has not yet produced substantial improvements in performance for this application. Our system uses a standard linear frequency cepstral coefficient-based feature extraction approach (LFCCs) [13]. We also use first and second derivatives of the features since these improve performance.

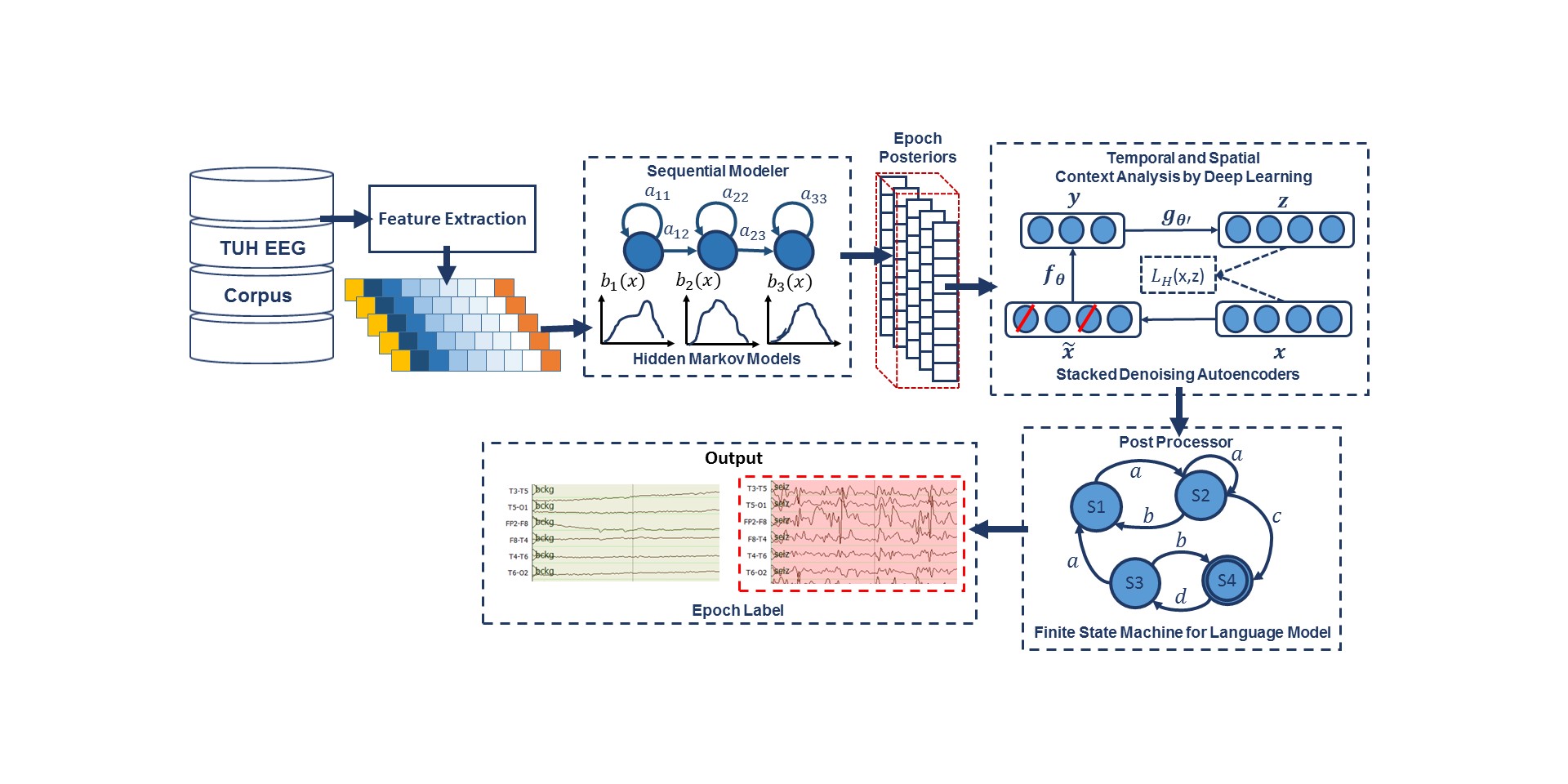


Figure 1: A hybrid architecture that integrates temporal and spatial context for sequential decoding of EEG events

Neurologists typically review EEGs in 10s windows and identify events with a resolution of approximately *1s*. We analyze the signal in *1s* epochs, further dividing this interval into *10* frames of *0.1s* (the frame duration), and use *0.2s* analysis windows (referred to as the window duration). The output of our feature extraction process is a feature vector of dimension *26* for each of *22* channels, with a frame duration of *0.1s*.

## Sequential Decoding Using HMMs

HMMs are powerful statistical modeling tools for signals that have both a time and frequency domain component [14]. In this study, a left-to-right channel-independent GMM-HMM, as illustrated in Figure 1, was used as a baseline system for sequential decoding. We divide each channel of an EEG into 1-second epochs, and further subdivide these epochs into a sequence of frames. Each epoch is classified using an HMM trained on the subdivided epoch, and then these epoch-based decisions are post-processed by additional statistical models in a process that parallels the language modeling component of a speech recognizer. Standard three-state left-to-right HMMs [14] with eight Gaussian mixture components per state were used to model each channel of the *22*-channel signal. A diagonal covariance matrix assumption was used for each mixture component. Channel-independent models were trained since channel-dependent models did not provide any improvement in performance.

Supervised training based on the Baum-Welch re-estimation algorithm was used to train two models – seizure and background. Models were trained on segments of data containing seizures based on manual annotations that are available as part of TUSZ. Since seizures comprise a small percentage of the overall data (*3%* in the training set; *8%* in the evaluation set), the amount of non-seizure data was limited to be comparable to the amount of seizure data, and non-seizure data was selected to include a rich variety of artifacts such as muscle and eye movements. Twenty iterations of Baum-Welch were used though performance is not very sensitive to this value. Standard Viterbi decoding (no beam search) was used in recognition to estimate the model likelihoods for every epoch of data (the entire file was not decoded as one stream because of the imbalance between the seizure and background classes – decoding was restarted for each epoch).

The output of the epoch-based decisions was postprocessed by a deep learning system. Our baseline system used a Stacked Denoising Autoencoder (SdA) [15] as shown in Figure 1. An N-channel EEG was transformed into *N* independent feature streams using a standard sliding window-based approach. The hypotheses generated by the HMMs were postprocessed using a second stage of processing that examines the temporal and spatial context. We apply a third pass of postprocessing that uses a stochastic language model to smooth hypotheses involving sequences of events so that we can suppress spurious outputs. This third stage provides a moderate reduction in false alarms.

The output of the first stage of processing is a vector of two likelihoods for each channel at each epoch. Therefore, if we have *22* channels, which is typical for an EEG collected using a standard *10/20* configuration, and *2* classes (seizure and background), we will have a vector of dimension *2 x 22 = 44* for each epoch. Each of these scores is independent of the spatial context (other EEG channels) or temporal context (past or future epochs). To incorporate context, we form a supervector consisting of *N* epochs in time using a sliding window approach. We find it beneficial to make *N* large – typically *41*. This results in a vector of dimension *1,804* that needs to be processed for each epoch. This input dimensionality is too high considering the amount of manually labeled data available for training and the computational requirements. To deal with this problem we used Principal Components Analysis (PCA) to reduce the dimensionality to *20* before applying the SdA postprocessing.

The parameters of the SdA model are optimized to minimize the average reconstruction error using a cross-entropy loss function. In the optimization process, a variant of stochastic gradient descent is used called “Mini-batch stochastic gradient descent” (MSGD). The SdA network has three hidden layers with corruption levels of *0.3* for each layer. The number of nodes per layer are: first layer = *800*, second layer = *500*, third layer = *300*. The parameters for pre-training are: learning rate = *0.5*, number of epochs = *150*, batch size = *300*. The parameters for fine-tuning are: learning rate = *0.1*, number of epochs = *300*, batch size = *100*. The overall result of the second stage is a probability vector of dimension *2* containing a likelihood that each label could have occurred in the epoch. A soft decision paradigm is used rather than a hard decision paradigm because this output is smoothed in the third stage of processing. A more detailed explanation about the third pass of processing is presented in [3].

## Two-Dimensional Decoding Using CNNs

Convolutional Neural Networks (CNNs) have delivered state of the art performance on highly challenging tasks such as speech and image recognition . Our overall architecture of a system that combines CNN and a multi-layer perceptron (MLP) is shown in , and is referred to as CNN/MLP. The network contains six convolutional layers, three max pooling layers and two fully-connected layers. A rectified linear unit (ReLU) non-linearity is applied to the output of every convolutional and fully-connected layer . Drawing on an image classification analogy, each image is a signal where the width of the image (*W*) is the window length multiplied by the number of samples per second, the height of the image (*H*) is the number of EEG channels and the number of image channels (*N*) is the length of the feature vector.

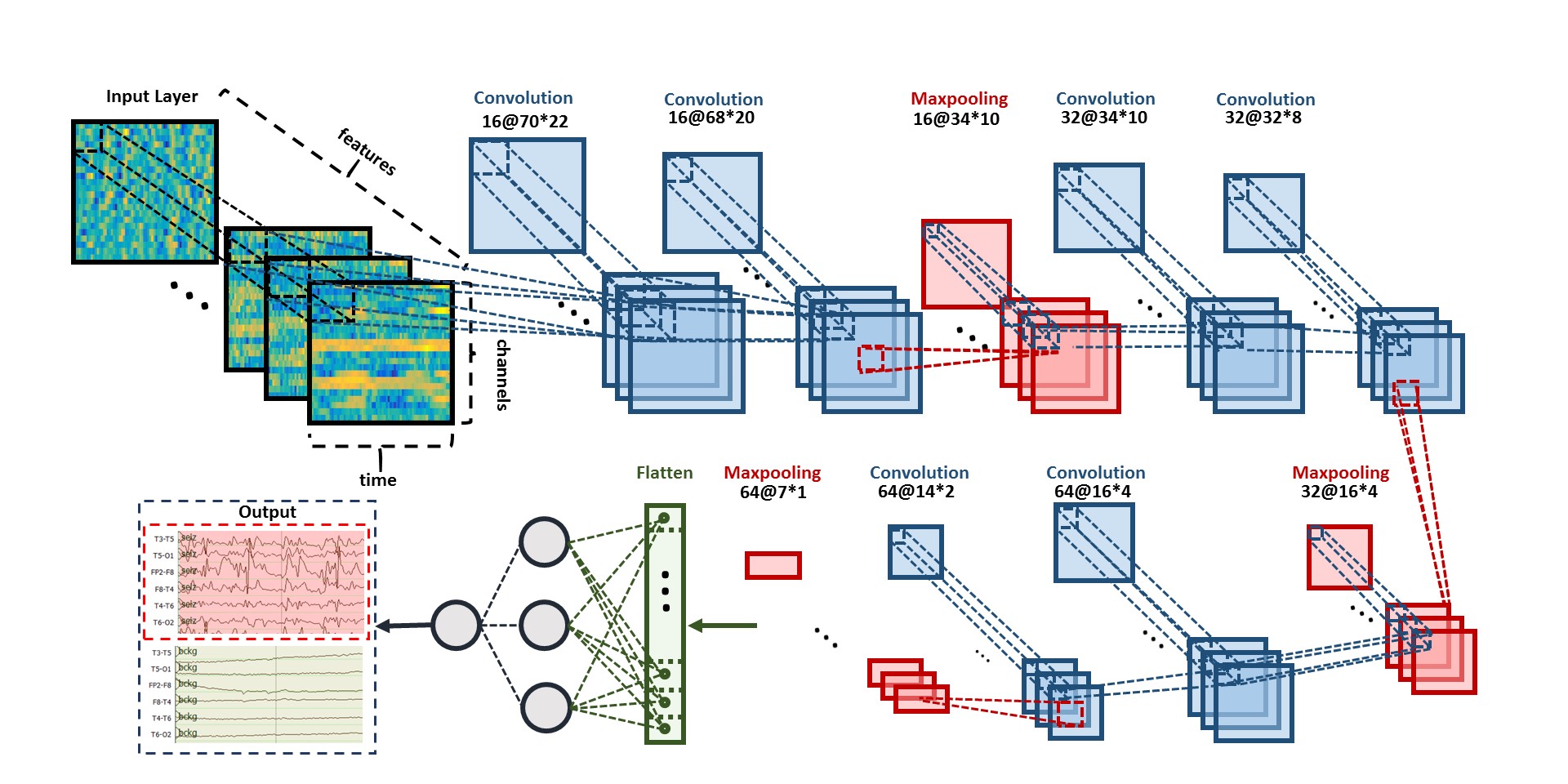


Figure 2: A two-dimensional decoding of EEG signals using a CNN/MLP hybrid architecture is shown that consists of  
six convolutional layers, three max pooling layers and two fully-connected layers.

In our optimized system, a window duration of *7* seconds is used. The first convolutional layer filters the input of size of *70 × 22 × 26* using *16* kernels of size *3 × 3* with a stride of *1*. The second convolutional layer filters its input using *16* kernels of size *3 × 3* with a stride of *1*. The first max pooling layer takes as input the output of the second convolutional layer and applies a pooling size of *2 × 2*. This process is repeated two times more with *32* and *64* kernels. Next, a fully-connected layer with *512* neurons is applied and the output is fed to a *2*-way sigmoid function which produces a two-class decision (the final epoch label).

## Recurrent Convolutional Neural Networks

In our final architecture, which is shown in Figure 3, we integrate 2D CNNs, 1-D CNNs and LSTM networks, which we refer to as a CNN/LSTM, to better exploit long-term dependencies. Note that the way that we handle data in CNN/LSTM is different from the CNN/MLP system presented in Figure 2. The input data is composed of frames distributed in time where for each frame *W* is the length of a feature vector, *H* is the number of EEG channels, and *N = 1*. The input data consists of *T* frames where *T* is equal to the window length multiplied by the number of samples per second. In our optimized system with a window duration of *21* seconds, the first 2D convolutional layer filters *210* frames (*T = 21 × 10*) of EEGs distributed in time with a size of *26 × 22 × 1* (*W = 26, H = 22, N = 1*) using *16* kernels of size *3 × 3* and with a stride of *1*. The first 2D max pooling layer takes as input a vector which is *260* frames distributed in time with a size of *26 × 22 × 16* and applies a pooling size of *2 × 2*. This process is repeated two times with two 2D convolutional layers with *32* and *64* kernels of size *3 × 3* respectively and two 2D max pooling layers with a *2 × 2* pooling size.

The output of the third max pooling layer is flattened to *210* frames with a size of *384 × 1*. Then a 1D convolutional layer filters the output of the flattening layer using *16* kernels of size *3* which decreases the dimensionality in space to *210 × 16*. We then apply a 1D maxpooling layer with a size of *8* to decrease the dimensionality to *26 × 16*. This is the input to a deep bidirectional LSTM network where the dimensionality of the output space is *128* and *256*. The output of the last bidirectional LSTM layer is fed to a *2*-way sigmoid function which produces a final classification of an epoch. To overcome the problem of overfitting and force the system to learn more robust features, dropout and Gaussian noise layers are used between layers [23]. To increase non-linearity, Exponential Linear Units (ELU) are used [24]. Adam is used in the optimization process along with a mean squared error loss function.

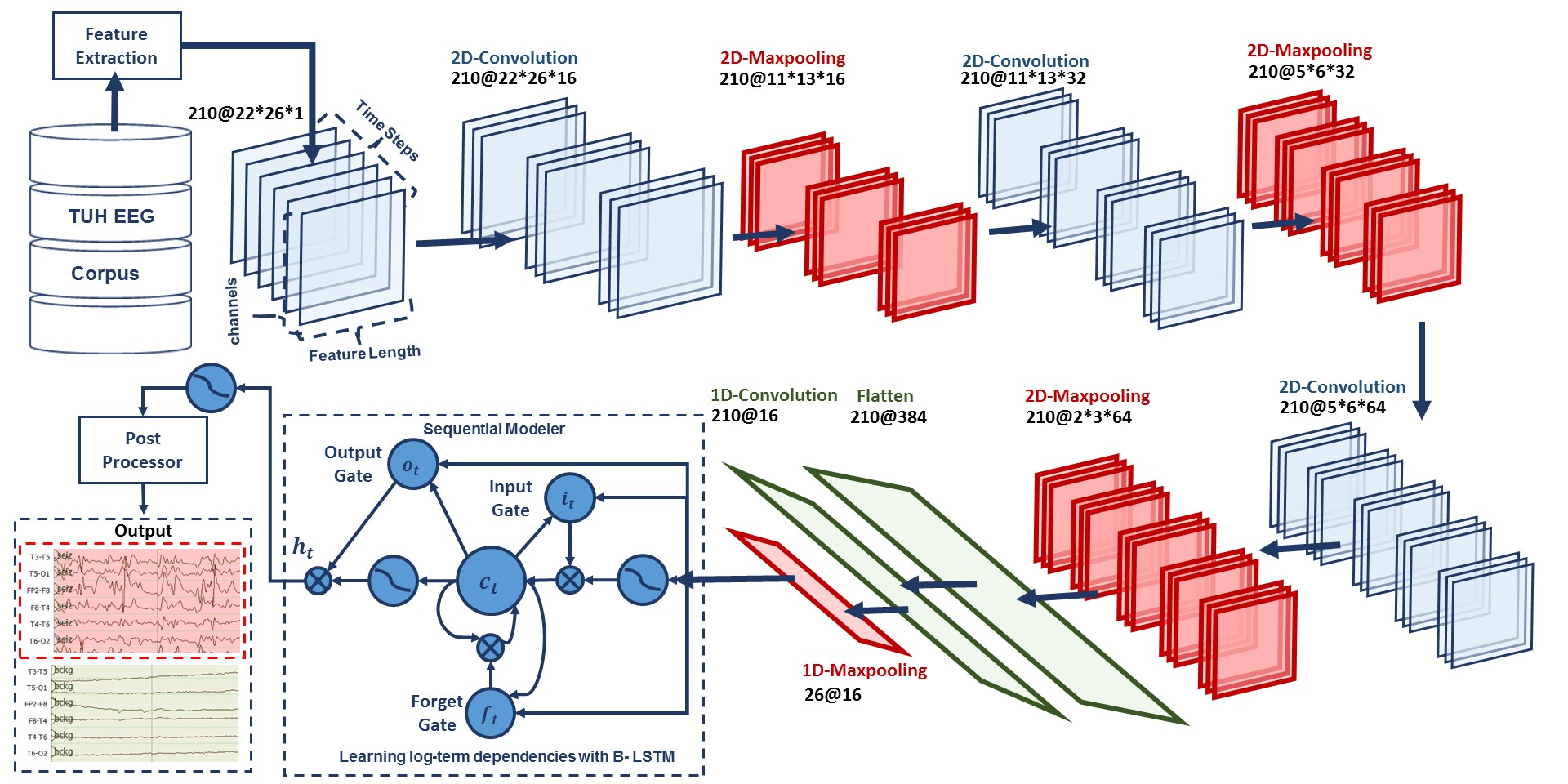


Figure 3: A deep recurrent convolutional architecture for two-dimensional decoding of EEG signals  
that integrates 2D CNNs, 1-D CNNs and LSTM networks is shown.

# EXPERIMENTS

The lack of big data resources that can be used to train sophisticated statistical models compounds a major problem in automatic seizure detection. Inter-rater agreement for this task is low, especially when considering short seizure events [1]. Manual annotation of a large amount of data by a team of certified neurologists is extremely expensive and time consuming. In this study, we are reporting results for the first time on TUSZ and a comparable corpus, the Duke University Seizure Corpus (DUSZ) [2]. TUSZ was used as the training and test set corpus, while DUSZ was used as a held-out evaluation set. It is important to note that TUSZ was collected using several generations of Natus EEG equipment, while DUSZ was collected using Nihon Kohden equipment. This is a true open-set evaluation since the data were collected under completely different recording conditions. A summary of these corpora is shown in Table 1.

A comparison of the performance of the different architectures presented in this paper, for sensitivity in range of *30%*, are shown in . The related DET curve is illustrated in Figure 4. These systems were evaluated using a method of scoring popular in the EEG research community known as the overlap method [25]. True positives (TP) are defined as the number of epochs identified as a seizure in the reference annotations and correctly labeled as a seizure by the system. True negatives (TN) are defined as the number of epochs correctly identified as non-seizures. False positives (FP) are defined as the number of epochs incorrectly labeled as seizure while false negatives (FN) are defined as the number of epochs incorrectly labeled as non-seizure. Sensitivity shown in is computed as TP/(TP+FN). Specificity is computed as TN/(TN+FP). The false alarm rate is the number of FPs per *24* hours.

It is important to note that the results are much lower than what is often published in the literature on other seizure detection tasks. This is due to a combination of factors including (1) the neuroscience community has favored a more permissive method of scoring that tends to produce much higher sensitivities and lower false alarm rates; and (2) TUSZ is a much more difficult task than any corpus previously released as open source. The evaluation set was designed to be representative of common clinical issues and includes many challenging examples of seizures.

Table 1: An overview of TUSZ and DUSZ

|  |  |  |  |
| --- | --- | --- | --- |
| **Description** | **TUSZ** | | **DUSZ** |
| **Train** | **Eval** | **Eval** |
| Patients | 64 | 50 | 45 |
| Sessions | 281 | 229 | 45 |
| Files | 1,028 | 985 | 45 |
| Seizure (secs) | 17,686 | 45,649 | 48,567 |
| Non-Seizure (secs) | 596,696 | 556,033 | 599,381 |
| Total (secs) | 614,382 | 601,682 | 647,948 |

Also, note that the HMM baseline system, which is shown in the first row of Table 2, operates on each channel independently. The other methods consider all channels simultaneously by using a supervector that is a concatenation of the feature vectors for all channels. The baseline HMM system only classifies epochs (*1s* in duration) using data from within that epoch. It does not look across channels or across multiple epochs when performing epoch-level classification. The results of the hybrid HMM and deep learning structures show that adding a deep learning structure for temporal and spatial analysis of EEGs can decrease the false alarm rate dramatically. Further, by comparing the results of HMM/SdA with HMM/LSTM, we find that a simple one-layer LSTM performs better than *3* layers of SdA due to LSTM’s ability to explicitly model long-term dependencies. Note that in this case the complexity and training time of these two systems is comparable.

The best overall system is CNN/LSTM. This doubly deep recurrent convolutional structure models both spatial relationships (e.g., cross-channel dependencies) and temporal dynamics (e.g., spikes). For example, CNN/LSTM does a much better job rejecting artifacts that are easily confused with spikes because these appear on only a few channels, and hence can be filtered based on correlations between channels. The depth of the convolutional network is important since the top convolutional layers tend to learn generic features while the deeper layers learn dataset specific features. Performance degrades if a single convolutional layer is removed. For example, removing any of the middle convolutional layers results in a loss of *4%* absolute in sensitivity.

Table 2: Performance on TUSZ

|  |  |  |  |
| --- | --- | --- | --- |
| **System** | **Sensitivity** | **Specificity** | **FA/24 Hrs.** |
| HMM | 30.32% | 80.07% | 244 |
| HMM/SdA | 35.35% | 73.35% | 77 |
| HMM/LSTM | 30.05% | 80.53% | 60 |
| CNN/MLP | 39.09% | 76.84% | 77 |
| CNN/LSTM | 30.83% | 96.86% | 7 |

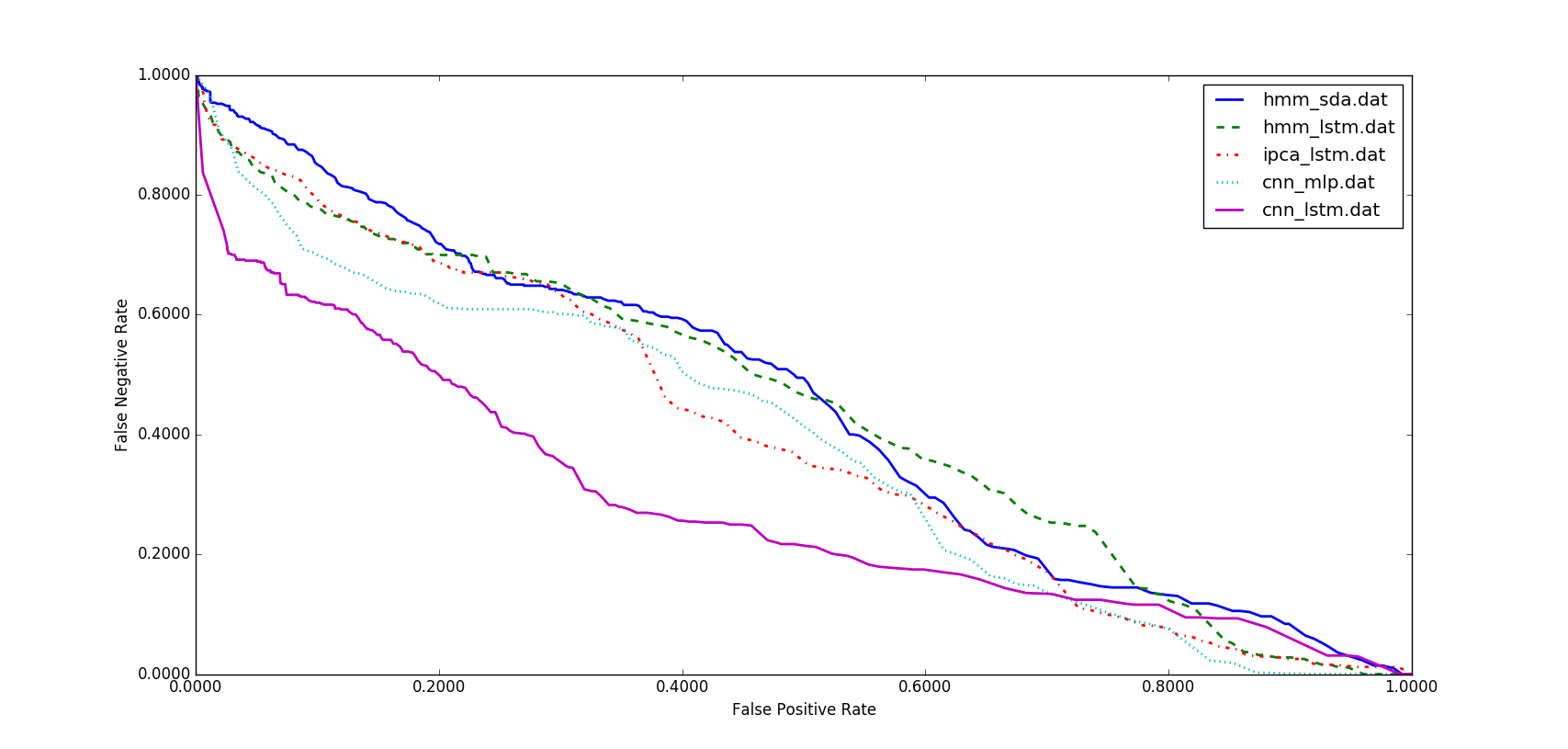


Figure 4: A DET curve comparing performance on TUSZ

We have also conducted an evaluation of our CNN/LSTM system on DUSZ. The results are shown in Table 3. A DET curve is shown in Figure 5. At high false positive rates, performance between the two systems is comparable. At low false positive rates, false positives on TUSZ are lower than on DUSZ. This suggests there is room for additional optimization on DUSZ.

In these experiments, we observed that the choice of optimization method had a considerable impact on performance. The results of our best performing system, CNN/LSTM, was evaluated using a variety of optimization methods. The best performance is achieved with Adam [20], using a learning rate of , a learning rate decay of *0.0001*, exponential decay rates of for the moment estimates and a fuzz factor of . The parameters follow the notation described in [20]. Similarly, we evaluated our CNN/LSTM using different activation functions. ELU delivers a small but measurable increase in sensitivity, and more importantly, a reduction in false alarms.

# Conclusions

In this paper, we introduced a variety of deep learning architectures for automatic classification of EEGs including a hybrid architecture that integrates CNN and LSTM. While this architecture delivers better performance than other deep structures, its performance still does not meet the needs of clinicians. Human performance on similar tasks is in the range of *65%* sensitivity with a false alarm rate of *12* per *24* hours [26]. The false alarm rate is particularly important to critical care applications since it impacts the workload experienced by healthcare providers.

The primary error modalities observed were false alarms generated during brief delta range slowing patterns such as intermittent rhythmic delta activity. A variety of these types of artifacts have been observed during inter-ictal and post-ictal stages. Training models on such events with diverse morphologies can significantly reduce the remaining false alarms. This is one reason we are continuing our efforts to annotate a larger portion of TUSZ.

Table 3: Performance of CNN/LSTM on DUSZ

|  |  |  |  |
| --- | --- | --- | --- |
| **Corpus** | **Sensitivity** | **Specificity** | **FA/24 Hrs.** |
| TUSZ | 30.83% | 96.86% | 7 |
| DUSZ | 33.71% | 70.72% | 40 |

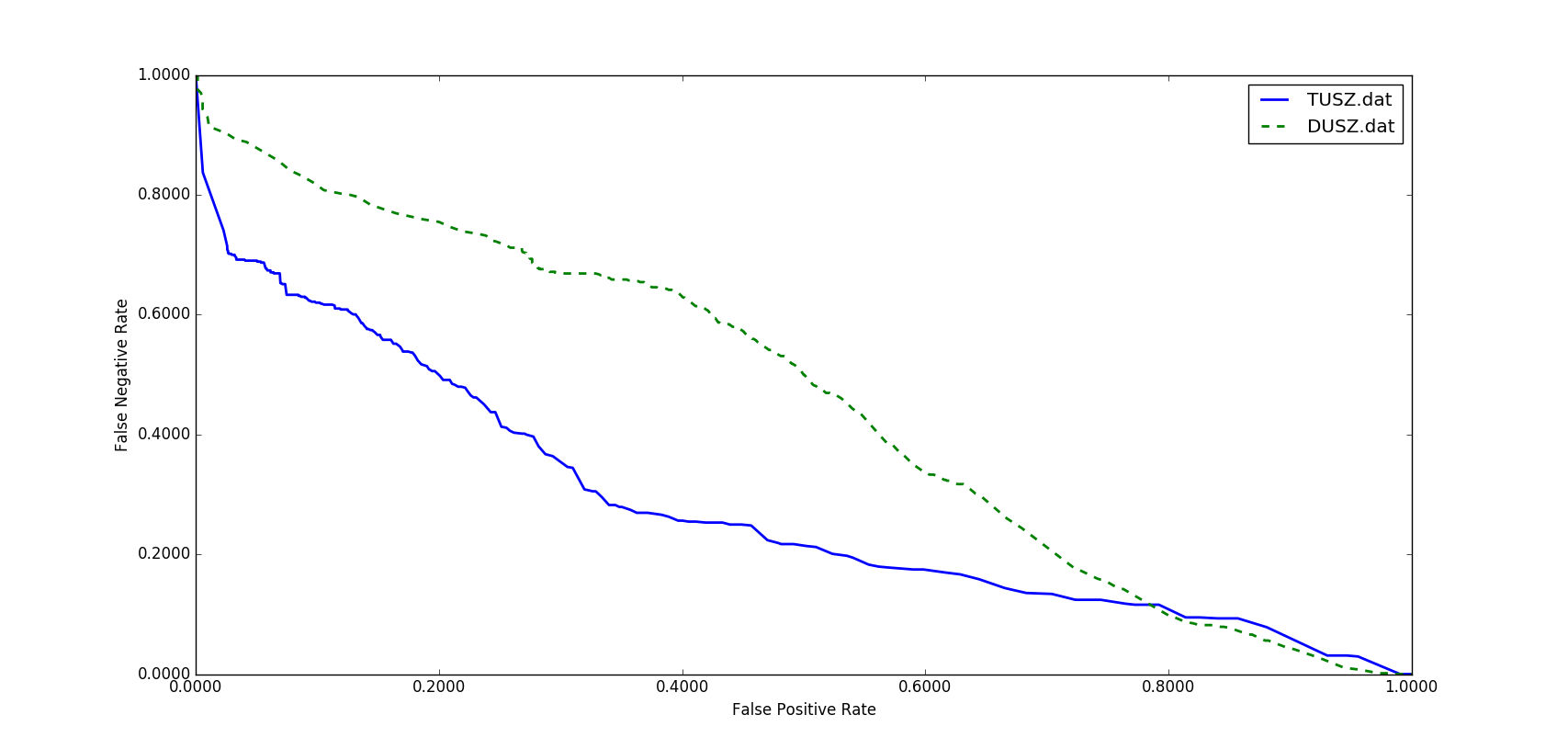


Figure 5: A performance comparison of TUSZ and DUSZ

##### Acknowledgments

Research reported in this publication was most recently supported by the National Human Genome Research Institute of the National Institutes of Health under award number U01HG008468. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This material is also based in part upon work supported by the National Science Foundation under Grant No. IIP-1622765. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The TUH EEG Corpus work was funded by (1) the Defense Advanced Research Projects Agency (DARPA) MTO under the auspices of Dr. Doug Weber through the Contract No. D13AP00065, (2) Temple University’s College of Engineering and (3) Temple University’s Office of the Senior Vice-Provost for Research.

##### References

1. I. Obeid and J. Picone, “Machine Learning Approaches to Automatic Interpretation of EEGs,” in *Biomedical Signal Processing in Big Data*, 1st ed., E. Sejdik and T. Falk, Eds. Boca Raton, Florida, USA: CRC Press, 2017, p. 70.
2. C. B. Swisher, C. R. White, B. E. Mace, and K. E. Dombrowski, “Diagnostic Accuracy of Electrographic Seizure Detection by Neurophysiologists and Non-Neurophysiologists in the Adult ICU Using a Panel of Quantitative EEG Trends,” *J. Clin. Neurophysiol.*, vol. 32, no. 4, pp. 324–330, Jan. 2015.
3. A. Harati, M. Golmohammadi, S. Lopez, I. Obeid, and J. Picone, “Improved EEG Event Classification Using Differential Energy,” in *Proceedings of the IEEE Signal Processing in Medicine and Biology Symposium*, 2015, pp. 1–4.
4. S. Lopez, G. Suarez, D. Jungreis, I. Obeid, and J. Picone, “Automated Identification of Abnormal EEGs,” in *IEEE Signal Processing in Medicine and Biology Symposium*, 2015, pp. 1–4.
5. A. Schad et al., “Application of a multivariate seizure detection and prediction method to non-invasive and intracranial long-term EEG recordings,” *Clin. Neurophysiol.*, vol. 119, no. 1, pp. 197–211, 2008.
6. P. Thodoroff, J. Pineau, and A. Lim, “Learning Robust Features using Deep Learning for Automatic Seizure Detection,” *Proc. Mach. Learn. Res.*, vol. 56, pp. 178–190, 2016.
7. L. Vidyaratne, A. Glandon, M. Alam, and K. M. Iftekharuddin, “Deep recurrent neural network for seizure detection,” in *Proceedings of the International Joint Conference on Neural Networks* (IJCNN), 2016, pp. 1202–1207.
8. M. Golmohammadi et al., “The TUH EEG Seizure Corpus,” presented at the American Clinical Neurophysiology Society (ACNS) Annual Meeting, 2017, p. 1.
9. S. Ramgopal, “Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy,” *Epilepsy Behav.*, vol. 37, pp. 291–307, 2014.
10. M. Christensen, A. Dodds, J. Sauer, and N. Watts, “Alarm setting for the critically ill patient: a descriptive pilot survey of nurses’ perceptions of current practice in an Australian Regional Critical Care Unit,” *Intensive and Critical Care Nursing*, vol. 30, no. 4. pp. 204–210, 2014.
11. J. Picone, “Signal modeling techniques in speech recognition,” *Proc. IEEE*, vol. 81, no. 9, pp. 1215–1247, 1993.
12. X. Zhang, Y. Liang, Y. Zheng, J. An, and L. C. Jiao, “Hierarchical Discriminative Feature Learning for Hyperspectral Image Classification,” *IEEE Geosci. Remote Sens. Lett.*, 2016.
13. S. Lopez, M. Golmohammadi, I. Obeid, and J. Picone, “An Analysis of Two Common Reference Points for EEGs,” in *Proceedings of the IEEE Signal Processing in Medicine and Biology Symposium*, 2016, pp. 1–4.
14. J. Picone, “Continuous Speech Recognition Using Hidden Markov Models,” *IEEE ASSP Mag.*, vol. 7, no. 3, pp. 26–41, Jul. 1990.
15. P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, “Extracting and composing robust features with denoising autoencoders,” in *Proceedings of the International Conference on Machine Learning* (ICMLA), 2008, pp. 1096–1103.
16. G. E. Hinton, S. Osindero, and Y.-W. Teh, “A Fast Learning Algorithm for Deep Belief Nets,” *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, May 2006.
17. I. Obeid and J. Picone, “The Temple University Hospital EEG Data Corpus,” *Front. Neurosci. Sect. Neural Technol.*, vol. 10, p. 00196, 2016.
18. D. A. Ross, J. Lim, R. S. Lin, and M. H. Yang, “Incremental learning for robust visual tracking,” *Int. J. Comput. Vis.*, vol. 77, no. 1–3, pp. 125–141, 2008.
19. S. Hochreiter, S. Hochreiter, J. Schmidhuber, and J. Schmidhuber, “Long short-term memory.,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–80, 1997.
20. D. P. Kingma and J. L. Ba, “Adam: a Method for Stochastic Optimization,” *Int. Conf. Learn. Represent.*, pp. 1–15, 2015.
21. K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *CoRR*, 2014.
22. V. Nair and G. E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines,” *Proc. Int. Conf. Mach. Learn.*, no. 3, pp. 807–814, 2010.
23. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.
24. D. Clevert, T. Unterthiner, and S. Hochreiter, “Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs),” arXiv Prepr. arXiv1511.07289, pp. 1–14, 2016.
25. B. Wilson, M. L. Scheuer, C. Plummer, B. Young, and S. Pacia, “Seizure detection: Correlation of human experts,” *Clin. Neurophysiol.*, vol. 114(11), pp. 2156–2164, 2003.
26. H. A. Haider et al., “Sensitivity of quantitative EEG for seizure identification in the intensive care unit,” *Neurology*, vol. 87, no. 9, pp. 935–944,2016.