Epileptic Seizure Detection Framework using EEG Signals

Fei Guo

Abstract—Electroencephalography (EEG) is the recording the electrical potential on the surface of the scalp which is widely applied to diagnose abnormal brain diseases, especially epileptic seizure. Methods that combined of conventional signal analysis techniques with supervised machine learning algorithms are well developed during past few decades. However researchers realizes that, the biggest limitation of supervised machine learning need to be solved, is tremendous annotation work by well-trained expertise. Moreover, the signal analysis techniques, which is essential component of conventional machine learning methods, will generate feature vectors used to classify the dataset. In this paper, a novel semi-supervised machine learning based framework is proposed to perform automatic epileptic seizure detection. The deep learning LSTM network will generate uncertainty scores for the instance selection scheme and the experiments evaluation shows that the uncertainly sampling can achieve a promising result compared to random sampling strategy.

Index Terms—EEG, Active Leanring, Semi-supervised, Seizure Detection, Epilepsy

1 INTRODUCTION

The epilepsy is one of the most common neurological disorder caused by abnormal synchronization of brain neurons. Current diagnosis and detection of epilepsy largely depends on the electroencephalograph (EEG) signals. The long-term EEG recordings of an epileptic patient obtained from the ambulatory recording systems contain a large volume of EEG data. Detection of the epileptic activity requires a time consuming analysis of the entire length of the EEG data by an expert. Nowadays, many signal analyzing and processing techniques such as machine learning algorithmhave been implemented for automatically epilepsy detection and prediction., such as SVM [1], ANN [13], decision tree [14]. Whats more, recently more deep active learning methods, such as recurrent neural network (RNN) [16] have been designed and implemented to perform seizure detection due to its unbeatable classification performance. Although overall average detection accuracy rates of current classification methods are relative high and satisfying, the demands of all data to be correct labeled is one of the significant shortcoming and limitation. Labeling massive data is time consuming and labor intensive process, which should not be overlooked anymore. In many practical machine learning domains, unlabeled data is available and abundant, but labeled data is relative rare due to the label effort is expensive and time consuming. To overcome this shortcoming and reduce the demand for label data, the active learning comes into play. Contrast with supervised learning, active learning is capable to actively request small sized unlabeled instances to be labeled by an experienced expertise, and thus, reducing the labeling cost significantly without loss of classification accuracy [15]. Another essential component of supervised machine learning algorithms is feature extraction, which the classification accuracy relay on. In particular, the selected features will affect the classification accuracy on a large extent, because it requires cross-disciplinary knowledge to find a well fit features.

In order to overcome the previous mentioned shortcomings of existing algorithm, this paper puts forward an active learning based seizure detection framework using EEG signals, and combined with LSTM network to select unlabel samples with uncertainty sampling scheme.

The reamining part of this paper is organized as follows. We first give a brief introduction to the related works on active learning and deep learning in Section 2, Then the system design scheme is present step by step in Section 3. Section 4 is the evaluation of the performance of our proposed system. The last but not the least, conclusion and brief future prospective is discussed in Section 5.

2 RELATED WORK

The epilepsy is one of the most common neurological disorder caused by abnormal synchronization of brain neurons. Current diagnosis and detection of epilepsy largely depends on the electroencephalograph (EEG) signals. As technology advances, especially the neural network and machine learning technology, computerized algorithms in epilepsy detection are on the increase. Nowadays, many machine learning algorithms have been well implemented for epilepsy detection, such as SVM, ANN, decision tree [4] [9] [1]. Based on the property of EEG signal of epilepsy, Vairavan Srinivasan et al. [17] adopted the ApEn features to measures the predictability of EEG signals between current amplitude and the previous amplitude values. ApEn is a statistical parameter that calculate the regularity of signal in time domain. V. Srinivasan at al also proposed a similar algorithm on the basis of previous work, but concentrated on the combination of both time domain and frequency domain features. The network used in both works is recurrent Elman neural networks. U. Rajendra Acharyaa focuses on more entropy features, such as Approximate Entropy (ApEn), Sample Entropy (SampEn), and two Phase Entropy, and also explored more comprehensive and deeper application by feeding

- *Fei Guo is with the Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH 44106.*
E-mail: fxg104@case.edu

the entropy features to 7 different classifier: Fuzzy Sugeno Classifier, Support Vector Machine, K-Nearest Neighbour, Probabilistic Neural Network, Decision Tree, Gaussian Mixture Model, and Naive Bayes Classifier [1] and overall classifiers performance were evaluated in the works. Although proper entropy features can generate great classification results based on the papers reviewed above, but in the real practice, the less signal process, the more information will be left in the signals and more time will be saved. Hence researchers based on the EEGs property, decompose the EEG signal into five sub bands according to the frequency band, among all the decomposition algorithm, discrete wavelet transform(DWT) is one of the most efficient and wide applied theory. Orhan et al. adopted DWT to capture features from the EEG signals and then combined with ANN to formulate the entire classification system. H. Ocak [12], L. Guo et al. [6], H. Adeli et al. [2] and Kumar et al. [7] detected epilepsy with the use of discrete wavelet transform (DWT) along with non-linear features. Active learning has been widely applied in different areas and demonstrated to be one of the most cost efficient methods, which can request and label most informative queries and thus reduce the labeling cost significantly. Keze Wan et al. Peng Liu et al Yarin Gal et al [18] [10] [19] [5] applied active learning in image recognition and classification. As shown in their works, using active learning, promising performance can be achieved using less training data. More than in the image recognition area, active learning can also expertise in learning any other media (e.g., images audios and videos). Other applications of active learning that being investigated currently are function optimization [11] and experimental design optimization [4].

3 System Framework

In this section, the system design of Deep Active Learning for Epileptic Seizure Detection framework is discussed. It consists of six parts: data pre-processing, design and implementation of LSTM network and visualization UI design. The system flowchart is shown in Figure 1. In this paper, the goal is to detect and classify any given unknown EEG signal into Normal or Epilipsy. First of all, EEG dataset will be processed through series of data processing techniques, including De-nosing, data segmentation and normalization. Then these processed EEG data are inputted into LSTM framework to train the network. Next, the trained network will be tested with entire 5 data subsets. In the end, the trained model is exported to android studio project, and an andriod application is developed, which can detect any given 2 second raw EEG data.

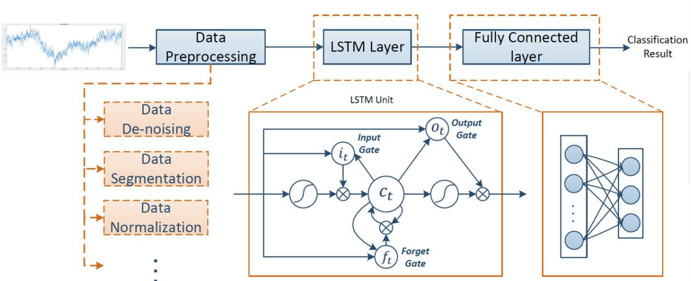

Fig. 1. system overview of LSTM network based seizure detection framework

3.1 Data Pre-processing

EEG data tested in this paper is provided by Department of Epileptology, University of Bonn [3], which is open to download for public. The entire dataset is consist of 5 subsets of data, named from Set A to E correspondingly. Each subset contains 100 single channel EEG segments records with 4097 samples in each segments, the time duration of each segment is each of 23.6 seconds length. All EEG signals were recorded with the same 128- channel amplifier system, using an average common reference, with 12 A/D conversion bit rate of 12, sampling rate of 173.61 Hz. The whole dataset can be divided into 2 category according to the testers. Dataset A and B were recorded from 5 healthy volunteers with international 10-20 electrodes placement scheme on the volunteers scalp, while we choose Dataset C, D and E are selected as interested data, which were recorded from 5 patients diagnosed with epilepsy with implanted electrodes inside the patients cranial. Specifically, dataset C and D were recorded from 5 patients seizure free interval, while C is from intracranial electrodes implanted within epileptic zone and D is from non-epileptic zone, and together C and D are labeled as Normal. Dataset E also obtained by intracranial electrodes and only contains seizure activities and labeled as Epilepy. First, each EEG records will be processed by 5 levels discrete wavelet with Daubechies in order 4 (db4) which has been proved to be proper for this specific EEG dataset [9]. Figure 2 illustrates the structure of fifth level wavelet decomposition. The Table 1 shows the actural frequency range of each subbands. EEG signal processed by the

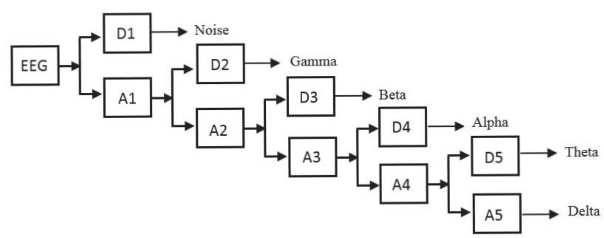

Fig. 2. Structure of the fifth level wavelet decomposition. According to frequency property of EEG, the subband wave Delta, Theta, Alpha, Beta and Gemma can be seperated accordingly

DWT will be de-noised by a series combination of high pass and low pass filters as shown in 2. After discarding first 1.6 seconds data, rest 22 seconds data will be segmented evenly into 11 sub signals each of 2 seconds length. Total 3300 segments with either labeled Normal or Epilepy will be normalized and in the next step be fed as input of the LTSM classifier.

3.2 Deep Active Learning LSTM network Design

Recurrent neural networks are widely applied in deep active learning based EEG seizure detection methods with promising outcomes. Hence, based on the properties of EEG, in this paper, LSTM, which is a special kind of RNN, is applied to classify EEG signals to normal and epileptic seizure. In general, LSTM are designed to avoid the long-term dependency problems, and the general mathematics formulations of LSTM are defined with

Decomposition Levels	Sub-bands	Frequency Range(Hz)
Gamma	D2	24.5-40
Beta	D3	12.5-27.5
Alpha	D4	5.5-14.5
Theta	D5	2.5-6.2
Delta	A5	0-3.5

TABLE 1
Frequency sub-bands of EEG signal using 5 level decomposition.

The office definition of subband of Delta, Theta, Alpha, Beta and Gemma should be 0.1-3Hz, 4-7Hz, 8-12.5Hz, 12.5-30Hz and 32-100Hz. The actural DWT result in this experiment is slightly different from the offical definition, which is caused by the DWT property. DWT method decomposes the frequency strict by half while frequency distribution of EEG signal is continously distributed.

following equations:

$$\mathbf{i}_t = \sigma\left(\mathbf{W}_{ix}\mathbf{x}_t + \mathbf{W}_{ih}\mathbf{h}_{t-1} + \mathbf{W}_{ic}\mathbf{c}_{t-1} + \mathbf{b}_i\right) \tag{1}$$
$$\mathbf{f}_t = \sigma\left(\mathbf{W}_{fx}\mathbf{x}_t + \mathbf{W}_{fh}\mathbf{h}_{t-1} + \mathbf{W}_{fc}\mathbf{c}_{t-1} + \mathbf{b}_f\right) \tag{2}$$
$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \phi\left(\mathbf{W}_{cx}\mathbf{x}_t + \mathbf{W}_{ch}\mathbf{h}_{t-1} + \mathbf{b}_c\right) \tag{3}$$
$$\mathbf{o}_t = \sigma\left(\mathbf{W}_{ox}\mathbf{x}_t + \mathbf{W}_{oh}\mathbf{h}_{t-1} + \mathbf{W}_{oc}\mathbf{c}_t + \mathbf{b}_o\right) \tag{4}$$
$$\mathbf{h}_t = \mathbf{o}_t \odot \phi\left(\mathbf{c}_t\right) \tag{5}$$

In the formulations, t is time step number, and $\mathbf{x}_t$, $\mathbf{i}_t$, $\mathbf{f}_t$, $\mathbf{o}_t$, $\mathbf{c}_t$ are inputs, input gate, a forget gate, an output gate and a memory cell respectively. All these gates and memory cell variables are defined in the dimension of $\mathbb{R}^d$ as the hidden vector $\mathbf{h}_t$, while the input variables $\mathbf{x}_t$ is in dimension $\mathbb{R}^e$. $\mathbf{W}_{ix}$, $\mathbf{W}_{ih}$, $\mathbf{W}_{ic}$, $\mathbf{W}_{fx}$, $\mathbf{W}_{fh}$,$\mathbf{W}_{fc}$, $\mathbf{W}_{ox}$, $\mathbf{W}_{oh}$,$\mathbf{W}_{oc} \in \mathbb{R}^{2d}$ are weighted matrices, and $\mathbf{b}_i, \mathbf{b}_f, \mathbf{b}_c, \mathbf{b}_o \in \mathbb{R}^d$ are biases of LSTM, which are obtained from training process. σ is the sigmoid function and sign $\odot$ denotes element-wise multiplication. The current hidden state $\mathbf{h}_t$ has the following relationship with previous hidden state $\mathbf{h}_{t-1}$.

$$\mathbf{h}_t = LSTM(\mathbf{h}_{t-1}, \mathbf{x}_t, \theta). \tag{6}$$

EEG segment sample vector denotes as $x = \{x_1, x_2, ..., x_T\}$, later input vector x will be input the LSTM network and generate the output vector $\mathbf{h}_t$, and straight after feeding into a fully connected layer, a global probability distribution matrix $\hat{y}$ will be generated with equation 6

$$\hat{y} = softmax\left(\mathbf{W}\mathbf{h}_T + b\right) \tag{7}$$

The loss function formulation is shown in the equation 8 as follows:

$$L\left(\hat{y}, y\right) = -\sum_{i=1}^{N}\sum_{j=1}^{C} y_i^j \log\left(\hat{y}_i^j\right) \tag{8}$$

N is the EEG sample dimension while C is the variety of classes. With the basic formulation of each LSTM layer, the entire training frame will be performed as shown in Algorithm 1.

With the basic mathmatics setup, the proposed active learning architecture is exhibited in Figure 3

3.3 Sampling Scheme

Since active learning requires only a small mount of data to be annoted, the effective selection of sample will become crucial adjective to the performance. In this framework, the uncertainty sampling scheme is implemented [8]. To be more specific, if

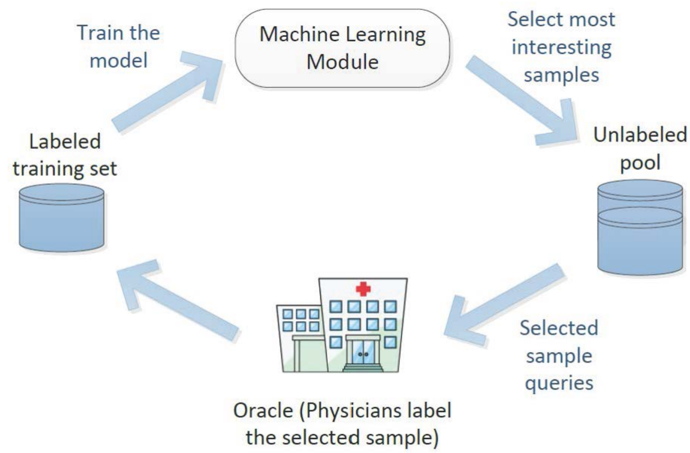

Fig. 3. The architecture of active learning framework

the classification score is high, which means there is less uncertainty in labeling this sample, on the contrary, if the classfication confidence score is relatively low, which means there is high possibility that the classfication result could be wrong, therefore this sample will be quaried to be labeled by expertises. In the LSTM framework, the unility function is defined as D^U to represent the possibility of a sample selected to be queried. The uncertainty sampling scheme guarantees the most uncertain unlabeled sample will always be selected and labeled correctly, which is a well-fit scheme for active learning. In the algorithm presented previous statement, the uncertainty equation respected to classification threshold θ is expressed in Equation 9.

$$U_M\left(x_i\right) = -u_1(x_i) = -\frac{|\hat{y}(x_i) - \theta|}{\theta} \tag{9}$$

After all the mathmatics formulations, primary steps of the LSTM algorithm are shown in Algorithm 1.

Algorithm 1 Active learning training process

Given:
D^L: Set of labeled examples
D^U: Set of unlabeled data pool
B: number of examples to be selected in each iteration
U_M: Utility function
Algorithm:
Initialization: Initialize the network. Loop step 1-7 until stopping criterion is met:
1. Train the primary LSTM model M on dataset D^L
2. For all $x_i \in D^U$: $u_{x_i} \leftarrow U_M\left(x_i\right)$
3. Select B examples $x_i \in D^U$ with highest utility u_{x_i}.
4. Retrieve the selected samples with their labels (x_i, y_i)
5. Move the selected samples from D^U to D^L
Return: Trained LSTM model M.

3.4 Android Application Design

With the support of TensorFlow Mobile or Lite, machine learning model can be implemented on a smartphone and run those algorithms to determine new outcomes on the go. The android application UI includes three components: GraphView to display the test EEG signal, Button functions to perform ”Clear”, ”Detect”, ”Load Signals” ”Next” functions and the TextView to display the classification results, as shown in Figure 4. Designed application could display the given test EEG signal and detect the probability of seizure with high accuracy and promptness.

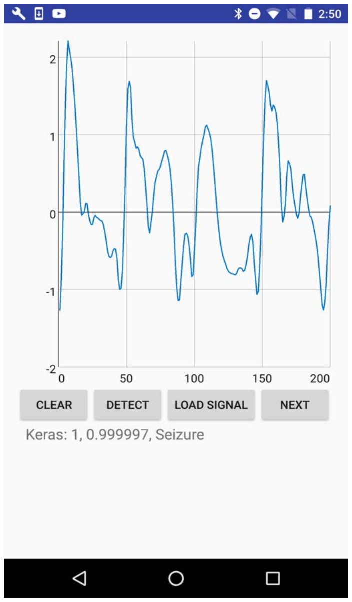

Fig. 4. Android Application User interface Design

4 EXPERIMENT

The detailed evaluation of the framework presented above is presented in this section. The start point of the entire framework is random selecting of sample. For the proposed active learning training process. The LSTM network consists of a layer with 256 LSTM nodes and one fully connected layer as shown in the Figure 1. The loss function is binary cross-entropy. Batch size is set to 20, the epoch is set to 10, while the threshold for classification $\theta = 0.5$.

4.1 Uncertain Sampling VS Random Sampling

According to Figure 5, the classification results of both uncertain sampling and random sampling are increaseing along with the increase of the query sample size. The uncertainty sampling can reach stable classification performance with shorter time and smaller query sample size. In particular, the uncertainty nearly stable with 500 query samples which is 1/5 of the query size, around 2400, needed for random sampling scheme. This result shows that the uncertainty sampling active learning algorithm can largely reduce the annotation labor even with much better performance.

4.2 Uncertain Sampling VS Matlab build-in Classifier

In addition horizontal comparison, the evaluation in vertical comparison with Matlab build-in Classifiers will also be demostrated in the following section. Among all the available Matlab classifiers, (KNN, Decision Tree, SVM...), the highest is Bagged Tree with accuracy at 95.9%, while the averaged accuracy of LSTM classifier is 96.3% with 5 cross-validation. Overall, LSTM outstands the Matlab classifiers with higher precision rate and lower recall rate.

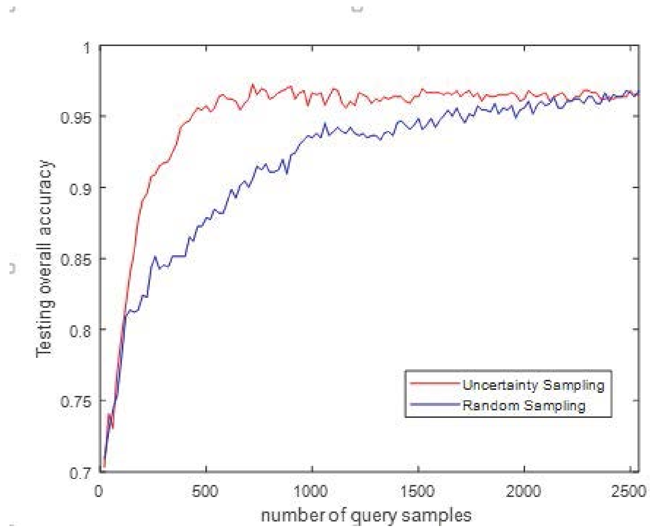

Fig. 5. The classification performance of LSTM seizure detection framework. x axis is the quary samples, y axis is the overall classification accuracy. Red line is ROC of Uncertainty Sampling, blue line is ROC of Random Sampling

TABLE 2
LSTM Precision = 97.75%, Recall = 91% Highest Accuracy = 95.9%; Bagged Trees Precision = 94.23%, Recall = 93.55%, Overall Accuracy = 96.3%

(a) Confusion Matrix of LSTM

		True Class	
		1	0
Predicted Class	1	1001	23
	0	99	2177

(b) Confusion Matrix of Bagged Tree in Matlab

		True Class	
		1	0
Predicted Class	1	1029	63
	0	71	2137

5 CONCLUSION AND FUTURE WORK

In this paper, a deep active learning LSTM based seizure detection framework using EEG signals is presented. The Long Short Term Memory based network with active learning can perform epileptic seizure detection effectively and accurately. In addition, the experiment results on EEG database validate the efficiency and accuracy of the entire system.

Future work is to reduce the annotation labor entirely with unsupervised machine learning algorithm with no trade of the accuracy and efficiency. To further look at the real pratical medical environment, the setting are much more complicated, massive valuable data are generated without enough correlation analysis. In the near future, the framework could be more smart in the sense of detecting multi-type of bio-signals, such as ECG, EKG, etc.

REFERENCES

[1] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K.-H. Ng, and J. S. Suri. Automated diagnosis of epileptic eeg using entropies. *Biomedical Signal Processing and Control*, 7(4):401–408, 2012.

[2] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr. A wavelet-chaos methodology for analysis of eegs and eeg subbands to detect seizure and epilepsy. *IEEE Transactions on Biomedical Engineering*, 54(2):205–211, 2007.

[3] 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*, 64(6):061907, 2001.

[4] A. C. Atkinson and R. Bailey. One hundred years of the design of experiments on and off the pages of biometrika. *Biometrika*, 88(1):53–97, 2001.

[5] Y. Gal, R. Islam, and Z. Ghahramani. Deep bayesian active learning with image data. *arXiv preprint arXiv:1703.02910*, 2017.

[6] L. Guo, D. Rivero, and A. Pazos. Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. *Journal of neuroscience methods*, 193(1):156–163, 2010.

[7] Y. Kumar, M. Dewal, and R. Anand. Epileptic seizure detection using dwt based fuzzy approximate entropy and support vector machine. *Neurocomputing*, 133:271–279, 2014.

[8] D. D. Lewis and J. Catlett. Heterogeneous uncertainty sampling for supervised learning. In *Proceedings of the eleventh international conference on machine learning*, pages 148–156, 1994.

[9] M. Li, W. Chen, and T. Zhang. Classification of epilepsy eeg signals using dwt-based envelope analysis and neural network ensemble. *Biomedical Signal Processing and Control*, 31:357–365, 2017.

[10] P. Liu, H. Zhang, and K. B. Eom. Active deep learning for classification of hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(2):712–724, 2017.

[11] A. W. Moore, J. G. Schneider, J. A. Boyan, and M. S. Lee. Q2: Memory-based active learning for optimizing noisy continuous functions. In *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, volume 4, pages 4095–4102. IEEE, 2000.

[12] H. Ocak. Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy. *Expert Systems with Applications*, 36(2):2027–2036, 2009.

[13] U. Orhan, M. Hekim, and M. Ozer. Eeg signals classification using the k-means clustering and a multilayer perceptron neural network model. *Expert Systems with Applications*, 38(10):13475–13481, 2011.

[14] K. Polat and S. Güneş. Classification of epileptiform eeg using a hybrid system based on decision tree classifier and fast fourier transform. *Applied Mathematics and Computation*, 187(2):1017–1026, 2007.

[15] B. Settles. Active learning literature survey. *University of Wisconsin, Madison*, 52(55-66):11, 2010.

[16] V. Srinivasan, C. Eswaran, and Sriraam. Artificial neural network based epileptic detection using time-domain and frequency-domain features. *Journal of Medical Systems*, 29(6):647–660, 2005.

[17] V. Srinivasan, C. Eswaran, and N. Sriraam. Approximate entropy-based epileptic eeg detection using artificial neural networks. *IEEE Transactions on information Technology in Biomedicine*, 11(3):288–295, 2007.

[18] K. Wang, D. Zhang, Y. Li, R. Zhang, and L. Lin. Cost-effective active learning for deep image classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016.

[19] L. Yang, Y. Zhang, J. Chen, S. Zhang, and D. Z. Chen. Suggestive annotation: A deep active learning framework for biomedical image segmentation. *arXiv preprint arXiv:1706.04737*, 2017.