Reduction of Massive EEG Datasets for Epilepsy Analysis using Artificial Neural Networks

Howard J. Carey, III, Kasun Amarasinghe, Milos Manic
Department of Computer Science
Virginia Commonwealth University
Richmond, VA USA
careyhj@vcu.edu, amarasinghek@vcu.edu , misko@ieee.org

***Abstract*— Epileptic seizure source identification involves neurologists combing through a substantial amount of data manually, which sometimes takes weeks per patient. This paper presents a methodology for minimizing the amount of data a neurologist has to analyze to identify the seizure focus. The method keeps the neurologist as the final decision maker and aids in the decision making process. It has to be noted that the primary focus of the work was not improving the accuracy of interictal spike detection but reduction of the volume of data. The presented methodology is based on Artificial Neural Networks (ANN) and is implemented on EEG data collected on 5 patients using a dense array EEG reader. As a baseline, a simple template matching was implemented on the same dataset. Experimental results showed that the ANN based methodology was able to reduce the dataset by 98%, a significant improvement on the template matching method.**

Keywords—EEG, Interictal Spike Detection, Epilepsy, Neural Network

I. INTRODUCTION

Epilepsy is a neurological disorder that is caused by an abnormal firing of a cluster of neurons in the brain [17, 30]. According to the World Health Organization (WHO), it affects approximately 50 million people worldwide across all age groups [32]. People who have epilepsy experience debilitating seizures that come without warning, interrupting their daily life and potentially endangering themselves. Epilepsy is treatable through various forms of medication, but has no known cure [32].

Doctors use electroencephalography (EEG) to analyze the electromagnetic radiation given off by the brain, commonly called brainwaves. An EEG reader consists of electrodes that are placed in strategic locations on a patient's scalp. Each electrode, gives out a voltage over time readout (see Fig. 1).

The EEG readings can be used to identify abnormal brainwave patterns, such as an epileptic seizure or interictal spike. An ictal event is another name for a seizure event, and an interictal spike is a spike that occurs between ictal events (between seizures). In EEG data, an epileptic seizure is identified by a period of very high amplitude, short duration pulses. An interictal spike, is a high amplitude pulse that occurs sporadically. Interictal spikes are not seizures themselves, but are generated by the same group of neurons that cause the seizures [6, 29]. Therefore, if the source of the interictal spike can be identified, the patient's seizure source (known as the seizure focus) can be found as well.

To identify these spikes, a neurologist must manually analyze the EEG output across multiple channels and be able to discern a spike from the noisy background brainwave signals. Such noise can include eye blinks, facial muscle movement, errant EM waves from electronic devices, etc. The neurologist must identify enough spikes to be confident in where the seizure focus is located. This is a time-consuming process, taking up to 10 days of analysis per patient.

This paper presents a method to minimize the amount of data a neurologist has to analyze to identify the seizure focus. An artificial neural network was used to analyze the EEG data from 5 patients and identify interictal spike patterns to filter out those regions which do not contain interictal spikes. This implementation was designed to keep the neurologist in the loop as the final arbiter of whether a given signal is an interictal spike. The primary focus of this method was not accuracy of interictal spike detection, but reduction of data that a neurologist has to manually analyze.

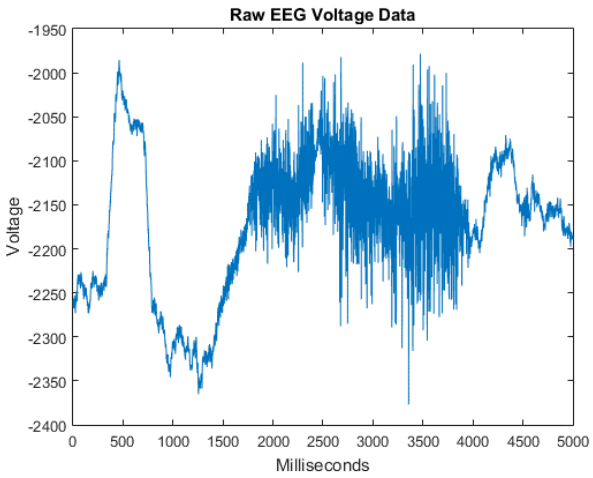


Fig. 1: Raw unfiltered EEG data

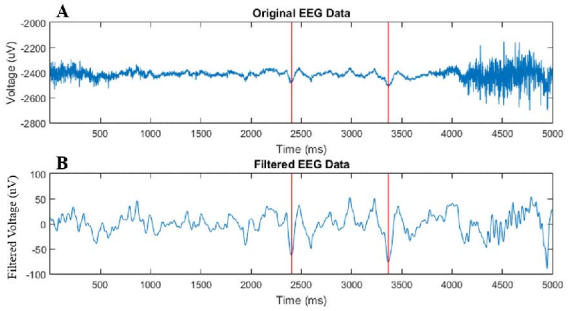


Figure 2: Original EEG data plotted with EEG data after being run through the bandpass filter.

This paper is organized as follows: Section II provides background on EEG and interictal spikes. Section III outlines the data preprocessing steps. Section IV details the methods used to analyze the effectiveness of the presented method. Section V analyzes the results of these methods and Section VI finishes with some closing remarks and discussion about future directions of this research.

II. EEG and Interictal Spikes

This section introduces Electroencephalography and interictal spikes.

Every living person's brain generates pulses of electricity, creating an electromagnetic field around the brain. EEGs are devices capable of measuring this field by way of numerous sensors placed around the brain. Different devices can have differing numbers of electrodes, with more electrodes leading to higher resolution data being collected.

Each electrode of the EEG network measures the voltage potential across the patient's brain. Different thought patterns, muscle movement, and even emotions can be detected using EEG [19]. The electrodes are also very sensitive to various sources of noise, such as muscle movement around the face, or external sources of electromagnetic radiation.

Neurologists use the EEG to identify where in the brain a patient's seizures are coming from. Epileptic seizures themselves are characterized by excessive, synchronous abnormal firing of neurons in the brain [22]. The EEG signal during an actual seizure makes it difficult to determine the origin of these signals due to the strong, chaotic signals. One pattern that arises during monitoring of patients in between seizures is the interictal spike.

The interictal spike tends to occur in the region of the brain where the patient's seizures emanate from [24], [25]. The interictal spike is characterized by a sharp change in amplitude of the voltage in a relatively short amount of time (typically 60-100ms, about 100ms for the patients in this study). This pattern is similar between patients, though not exactly the same. The interictal spike allows neurologists to identify where a patient's seizures are coming from and prepare the patient for surgery.

The current detection method involves a doctor manually analyzing hours of EEG data, looking for the interictal spike by eye. This is an exhaustive, error-prone process, taking up to ten days per patient.

III. Preprocessing of EEG Data

This section discusses the preprocessing techniques used in the paper.

A. Raw EEG Data

This experiment examined patients from the Virginia Commonwealth University Health's (VCU Health) Department of Neurology. The EGI Dense Array EEG was used to examine five separate patients, with EEG readings between 5-24 hours. The Dense Array has a sampling rate of 1000Hz, generating millions of samples per electrode. Due to memory and processing constraints, only a small portion of the data from each patient was analyzed.

B. Signal Preprocessing

The information output from the EEG data is a voltage signal propagating over time. Figure 1 shows the raw voltage as a function of time from a single electrode of EEG data. The signal is visibly noisy, and any distinct patterns would be difficult for a human eye to discern. To reduce this noise, an infinite impulse response (IIR) Butterworth band pass filter [21] from 5-30 Hz is used. This creates a filter that only allows signals between 5 and 30 Hz to pass through, attenuating the high frequency noise and providing a de-trending of low

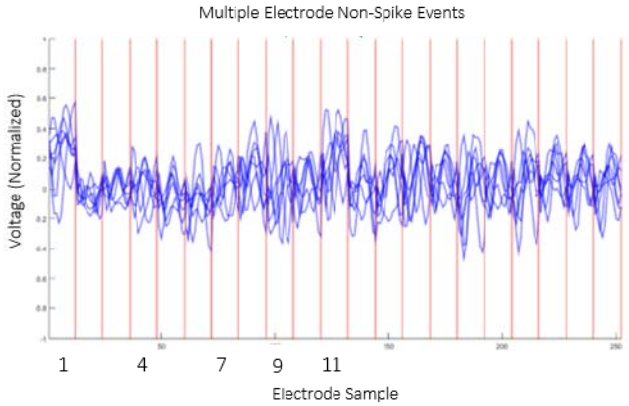


Figure 3: Multiple electrode input values showing 21 concatenated electrode windows from 7 random non-interictal spike data samples

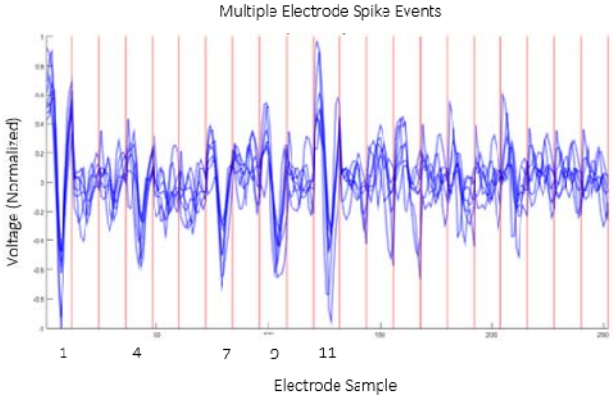


Figure 4: Multiple electrode input values showing 21 concatenated electrode windows of interictal spike data.

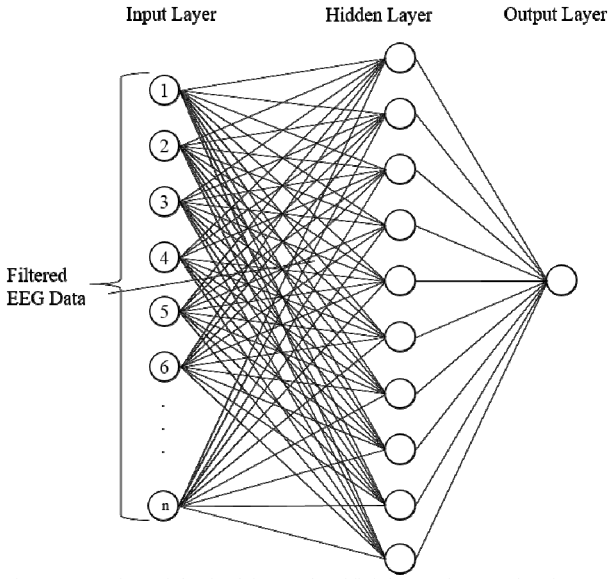


Figure 5: Design of the feed forward artificial neural network using error back-propagation.

frequency travel. Figure 2 shows two interictal spikes both before and after being run through this filtering process.

C. Data Extraction

The raw data format consists of 256 electrodes and 1,000 records per second for each electrode. To analyze the data for interictal spikes, a sliding window is used to analyze each potential window of data for the interictal spike pattern. The width of this window was chosen to be 120ms to capture the entire 100ms interictal spike, as well as information before and after the spike. Each window is extracted and considered a single data sample, allowing the windows that contain interictal spikes to be separated from those that do not.

IV. METHODS

A. Electrode Selection

The EGI Dense Array consists of 256 electrodes, each recording 1,000 samples per second. Due to memory constraints of the neural network that will be explained in a later section, not every electrode can be analyzed at once. To test the effectiveness of analyzing multiple electrodes, three different electrode sets were chosen with an increasing number of electrodes per set. With the window size of 120 samples ;per electrode, this would end up creating very large concatenated signals as input to the classification algorithms. To reduce the size of the input layers for each electrode test set, each electrode was downsampled to a tenth of the original number of values. More specifically, every 10^{th} data point from each electrode was kept, leaving 12 samples from each electrode instead of 120. The overall shape of the signal is kept, keeping much of the original information but reducing the over input size by an order of magnitude.

The Data Extraction method, detailed in Section IV C, discussed the sliding window technique used to analyze the data. For a single electrode, this would mean comparing a single window at a given time step to some separate signal for comparison. In the case of the multiple electrode analysis, each electrode was concatenated together to form a single, composite signal for the time step, as seen in Figures (3,4). Figure 3 shows a single time step across a number of electrodes, with the vertical red lines separating each electrode. The time step shown for this figure is not during an interictal spike, and there seems to be no discernable correlation of signals across the electrodes. However, in Figure 4 there is a similar spike signal seen on electrodes 1, 4, 7, 9, and 11. This correlation of spike signals between electrodes is the key feature of an interictal spike. Identifying signals such as those in Figure 4 from those in Figure 3 is the primary purpose of the two methods that will be explained in the next section.

B. Template Matching

Machine learning algorithms are computationally expensive, requiring training time, testing, and tuning of the hyper-parameters that are the design of each individual network. Signal analysis is not a new field, and machine learning algorithms are not the only method available to match signals to one another. A more straight-forward approach is to use simple template matching.

As a baseline for comparison, a spike template for each patient was constructed from the average of half of a patient's annotated spikes. This template was the width of the spike (120 samples). This average template was slid across the entire dataset for each patient, with the mean-squared-error (MSE) being recorded for each window. To determine whether a given signal was an interictal spike, a threshold was set so that any window with an MSE lower than this threshold would be classified as a spike. Five thresholds were tested, ranging from the standard deviation of the spikes used to create the average set, to its variance.

This threshold can be changed to best fit the patient's spikes, reducing the number of false positives and increasing the number of true positives, however choosing this threshold is a difficult task, and one that cannot be easily generalized across patients.

C. Artificial Neural Network

There is no simple, easy process to choose a good threshold for the template matching system described above. A learning algorithm, such as an artificial neural network (ANN), is capable of using training data to select a number of thresholds for each feature of the input space, and classify each sample accordingly.

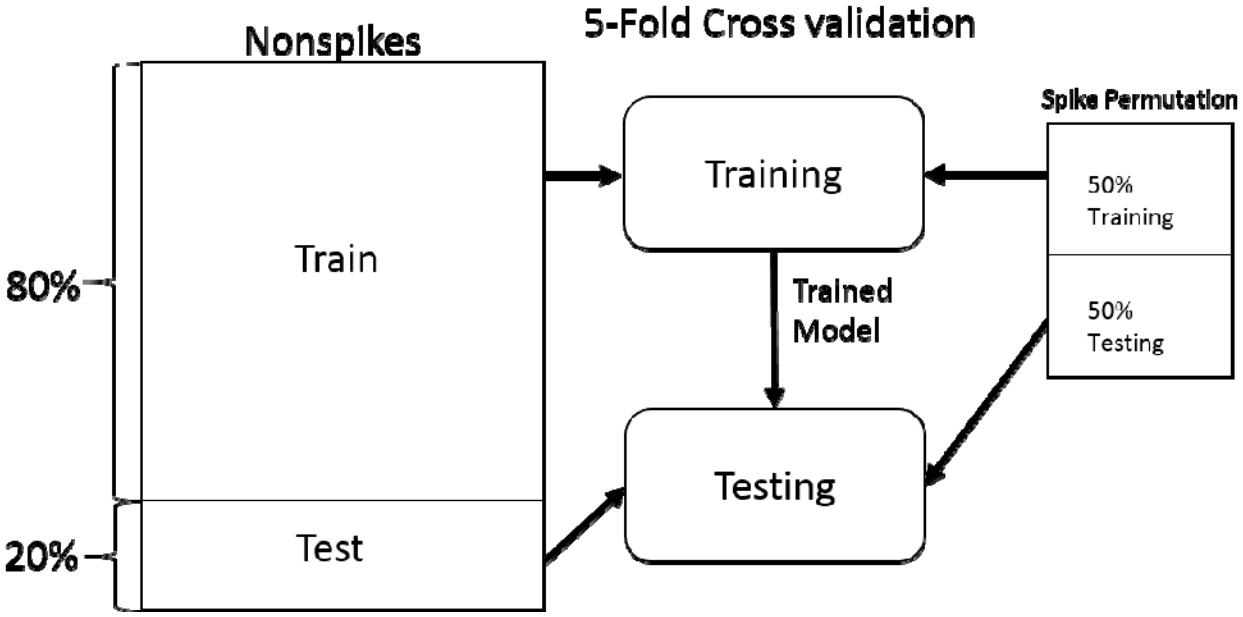


Figure 6: Diagram showing the 5-fold cross-validation technique used for training and testing.

Figure 6 shows the training process for the ANN used in this study. A 5-fold cross validation was used to train the network. The number of annotated spikes per patient is significantly smaller than the total number of samples to be analyzed. This imbalance in data requires a modification of the standard cross validation techniques. As mentioned in subsection A, the data was split into windows that contain spikes and windows that do not. To create the training and testing datasets, the nonspike data used normal 5-fold cross validation methods, taking 80% as training data and 20% as testing. The spike windows were split into 5 separate randomized permutations. For each fold of nonspike data, the corresponding spike permutation would be used, with the first half being used for training and second half for testing. This guaranteed that there was no contamination of the testing data during training, and that all nonspike and spike data was eventually trained and tested upon.

The training of the ANN used a batch size of 50 samples, with a 50% chance of each sample being a spike or a nonspike from the training sets. This equal weighting of spike and nonspike helps alleviate the problem of the massively imbalanced dataset. Since a given training size of spikes is anywhere between 3-7 spikes, each batch will contain multiples of the same spike, essentially weighting the spikes higher than a nonspike.

The ANN architecture, diagrammed in Figure 5, consisted of an input layer for each electrode test set, a single hidden layer with 100 neurons using the sigmoid activation function, and a single output neuron also using the sigmoid activation function. Error back-propagation method was used to train the ANN. Several ANN architectures were tested and the single hidden layer network with 100 neurons in the hidden layers produced the most accuracy with best generalization. The number of iterations used for training was 10,000, making a total of 500,000 total samples used for training with the batch size of 50.

V. Experimental Results

A. Evaluation Metrics

In this study, True Positives (TP) are spikes that are identified as spikes, True Negatives (TN) are non-spikes that are identified as non-spikes, False positives (FP) are non-spikes classified as spikes and False Negatives (FN) are spikes that are classified as non-spikes. The evaluation metrics used in this study are precision, recall, F1 score, and data reduction percentage. The precision is defined as:

$$precision = \frac{TP}{(TP+ FP)} \tag{1}$$

This determines how accurate the system is at filtering out actual spikes from falsely classified spikes.

Recall measures how accurate the system is at extracting spikes that are actually occurring in the data, and is defined as:

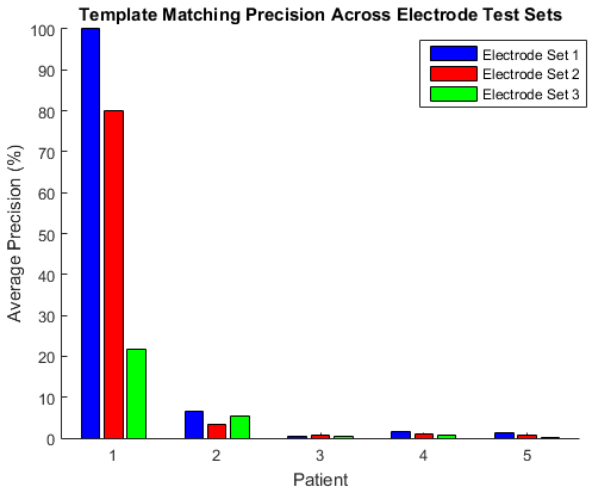
$$recall = \frac{TP}{(TP+FN)} \tag{2}$$

where TP is true positives and FN is false negatives.

The F1 score, defined as:

$$F1\_Score = 2\frac{precision * recall}{precision + recall}$$

The F1 score is a weighted average of the precision and recall, allowing a single value to represent the accuracy of the system.



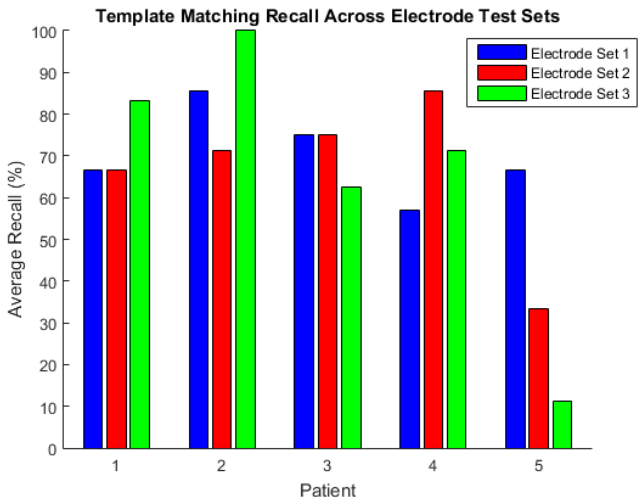
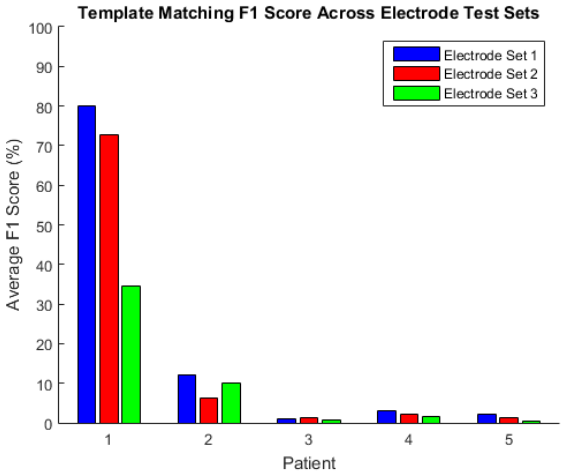


Figure 7: Precision and recall scores across all three electrode sets for the single patient analysis.



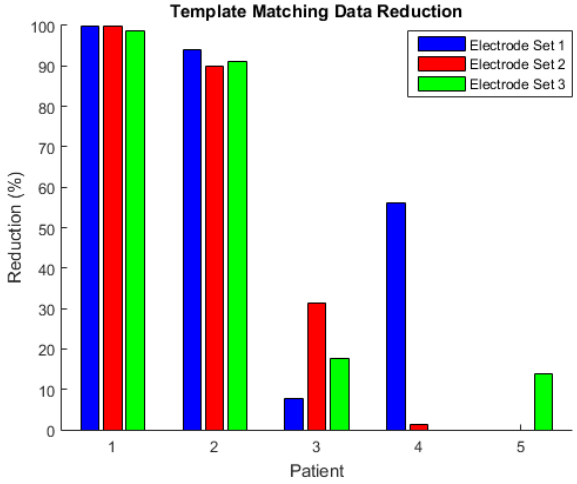


Figure 8: F1 score and data reduction percentage across all three electrode sets for single patient analysis.

The metric used to determine the data reduction percentage is defined as:

$$reduction\,\% = 100 - \left(100\,\frac{TP + FP}{\frac{Total\_EEG\_Data\_Size}{Spike\_Window\_Size}} \right)$$

with the Total_EEG_Data_Size being patient-specific data size, and the Spike_Window_Size equaling 120. This metric allows us to determine how much information remains that a neurologist would have to analyze.

B. Results

1) Template Matching Results

Figures (7,8) show the precision, recall, and F1 scores for the template matching analysis. Immediately obvious is the poor precision for all patients except Patient 1. Patient 1's performance is likely due to the quality of the annotated spikes. The more similar a set of annotated spikes are, the better a template matching system will work. This is because the standard deviation used to calculate the threshold value yield a more precise value, cutting out more false data. The poor precision means the template matching system does not discriminate between true positives and false positives very effectively. In an implemented system, this would yield a large dataset for the neurologist to analyze.

Adding more electrodes did not yield better results in most patients. A template matching system that relies on a threshold will likely perform worse when adding more electrodes unless the majority of the electrodes added contain spikes. The F1 score, Figure 8 shows the overall performance of the template matching system. Patient 1 performs well, while the remaining 4 do not perform nearly as well.

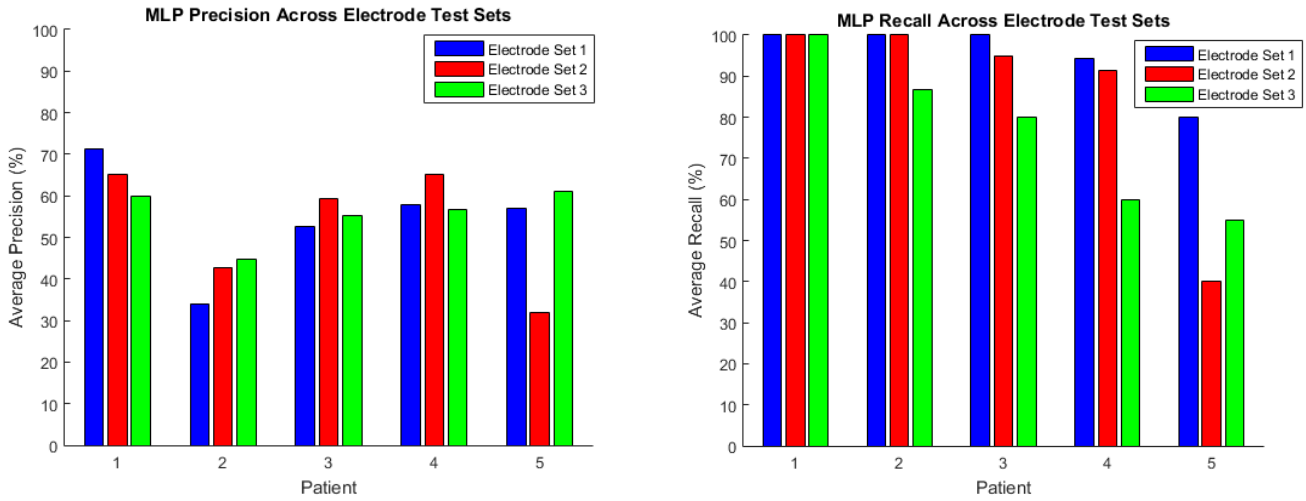


Figure 9: Precision and recall scores across all three electrode sets for the single patient analysis.

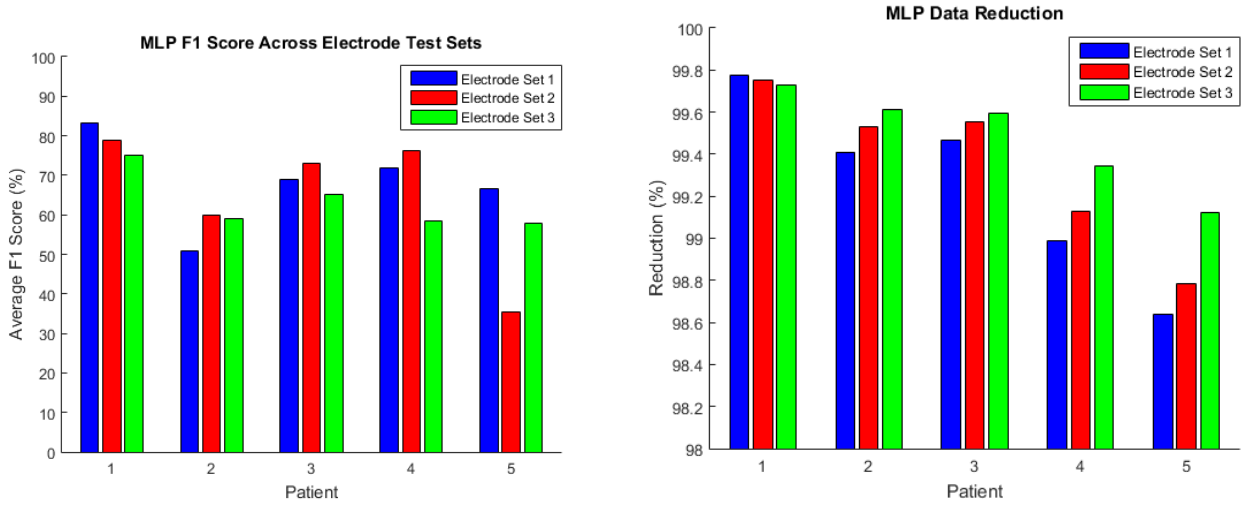


Figure 10: F1 score and data reduction percentage across all three electrode sets for single patient analysis.

Figure 8 also shows the data reduction % for the template matching. Patient 1 shows a 99% data reduction across all three electrode sets, with Patient 2 closely behind around 90%. Patient 2's case is interesting, while the precision for patient 2 is under 10% on all three electrode sets, the data reduction is still a significant amount. However, patients 3-5 show very inconsistent results in comparison to the first two patients, with patients 4 and 5 exhibiting no data reduction on some electrode sets.

2) Artificial Neural Network Results

Figure 9 shows the average precision across the 5 folds used for testing, giving a useful indication of how well the ANN can distinguish between true and false positives. Figure 9 also shows the average recall, giving a good indication of how well the ANN can identify the annotated spikes from the testing set. The F1 score averages across all 5 patients are consistently higher than the template matching, with the exception of Patient 1. Most importantly is the data reduction percentage shown in Figure 10 notice the y-axis beginning at 98%. While the F1 scores are not approaching 100% for any of the patients, the amount of data reduction is significant for all patients, which was the main goal for this project.

The analysis between the differing sizes of electrode sets allows us to gain useful insight into the effectiveness of the multiple electrode approach. By analyzing the precision, recall, and F1 scores, it is seen that increasing the number of electrodes, in general, decreased the performance of the ANN. This is likely due to the ANN not being able to extract any extra useful information out of the new electrodes. The increased complexity and noise introduced by the new electrodes outweighed any extra useful information contained in the electrodes.

VI. Conclusion and Future Work

The final results showed that the artificial neural network was able to accurately reduce the amount of data across multiple patients by over 98%, a significant improvement over the template matching system used as a baseline. Furthermore, it was found that increasing the number of electrodes did not increase the accuracy of the system, but decreased it instead.

ACKNOWLEDGMENT

The authors would like to thank Dr. Ken Ono, Dr. Victor Gonzalez and the MCV Department of Neurology for their help in supplying the data as well as valuable expert knowledge regarding EEG reading and epileptic spike information.

REFERENCES

[1] C.W Ko, Y.D. Lin, H.W. Chung, G.J. Jan, "An EEG spike detection algorithm using artificial neural network with multi-channel correlation" in Proc. of Intl. conf. on *Engineering in Medicine and Biology Society,* vol. 4, pp. 2070-2073, 1998.

[2] O. Özdamar and T. Kalayci, "Detection of spikes with artificial neural networks using raw EEG," in *Computers and Biomedical Research*, vol. 31, issue no. 2, pp.122-142, 1998.

[3] H.K Garg and A.K. Kohli, "EEG Spike Detection Technique Using Output Correlation Method: A Kalman Filtering Approach," in *Circuits, Systems, and Signal Processing*, vol. 34, no. 8, pp. 2643-2665, 2015.

[4] D.T. Barkmeier, A.K. Shah, D. Flanagan, M.D. Atkinson, R. Agarwal, D.r. Fuerst, K. Jafari-Khouzani, J.A. Loeb, "High inter-reviewer variability of spike detection on intracranial EEG addressed by an automated multi-channel algorithm," in *Clinical Neurophysiology*, vol. 123, no. 6, pp.1088-1095, 2012.

[5] S. Chaibi, T. Lajnef, A. Ghrob, M. Samet, A. Kachouri, "A Robustness Comparison of Two Algorithms Used for EEG Spike Detection," in *The open biomedical engineering journal*, vol. 9, p.151, 2015.

[6] Y.C. Liu, C.C.K. Lin, J.J. Tsai, Y.N. Sun, "Model-based spike detection of epileptic EEG data," in *Sensors*, vol. 13, no. 9, pp.12536-12547, 2013.

[7] M.H. Zarifia, N.K. Ghalehjogh, M. Baradaran-nia, "A new evolutionary approach for neural spike detection based on genetic algorithm, " in *Expert Systems with Applications*, vol. 45, no. 1, pp.462-467, 2015.

[8] T. Kalayci, Ö Özdamar, "Wavelet preprocessing for automated neural network detection of EEG spikes," in *Engineering in Medicine and Biology Magazine,* vol. 14, no. 2, pp.160-166, 1995.

[9] V.P. Nigam, D. Graupe, "A neural-network-based detection of epilepsy," in *Neurological Research*, vol. 26, no. 1, pp.55-60, 2004.

[10] A.J. Gabor, R.R. Leach, F.U. Dowla, "Automated seizure detection using a self-organizing neural network," in *Electroencephalography and clinical Neurophysiology*, vol. 99, no. 3, pp.257-266, 1996.

[11] W.R.S. Webber, R.P. Lesser, R.T. Richardson, K. Wilson, "An approach to seizure detection using an artificial neural network (ANN),," in *Electroencephalography and clinical Neurophysiology*, vol. 98, no. 4, pp.250-272, 1996.

[12] R.Q. Quiroga, Z. Nadasdy, Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," in *Neural computation*, vol. 16, no. 8, pp.1661-1687, 2004.

[13] C.J. James, R.D. Jones, P.J. Bones, G.J. Carroll, "Detection of epileptiform discharges in the EEG by a hybrid system comprising mimetic, self-organized artificial neural network, and fuzzy logic stages," in *Clinical Neurophysiology*, vol. 110, no. 12, pp.2049-2063, 1999.

[14] A. Subasi, E. Erçelebi, "Classification of EEG signals using neural network and logistic regression," in *Computer methods and programs in biomedicine*, vol. 78, no. 2, pp.87-99, 2005.

[15] Y. Kumar, M.L. Dewal, R.S. Anand, "Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network," in *Signal, Image and Video Processing*, vol. 8, no. 7, pp.1323-1334, 2014.

[16] R. Aliabadi, F. Keynia, M. Abdali, "Epilepsy Seizure Diagnosis in EEG by Artificial Neural Networks," in *Majlesi Journal of Multimedia Processing*, vol. 2, no. 2, 2013.

[17] S.M. Akareddy, P.K. Kulkarni, "EEG signal classification for epilepsy seizure detection using improved approximate entropy," in *International Journal of Public Health Science* , vol. 2, no. 1, pp.23-32, 2013.

[18] R. Kottaimalai, M.P. Rajasekaran, V. Selvam, B. Kannapiran, "EEG signal classification using Principal Component Analysis with Neural Network in Brain Computer Interface applications," in *Intl. conf. on Emerging Trends in Computing, Communication and Nanotechnology,* pp.227-231, March 2013.

[19] R. Khosrowabadi, Chai Quek; Kai Keng Ang; A. Wahab, "ERNN: A Biologically Inspired Feedforward Neural Network to Discriminate Emotion From EEG Signal," in *Neural Networks and Learning Systems*, vol.25, no.3, pp.609-620, March 2014.

[20] Yuanfang Ren; Yan Wu, "Convolutional deep belief networks for feature extraction of EEG signal," in *Intl. Joint Conf. on Neural Networks*, pp.2850-2853, July 2014.

[21] M. Teplan, "Fundamentals of EEG measurement", in *Measurement science review*, vol. 2, no. 2, pp.1-11, 2002.

[22] A.C. Guyton, Text Book of Medical Physiology Saunders, Philedelphia, PA, 1986

[23] E. Niedermeyer, F.D. Silva, Electroencephalography: Basic Principals, Clinical Applications and Related Fields, Baltimore, MD. Williams and Wilkins, 1999.

[24] E. Asano, O. Muzik, A. Shah, C. Juhasz, D.C. Chugani, S. Sood, et al., "Quantitative interictal subdural EEG analyses in children with neocortical epilepsy," in *Epilepsia*, vol. 44, pp. 425-434, 2003.

[25] E.D. Marsh, B. Peltzer, M.W. Brown III, C. Wusthoff, P.B. Storm Jr, B. Litt, et al., "Interictal EEG spikes identify the region of electrographic seizure onset in some, but not all, pediatric epilepsy patients," in *Epilepsia*, vol. 51, pp. 592-601, 2010.

[26] A.T. Tzallas, D.G. Tsalikakis, E.C. Karvounis, L. Astrakas,M. Tzaphlidou, M.G. Tsipouras, S. Konitsiotis, "Automated epileptic seizure detection methods: a review study," in *INTECH Open Access Publisher*, 2012.

[27] K.S. Anusha, M.T. Mathews, S.D. Puthankattil, "Classification of Normal and Epileptic EEG Signal Using Time & Frequency Domain Features through Artificial Neural Network," in *Intl. Conf. on Advances in Computing and Communications*, pp.98-101, Aug. 2012

[28] EGI. (2016, April). Dense Array EEG Neuroimaging [online]. Available: https://www.egi.com/clinical-division/clinical-division-care-center/clinical-division-dense-array-neuroimaging

[29] S.B. Wilson, R. Emerson, "Spike detection: a review and comparison of algorithms," in *Clinical Neurophysiology*, vol. 113, no. 12, pp. 1873-1881, 2002.

[30] J.J. Halford, "Computerized epileptiform transient detection in the scalp electroencephalogram: Obstacles to progress and the example of computerized ECG interpreteation," in *Clinical Neurophysiology*, vol. 120, no. 11, pp. 1909-1915, 2009.

[31] N. Acir, I. Öztura, M. Knutalp, B. Baklan, "Automatic detection of epileptiform events in EEG by a three-stage procedure based on artificial neural networks," in *Intl. Trans. on Biomedical Engineering*, vol. 52, no. 1, pp. 30-40, 2005.