Krig-Detect: Exploring Alternative Kriging Methods for Real-Time Seizure Detection from EEG Signals

I. L. Olokodana¹, S. P. Mohanty², and E. Kougianos³
Smart Electronic Systems Laboratory (SESLS)¹,²,³
University of North Texas, TX, USA.¹,²,³
Email: IbrahimOlokodana@my.unt.edu¹, saraju.mohanty@unt.edu², elias.kougianos@unt.edu³
Outline of the talk

- Epileptic Seizures
- Kriging Methods
- Novel Contributions
- Brain as a Spatial Object
- Proposed Seizure Detection Models
- Experimental Results
- Conclusion & Future Research
What is Seizure?

- A seizure is an abnormal activity in the nervous system which alters the functioning of the brain and causes victims to lose consciousness and control.
The Research Vision

- EEG Signal Preparation
- Feature Extraction
- Seizure Detection

Cloud

Persistent Storage

Quick Response Due to Early Detection

Smart EEG Cap

Wearable Edge Computing Device

Seizure Crisis
Characteristics of EEG Signal

- **Complexity**: Highly complex signals depicting the cortical electrical activities in the brain.
- **Intensity**: Low intensity signal measured in µV.
- **Frequency**: Frequency ranges from 0.5–30Hz. Signals are classified as delta, theta, alpha and beta based on frequency.

Source: https://www.researchgate.net/figure/Scalp-EEG-signal-for-a-tonic-clonic-epileptic-seizure-TCES-recorded-at-the-central_fig2_283645061
What are the Research Problems?

- Automatic Seizure Detection.
- Seizure Detection Latency.
- Extended Training Time.
- Patient-Specific Seizure Detection.
- Mobile and Portable Seizure Detection.
- Seizure Crisis Intervention Mechanism.
- Seamless, consistently accurate seizure detection system.
- Low-power seizure detection system.
What are the Challenges?

- Collecting a custom data-set due to the stringent regulations that are involved in collecting data from animals or human subjects.

- The same also applies to testing our models directly on human or animal subjects.

- It is difficult to estimate the level of noise in public data-sets since the conditions of the environment in which they are collected are not known.
## Related Research in Seizure Detection – EEG/ML

<table>
<thead>
<tr>
<th>Reference</th>
<th>Dataset</th>
<th>Processing</th>
<th>Classifier Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alejandro et al (2017)</td>
<td>CHB-MIT</td>
<td>Time Analysis</td>
<td>Thresholds</td>
</tr>
<tr>
<td>Zhou et al. (2012)</td>
<td>Freiburg Database</td>
<td>Wavelet</td>
<td>Bayesian Method</td>
</tr>
<tr>
<td>Acharya et al. (2012)</td>
<td>Self-recorded</td>
<td>Freq. Analysis</td>
<td>SVM, KNN</td>
</tr>
<tr>
<td>Khan et al. (2012)</td>
<td>Self-recorded</td>
<td>Wavelet</td>
<td>LDA</td>
</tr>
<tr>
<td>Runarsson et al (2005)</td>
<td>Self-recorded</td>
<td>Time Analysis</td>
<td>SVM</td>
</tr>
<tr>
<td>Rezvan et al. (2017)</td>
<td>Bonn Dataset</td>
<td>Wavelet</td>
<td>MLP</td>
</tr>
<tr>
<td>Mursalin et al. (2017)</td>
<td>Bonn Dataset</td>
<td>Time Analysis</td>
<td>Random Forest</td>
</tr>
<tr>
<td>Guo et al. (2010)</td>
<td>Bonn Dataset</td>
<td>Wavelet</td>
<td>ANN</td>
</tr>
<tr>
<td>Mitra et al. (2009)</td>
<td>Texas’ Children</td>
<td>Freq. Analysis</td>
<td>ANN</td>
</tr>
<tr>
<td>Zandi et al. (2010)</td>
<td>Vancouver GH</td>
<td>Wavelet</td>
<td>Thresholds</td>
</tr>
</tbody>
</table>
Related Research in Seizure Detection – Non-EEG

IBM’s Implantable Seizure Detector

- IBM is developing an implantable seizure detector by leveraging on their neurosynaptic computing hardware called TrueNorth.
- The TrueNorth chip is postage stamp-sized and consumes over 1,000 times less power than a conventional processor of similar size.

Source: http://uberveillance.squarespace.com/?category=health_care
Consumer Electronics for Seizure Detection

■ Embrace2: Smart-band which uses machine learning to detect convulsive Seizures and notifies caregivers.

Source: https://spectrum.ieee.org/the-human-os/biomedical/diagnostics/this-seizuredetecting-smartwatch-could-save-your-life

■ Medical grade smart watch: It detects generalized clonic-tonic Seizures and notifies physicians.

Source: https://www.empatica.com/embrace2/
What are the Drawbacks of Existing Works?

- High seizure detection latency.
- Models are complex and unrealistic for real-time deployment in the Internet of Medical Things (IoMT).
- Lack of adequate intervention mechanism after detection.

Research Question and Hypothesis

- How effective will Kriging Methods perform on a seizure detection problem given that the brain is structured like a geo map?
- Which Kriging method is best-suited for seizure detection?
- Is it possible to run a seizure detection algorithm on the edge rather than the cloud to achieve a better latency, without significant compromise on accuracy?
Novel Contributions of the Current Paper

- To the best of the authors’ knowledge, this is the first work where multiple Kriging methods have been used for real-time seizure detection in an edge computing paradigm.

- A novel achievement of an epileptic seizure detection latency of less than 1 second while maintaining a comparable accuracy with existing models and $O(1)$ time and space complexity for edge computation.
Kriging was originally developed as a geo-statistical model for spatial prediction.

It is a stochastic process that is governed by a mean value and the relative co-variances of known data points with respect to an unknown.
Why Kriging?

- The brain can be modeled as a spatial map on which spatial data processing methods can be applied.
- Kriging method performs very well even on a relatively small dataset unlike machine learning algorithms. This is very important because of the difficulty in obtaining biomedical datasets.
- Kriging model comes with a variance estimate which gives the level of confidence of the model in a given prediction.
- Kriging model is very reliable without requiring the use of many hyperparameters.
Brain as a Spatial Map

Existing Applications of Kriging

- Seismic intensity analysis (De Rubeis et al., 2005).
- Hydrology and well selection (Virdee et al., 1984).
- Geodesy and geology (Reguzzoni et al., 2005).
- Structural reliability (Kaymaz et al., 2005).
- Mixed signal design optimization (Mohanty et al., 2015).
- Cellular network optimization (Braham et al., 2014).
Types of Kriging

- Simple
- Ordinary
- Universal
- Indicator
- Lognormal
- Disjunctive
- CoKriging
- Bayesian
The Kriging Process

- There are three important steps in the application of Kriging methods.
- First is the establishment of spatial continuity through the semi-varioogram which is a function of the variations in values over distance.
- Second is fitting a model to the generated semi-variogram.
- The final step is the actual estimation through the fitted model.
The Semi-Variogram

- The semi-variogram is merely a scatter plot with each point representing the average variation among a group of location pairs with common distance known as the lag vector \( h \).

- where \( \gamma(h) \) represents the semi-variogram at the lag vector \( h \) between two points, \( N(h) \) is the number of lag vectors \( h \) considered for a single point on the semi-variogram plot.

\[
\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z(x_i) - Z(x_i + h))^2,
\]

The semi-variogram model simply fits a line or curve on the scatter plot represented by the semi-variogram.

There are different types of Semi-Variogram Models as shown below:

- Linear
- Nugget Effect
- Spherical
- Exponential
- Gaussian
Gaussian Semi-Variogram Model

- EEG time-series recorded from normal and epileptic patients were congruent with Gaussian stochastic process. Hence the choice of Gaussian Semi-Variogram Models.

\[
\gamma(h) = \begin{cases} 
C \left[ 1 - \exp \left( -\frac{h^2}{a^2} \right) \right] & h > 0 \\
0 & h = 0 
\end{cases}
\]


- C is the sill (total variance contribution) and a is the range (distance on the horizontal axis corresponding to the sill).
Kriging Estimates

- Given the following set of observations $x_1, x_2, ..., x_n$ as inputs, and $y(x_1), y(x_2), ..., y(x_n)$ as outputs, the input-output relationship based on Kriging is given by:

$$y(x_i) = \mu + Z(x_i),$$

- Where $i$ is the data point index, $\mu$ is a mean constant and $Z(x_i)$ is a Gaussian process.

- The weights between the unknown and each of the known can be obtained by solving the following equation, where $C(.)$ is covariance between two points:

$$\sum_{j=1}^{n} \lambda_j C(x_i, x_i) = 2C(x_o, x_i).$$

- Hence, the final estimate can be obtained as:

$$y(x_o) = \sum_{i=1}^{n} \lambda_i Z(x_i) + (1 - \sum_{i=1}^{n} \lambda_i) \mu_z,$$
Experimental Results - EEG Dataset

BONN DATASET

- The datasets were originally collected from 5 healthy volunteers & five epilepsy patients by the University of Bonn. 5 different sets of data were collected as sets A, B, C, D, & E.

- Sets A&B are healthy signals, C&D are inter-ictal signals while E is the only set with ictal signals. Each of the sets comprises 100 EEG segments which were collected with a 128-channel EEG system sampled at 173.61 Hz.
CHB-MIT SCALP EEG DATASET

- This dataset was collected at the Children Hospital Boston (CHB) in conjunction with the Massachusetts Institute Of Technology (MIT). It is therefore referred to as the CHB-MIT Scalp EEG Database.

- The EEG signals were collected from 22 epileptic patients of CHB using a 23-channel EEG, sampled at 256Hz and labeled according to the subjects as chb01 to chb23.

- The dataset consists of a total of 916 hours of continuous EEG recordings across all 22 subjects.
Figure shows the plot of the Discrete Wavelet Transformation (DWT) coefficients after decomposition using Daubechies Wavelet of order four (db4).

The final output of the decomposition is shown in the table below:

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>43.4 - 86.8</td>
</tr>
<tr>
<td>D2</td>
<td>21.7 - 43.4</td>
</tr>
<tr>
<td>D3</td>
<td>10.9 - 21.7</td>
</tr>
<tr>
<td>D4</td>
<td>5.4 - 10.9</td>
</tr>
<tr>
<td>D5</td>
<td>2.7 - 5.4</td>
</tr>
<tr>
<td>A5</td>
<td>0 - 2.7</td>
</tr>
</tbody>
</table>
Features of EEG Signal

- Signal Entropy
- Fractal Dimension
- Signal Power
- Standard Deviation
- Singular Value Decomposition Entropy
- Maximum Fractal Length
- Hjorth Parameters
- Hurst Exponent
- Lyapunov Exponent etc.

![Graph showing EEG signals with calculations](image)

\[ HJC = 2.231992, \quad ENT = 0.577971, \quad PFD = 1.011104 \]

\[ HJC = 1.685002, \quad ENT = 0.635090, \quad PFD = 1.007473 \]
Feature Representation of Dataset

![EEG Dataset](image)

- SVD Entropy
- Hjorth Complexity
- Fractal Dimensions
Experimental Results

![Graphs showing semivariance over lag vector in hours.](image-url)
Experimental Results

<table>
<thead>
<tr>
<th>Confidence Intervals</th>
<th>Simple Kriging</th>
<th>Ordinary Kriging</th>
<th>Universal Kriging</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.7% CI</td>
<td>97.50%</td>
<td>97.50%</td>
<td>80.00%</td>
</tr>
<tr>
<td></td>
<td>94.74%</td>
<td>94.74%</td>
<td>100.00%</td>
</tr>
<tr>
<td>95.4% CI</td>
<td>92.50%</td>
<td>92.50%</td>
<td>80.00%</td>
</tr>
<tr>
<td></td>
<td>94.74%</td>
<td>94.74%</td>
<td>90.48%</td>
</tr>
<tr>
<td>68.2% CI</td>
<td>90.00%</td>
<td>87.50%</td>
<td>80.00%</td>
</tr>
<tr>
<td></td>
<td>89.47%</td>
<td>84.21%</td>
<td>71.43%</td>
</tr>
</tbody>
</table>

Kriging Models

<table>
<thead>
<tr>
<th>Detection Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Kriging</td>
</tr>
<tr>
<td>Ordinary Kriging</td>
</tr>
<tr>
<td>Universal Kriging</td>
</tr>
</tbody>
</table>
## Comparison with Related Works

<table>
<thead>
<tr>
<th>Published Works</th>
<th>Extracted Features</th>
<th>Classification Algorithm</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Detection Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoeb, et al. 2010</td>
<td>Spectral, temporal and spatial features.</td>
<td>Support Vector Machine (SVM)</td>
<td>NA</td>
<td>96.00%</td>
<td>4.2 sec.</td>
</tr>
<tr>
<td>Zandi, et al. 2012</td>
<td>Regularity, energy &amp; combined seizure indices</td>
<td>Cumulative Sum (CUSUM) thresholding</td>
<td>NA</td>
<td>91.00%</td>
<td>9 sec.</td>
</tr>
<tr>
<td>Altaf, et al. 2015</td>
<td>Digital hysteresis</td>
<td>Linear Support Vector Machine (LSVM)</td>
<td>NA</td>
<td>95.70%</td>
<td>1 sec.</td>
</tr>
<tr>
<td>Vidyaratne, et al. 2017</td>
<td>Fractal dimension, spatial/temporal features</td>
<td>Relevance Vector Machine (RVM)</td>
<td>99.80%</td>
<td>96.00%</td>
<td>1.89 sec.</td>
</tr>
<tr>
<td>Our ICCE 2020 Paper</td>
<td>Petrosian fractal dimension</td>
<td>Kriging Classifier</td>
<td>100.00%</td>
<td>100.00%</td>
<td>0.85 sec.</td>
</tr>
<tr>
<td>Curr. Paper</td>
<td>Fract dim., Hjorth comp.&amp; Entropy</td>
<td>Kriging Classifier</td>
<td>97.50%</td>
<td>94.74%</td>
<td>0.81 sec.</td>
</tr>
</tbody>
</table>
Conclusions

- This results in this presentation demonstrate the effectiveness of Kriging method for accurate and early seizure detection.
- The detection of seizure onset takes place in near real time with an average detection latency of 0.81 second which is better than previous models in the literature.
- Three different Kriging methods were compared for Seizure Detection and results show that Simple Kriging is a slight favorite over Ordinary Kriging while Universal Kriging is far behind them.
Future Research

- In future, we will investigate seizure prediction, which means having prior knowledge that a seizure will occur before it actually does.

- Another future research is to have unified systems that detects seizure before it happens, and then injects drug or performs other control measures right after that.

- We also intend to add security and privacy features to the overall system as it is IoMT-enabled and always connected to Internet.

- We will also use more sophisticated and power-efficient edge devices such as IBM’s neurosynaptic hardware in validating our models.
References


THANK YOU