Kriging-Bootstrapped DNN Hierarchical Model for Real-Time Seizure Detection from EEG Signals

Ibrahim L. Olokodana
Computer Science and Engineering
University of North Texas, USA.
Email: IbrahimOlokodana@my.unt.edu

Saraju P. Mohanty
Computer Science and Engineering
University of North Texas, USA.
Email: saraju.mohanty@unt.edu

Elias Kougianos
Electrical Engineering
University of North Texas, USA.
Email: elias.kougianos@unt.edu

***Abstract*—The Deep Neural Network (DNN) model is known for its high accuracy in classification tasks due to its intrinsic ability to learn the underlying patterns existing in a set of data. Hence it has gained momentum in seizure detection research, as in many other fields. However, its high performance is at the expense of an extensive training time. This is not appropriate for a real-time application such as seizure detection in which a swift reaction is required to save the life of the patient. This paper presents a novel Kriging-Bootstrapped Deep Neural Network hierarchical model for early seizure detection in which Kriging is first used to generate a well-correlated intermediate data set from the original input. The correlated data is then fed into the DNN for the final training. Experiments were carried out using electroencephalogram (EEG) data from both normal and epileptic patients. Results show that, with the same architecture and data size, the cumulative training time of the Krigging-Bootstrapped DNN is about 75% lower than that of the ordinary DNN without a compromise in performance as the proposed hybrid model shows a slightly better accuracy than the baseline DNN model.**

***Index Terms*—Smart Healthcare, Brain, Seizure Detection, Epilepsy, Edge Computing, Kriging Methods, EEG**

I. Introduction

Epileptic seizure is a neurological disorder with incidence rate of over 100 in every 100,000 [1]. It is more prevalent in low-income countries [2] where there is minimal access to standard medical facilities and patients would most often resort to self-help for their medical needs. Such regions of the world will benefit immensely from a simple but fast, accurate and inexpensive seizure detection solution. Delayed attention to a suffering patient in a seizure crisis may cause severe injuries or even death in the worst cases.

Deep Neural Networks (DNN) are quite accurate when trained with sufficient data but are highly computationally expensive in time. This is because the DNN learns slowly the latent relationship between the input and output within the dataset via a complex optimization mechanism [3]. Kriging methods, on the other hand, are popularly used in geostatistics to estimate quantitative values using the principles of spatial continuity to establish correlation between locations in a geographical space [4], [5]. Recent studies have also shown that the brain is similar to a spatial map [6], [7]. A hierarchical blend of Kriging and DNN is therefore promising for seizure detection tasks. Since Kriging is suitably matched for the brain as a spatial map, it easily generates correlated data for the DNN, hence making the actual training by the DNN less laborious. Fig. 1 presents the concept of the proposed Kriging-Bootstrapped DNN model.

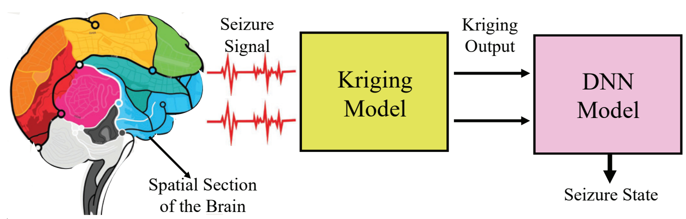

Fig. 1: Conceptual diagram of Kriging-Bootstrapped DNN.

The rest of this paper is organized as follows: Section II reviews related works in seizure detection. Section III highlights the novel contributions of this paper. The proposed novel Kriging-Bootstrapped DNN model for real-time seizure detection is described in section IV. Section V discusses the experimental validation of the proposed model while the conclusion and future works are stated in Section VI.

II. Related Works in EEG-based Seizure Detection

DNNs have been previously used for EEG-based seizure detection but without much attention to the training time [8], [9]. Gaussian process modeling was proposed for neonatal seizures in [10], although it was not in real-time. However, the ordinary Kriging method, which is also known as Gaussian process modeling, was proposed for real-time seizure detection using edge computing by the authors of this paper in a previous work [11]. They also explored different types of Kriging for seizure detection in another related work [12].

A variety of other methods apart from DNN and Kriging have been proposed in the current literature for seizure detection with EEG signals. These include Artificial Neural Network (ANN) [13], Support Vector Machines (SVM) [14], κ-Nearest Neighbors (κ-NN) [15] and Naive Bayes classifier [16] among others. This paper is however, to the best of the authors' knowledge, the first work that uses Kriging-Bootstrapped DNN for EEG-based seizure detection. The Kriging-Bootstrapped Neural Network was originally proposed in [17] and [18] for process variation analysis and design of mixed signal integrated circuits. [19] and [20] are seizure detection systems which are not based on EEG.

III. Novel Contributions of the Current Paper

A. Research Question

A DNN is somewhat accurate for classifying seizure signals [9]. However, it takes too long to train, which may limit its usefulness for real time seizure detection applications. Although training is not done in real time, a quick turn-around is needed to consistently update the already trained model that is used in a real time seizure detection system. How can the training time for a DNN be reduced in seizure detection applications without a compromise in performance? Since Kriging methods can estimate from the known data highly-correlated values which are previously unknown, will intermediate output from a Kriging model improve DNN training time? This paper aims to provide answers to these research questions.

B. Proposed Solution

This paper proposes a novel Kriging-Bootstrapped DNN hierarchical model for fast and accurate seizure detection where the Kriging model generates a correlated intermediate output that is used to train the DNN for the final output. Extensive experiments were carried out using a specific DNN model as a base line for comparing our proposed hierarchical Kriging-Bootstrapped DNN model.

C. The Novelty of the Proposed Solution

While there have been seizure detection models that proposed either a DNN or a Kriging solution in the literature, there was not found a seizure detection that combines both into a single model. This is the first Krigging-Bootstrapped hierarchical model for real-time seizure detection to the best of the authors' knowledge. Furthermore, the proposed hybrid model achieved a 75% reduction in training time and also improves the performance of the DNN by at least 2.5% after training on the same data size and the same DNN architecture.

IV. A Novel Kriging-Bootstrapped DNN Model for Real-Time Seizure Detection

A DNN is adept at learning a pattern in a body of data if any exists and then establishes a suitable relationship between the input and the corresponding output. It performs especially well where there is sufficient data that is commensurate with the complexity of its architecture [3]. Experience shows that some patterns are easier to learn than others. The more difficult it is for the DNN to find a pattern, the longer the training time it takes to build a model. Kriging on the other hand has been very effective in geostatistics for finding spatial estimates based on the spatial correlation that exists between different locations in the region of interest [5]. Given some data samples in different locations with some degree of spatial continuity, Kriging can be used to generate accurate estimates at other locations, thereby increasing the correlation within the data samples. Kriging has since found application in other fields for estimating values based on the correlation within the data samples. Kriging has been successfully applied to the seizure detection problem by modeling the brain as a spatial map.

Our proposed hierarchical model generates a highly correlated intermediate data set from a Kriging model. The intermediate data serves as an input to the DNN. Since the intermediate data is more correlated than the original dataset, it is expected that the DNN will spend less time in identifying a coherent pattern within the dataset from which an input-output relationship can be built, hence reducing the training time. The proposed Kriging-Bootstrapped DNN hierarchical model for real-time seizure detection is shown in Fig. 2. The collected EEG signals from the patient are first preprocessed to obtain relevant features. The intermediate Kriging model is then created from the extracted features to generate a highly correlated input for the DNN.

A. The Bootstrapped Kriging Model

The term "Bootstrap" as used in this work was borrowed from Bootstrap sampling which refers to a method of estimating the true statistical value of a population from some given samples [21]. It involves sampling from the available samples with replacement to create multiple sets of samples. The statistical value of interest is evaluated for each set to create a single sample set from which the final estimation of the population statistical value can be made. The number of sample sets that is created from the original samples is called Bootstrap Size. The higher the Bootstrap Size, the closer the statistical value is to the true population value. In other words, Bootstrap Sampling involves increasing the size of a sample set via Monte Carlo simulation in order to bootstrap to the true statistical value of the population.

In the context of this work, Bootstrapping refers to the process of increasing the correlation between data points in the sample field by increasing the number of data points per unit area via Kriging. Hence it is called a Kriging-Bootstrapped model. Fig. 3 is a schematic illustration of our proposed Kriging Bootstrap method where the red points represent the initial samples, the blue points represent the newly Kriged data points and the dashed lines signify the correlation between data points. Shorter dashed lines imply stronger correlation between the points while the longer dashed lines mean otherwise. Hence, higher correlation exists between points on the right than on the left.

A variety of Kriging methods are available in the literature [22]. In the current paper, we explore the use of ordinary Kriging. The ordinary Kriging estimate $y(\mathbf{x}_o)$ for each of the blue points in Fig. 3 is given by the following expression:

$$y(\mathbf{x}_o) = \sum_{i=1}^{n} \lambda_i \mathbf{R}(\mathbf{x}_i), \tag{1}$$

where y is the estimated value, $\mathbf{x}_o$ is the location of the bootstrap point for which the estimate is made (blue points

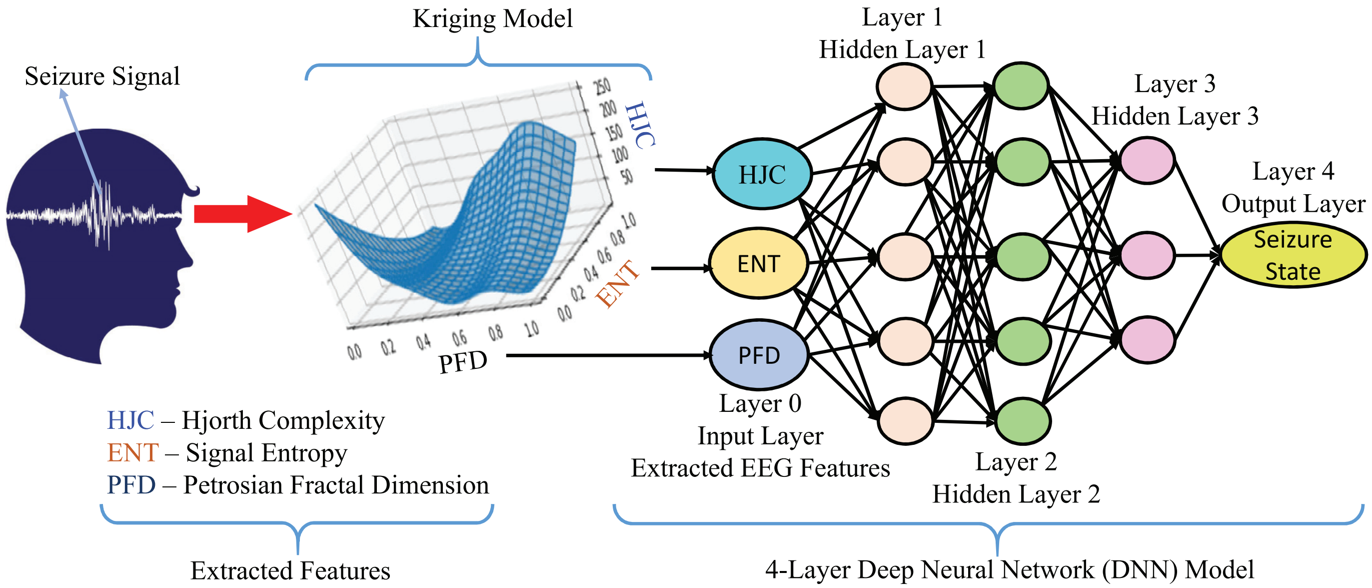

Fig. 2: Proposed Kriging-Bootstrapped DNN hierarchical model for real-time seizure detection.

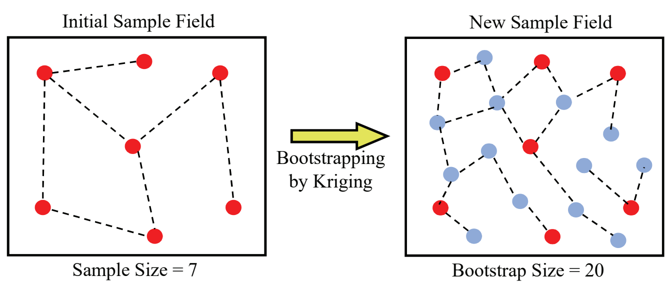

Fig. 3: Schematic illustration of proposed Kriging Bootstrap.

in Fig. 3), $\mathbf{x}_i$ are the locations of the original samples (red points in Fig. 3), and

$$\mathbf{R}(\mathbf{x}_i) = \sum_{i=1}^{n} (Z(\mathbf{x}_i) - \mu_z), \tag{2}$$

is the residual at each point $\mathbf{x}_i$, i.e. the difference between the Gaussian process $Z(\mathbf{x}_i)$ at $\mathbf{x}_i$ and the global mean μ_z. The weights λ_i in eq. (1) are obtained from:

$$\boldsymbol{\lambda}_{n\times 1} = (\mathbf{C}_{n\times n})^{-1}\mathbf{c}_{n\times 1}, \tag{3}$$

where $\mathbf{C}_{n\times n}$ is a covariance matrix of all sample pairs, $\mathbf{c}_{n\times 1}$ is a covariance vector of individual points with respect to the unknown, $\boldsymbol{\lambda}_{n\times 1}$ is the vector of weights and n is the number of points.

B. The Deep Neural Network

A deep neural network is a neural network with more than one hidden layer. The typical neural network has one input layer, one hidden layer and one output layer. The neural network operates by minimizing the cost function which represents the sum of the error margin between a predicted value and the true value via the gradient descent algorithm. The cost function is given by the following expression:

$$J(\omega, b) = -\frac{1}{n}\sum_{i=1}^{n}\big[y_i \log \hat{y}_i + (1 - y_i)\log(1 - \hat{y}_i)\big], \tag{4}$$

where ω refers to the weights, b is the bias, while y_i and $\hat{y}_i$ represent the true and the predicted outputs respectively. i refers to individual sample and n is the total number of samples.

The gradient descent iterative equations for minimizing the cost are the following:

$$\omega := \omega - \alpha\frac{\partial J(\omega, b)}{\partial \omega} \tag{5}$$

and

$$b := b - \alpha\frac{\partial J(\omega, b)}{\partial \omega}, \tag{6}$$

respectively, for finding the weight-bias pair at which the cost function $J(\omega, b)$ is minimum. In the above expression, α is called the learning rate.

The final prediction $\hat{y}$ for a single neuron are then given by:

$$\hat{y} = f(\omega^T\mathbf{X} + b), \tag{7}$$

where ω and b are the values obtained from gradient descent equations (5) and (6) respectively and f is the activation function. The cost function and its derivative are computed by forward propagation and backward propagation respectively.

The proposed deep neural network model for this work, as shown in Fig. 2, is a 4-layer fully connected deep neural network with three hidden layers. The input layer is not counted as a layer in the network. The first two hidden layers have five neurons each while the third hidden layer has three neurons.

V. Experimental Validation of the Proposed Model

A. The Dataset and Features Extracted

The dataset used for this work are EEG signals collected from patients with epilepsy as well as healthy individuals [23]. The extracted features from the EEG signals are Hjorth Complexity (HJC), Signal Entropy (ENT) and Petrosian Fractal Dimension (PFD). A single feature vector is created from these features and is then used to generate a Kriging model. Fig. 4 shows typical examples of a seizure EEG signal and a healthy EEG signal taken from the dataset with their corresponding features extracted.

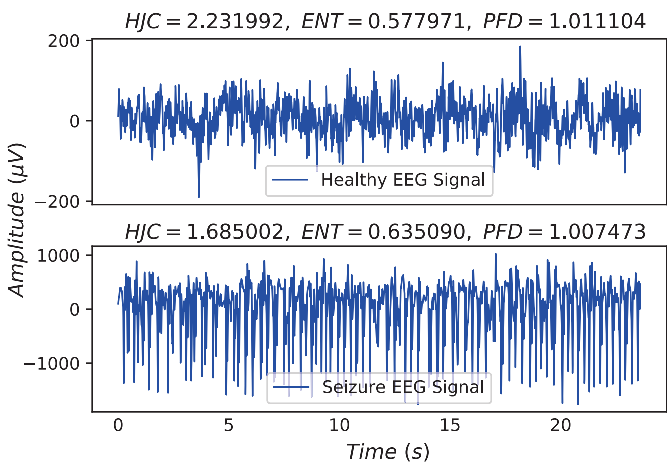

Fig. 4: EEG segments at healthy and seizure states.

B. Experimental Results and Their Analysis

Several experiments were performed using different data sizes and training epochs while comparing the proposed Kriging-Bootstrapped DNN model with the baseline ordinary DNN. Tables I and II show the performance of the proposed hierarchical model and the baseline DNN model respectively using the same DNN architecture while Table III compares both models' best performances with respect to training time, training epochs and accuracy. It is observed from Table I that more training time and more epochs are required to achieve a decent performance with the baseline DNN model whereas Table II indicates that the proposed Kriging-Bootstrapped model converges quickly to a very good performance within a short training time and reduced training epochs. Note that the recorded training time for the Kriging-Bootstrapped DNN in Tables II and III is a total sum of the Kriging time and the subsequent training time for the DNN.

As observed in Table III, the training time for Kriging-Bootstrapped DNN is reduced by 75%, testing accuracy is improved by 2.5% and 30 times fewer training epochs are required than that of the baseline DNN model. It is however noted that the ordinary DNN model has a slightly better training accuracy by a marginal 0.07%. Fig. 5 is a plot comparing the testing accuracy of both models with respect to their training times while Fig. 6 shows a plot of testing accuracy against the respective training epochs. It is clear from Figures 5 and 6 that the proposed Kriging-Bootstrapped DNN model achieved a higher accuracy in much shorter training time and fewer training epochs than the ordinary DNN model.

TABLE I: Baseline DNN model performance results with 10,000 samples.

Count	Training Accuracy	Testing Accuracy	Training Epochs	Training Time
1	99.82%	80.00%	800	4.29s
2	99.85%	82.50%	1000	5.13s
3	99.95%	92.50%	10000	37.46s
4	99.99%	97.50%	45000	173.57s
5	99.99%	97.50%	50000	199.66s

TABLE II: Kriging-Bootstrapped DNN model performance with 10,000 samples.

Count	Training Accuracy	Testing Accuracy	Training Epochs	Training Time
1	99.14%	97.50%	500	41.07s
2	99.76%	100.00%	800	41.73s
3	99.84%	100.00%	1000	42.02s
4	99.92%	100.00%	1500	43.83s
5	99.92%	100.00%	10000	80.99s

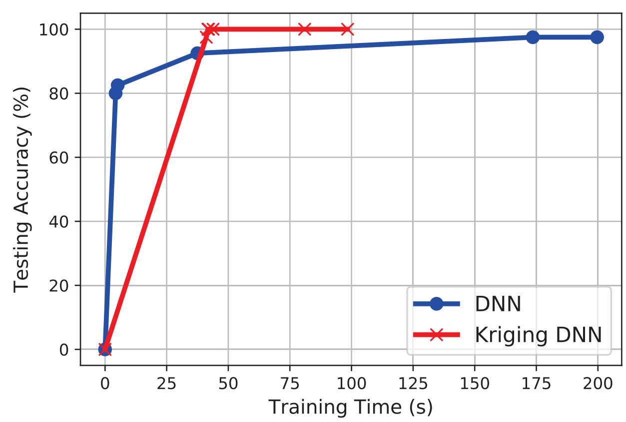

Fig. 5: Performance Comparison of DNN model vs Kriging-Bootstrapped DNN model with respect to training time.

VI. Conclusion and Future Work

A novel Kriging-Bootstrapped DNN hierarchical model for real-time seizure detection from EEG signals was presented in this paper. The proposed model was compared to a baseline DNN model and the performance results demonstrate that the proposed model trains in 75% less time and 30 times reduced training epoch size than the ordinary DNN, as well as a 2.5%

TABLE III: Comparing best performances for DNN and Kriging-Bootstrapped DNN models.

Models	Training Accuracy	Testing Accuracy	Training Epochs	Training Time
DNN	99.99%	97.50%	45000	173.57s
Kriging DNN	99.92%	100.00%	1500	43.83s

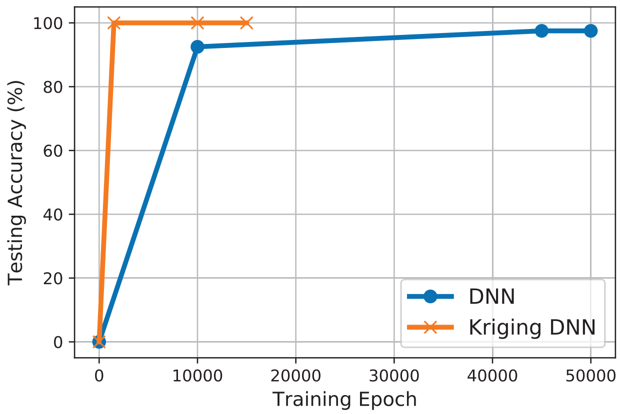

Fig. 6: Performance Comparison of DNN model vs Kriging-Bootstrapped DNN model with respect to training epoch.

improvement in testing accuracy. This proves that the proposed Kriging-Bootstrapped DNN model will be a better choice for real-time seizure detection.

While in the current method ordinary Kriging method has been used, in future work we will explore other Kriging methods. We will study the effectiveness of other Kriging methods for hierarchical machine learning (ML) modeling. Similarly other types of DNN models and with different number of hidden layers is a future research. The authors intend to implement this model in edge computing hardware for real time seizure detection. Predicting the onset of seizure rather than detection is another area the authors look to explore in the near future. Integration of seizure control methods along with the detector to have a unified seizure detection and control in the IoMT framework is another future research.

REFERENCES

[1] I. Commission Report, "The epidemiology of the epilepsies: future directions," *International League Against Epilepsy. Epilepsia*, vol. 38, pp. 614–618, 1997.

[2] C. R. Newton and H. H. Garcia, "Epilepsy in poor regions of the world," *The Lancet*, vol. 380, no. 9848, pp. 1193–1201, 2012.

[3] Q. V. Le, J. Ngiam, A. Coates, A. Lahiri, B. Prochnow, and A. Y. Ng, "On optimization methods for deep learning," in *Proceedings of the 28th International Conference on International Conference on Machine Learning*. Omnipress, 2011, pp. 265–272.

[4] N. Cressie, "The Origins of Kriging," *Math. Geology*, vol. 22, no. 3, pp. 239–252, Apr 1990.

[5] M. J. Pyrcz and C. V. Deutsch, *Geostatistical reservoir modeling*. Oxford university press, 2014.

[6] M. Zaleshina and A. Zaleshin, "The Brain as a Multi-layered Map. Scales and Reference Points for Pattern Recognition in Neuroimaging," *European J. of Geog.*, vol. 8, no. 1, pp. 6–31, February 2017.

[7] E. I. Moser, Y. Roudi, M. P. Witter, C. Kentros, T. Bonhoeffer, and M.-B. Moser, "Grid cells and cortical representation," *Nature Reviews Neuroscience*, vol. 15, no. 7, p. 466, 2014.

[8] M. A. Sayeed, S. P. Mohanty, E. Kougianos, and H. P. Zaveri, "Neuro-Detect: A Machine Learning-Based Fast and Accurate Seizure Detection System in the IoMT," *IEEE Transactions on Consumer Electronics*, vol. 65, no. 3, pp. 359–368, August 2019.

[9] H. G. Daoud, A. M. Abdelhameed, and M. Bayoumi, "Automatic epileptic seizure detection based on empirical mode decomposition and deep neural network," in *Proc. IEEE 14th Int. Collq. on Sig. Proc. & Its App. (CSPA)*, 2018, pp. 182–186.

[10] S. Faul, G. Gregorcic, G. Boylan, W. Marnane, G. Lightbody, and S. Connolly, "Gaussian process modeling of EEG for the detection of neonatal seizures," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2151–2162, 2007.

[11] I. L. Olokodana, S. P. Mohanty, and E. Kougianos, "Ordinary Kriging-based Real-Time Seizure Detection in an Edge Computing Paradigm," in *Proceedings of IEEE 38th International Conference on Consumer Electronics (ICCE)*, 2020, p. in Press.

[12] I. L. Olokodana, S. P. Mohanty, and E. Kougianos, "Krig-Detect: Exploring Alternative Kriging Methods for Real-Time Seizure Detection from EEG Signals," in *Proceedings of 6th IEEE World Forum on Internet of Things*, 2020, p. in Press.

[13] V. P. Nigam and D. Graupe, "A neural-network-based detection of epilepsy," *Neurological Res.*, vol. 26, no. 1, pp. 55–60, January 2004.

[14] A. Subasi and M. I. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert sys. with app.*, vol. 37, no. 12, pp. 8659–8666, December 2010.

[15] M. Behnam and H. Pourghassem, "Singular Lorenz Measures Method for seizure detection using κNN-Scatter Search optimization algorithm," in *Proceedings of the Signal Processing and Intelligent Systems Conference (SPIS)*, 2015, pp. 67–72.

[16] A. Sayeed, S. P. Mohanty, E. Kougianos, V. P. Yanambaka, and H. Zaveri, "A Robust and Fast Seizure Detector for IoT Edge," in *Proc. IEEE Int. Conf. Smart Elect. Sys. (iSES)*, 2018, pp. 156–160.

[17] O. Okobiah, S. P. Mohanty, and E. Kougianos, "Kriging bootstrapped neural network training for fast and accurate process variation analysis," in *Proc. 15th International Symposium on Quality Electronic Design*, 2014, pp. 365–372.

[18] S. P. Mohanty, E. Kougianos, and V. P. Yanambaka, "Ultra-fast variability-aware optimization of mixed-signal designs using bootstrapped kriging," in *Proc. 16th International Symposium on Quality Electronic Design*, 2015, pp. 239–242.

[19] A. Marquez, M. Dunn, J. Ciriaco, and F. Farahmand, "iSeiz: A low-cost real-time seizure detection system utilizing cloud computing," in *Proc. IEEE Glob. Hum. Tech. Conf.*, 2017, pp. 1–7.

[20] P. M. Vergara, E. de la Cal, J. R. Villar, V. M. González, and J. Sedano, "An IoT platform for epilepsy monitoring and supervising," *J. Sensors*, vol. 2017, July 2017.

[21] F. A. Wichmann and N. J. Hill, "The psychometric function: II. Bootstrap-based confidence intervals and sampling," *Perception & psychophysics*, vol. 63, no. 8, pp. 1314–1329, 2001.

[22] S. P. Mohanty, *Nanoelectronic mixed-signal system design*. McGraw-Hill Education New York, 2015, no. 0071825711.

[23] 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*, vol. 64, no. 6, p. 061907, November 2001.