Performance Comparison of Classification Algorithms for EEG-based Remote Epileptic Seizure detection in Wireless Sensor Networks

Khalid Abualsaud[2, 1], Massudi Mahmuddin[2], Mohammad Saleh[1], Amr Mohamed[1]
[1]Department of Computer Science & Engineering, College of Engineering,
Qatar University, P.O. Box 2713 Doha, Qatar
[2]Computer Science Dept., Graduate School of Computing,
University Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia
k.abualsaud@qu.edu.qa, ady@uum.edu.my, mohd.saleh@qu.edu.qa, amrm@qu.edu.qa

***Abstract* - Identification of epileptic seizure remotely by analyzing the electroencephalography (EEG) signal is very important for scalable sensor-based health systems. Classification is the most important technique for wide-ranging applications to categorize the items according to its features with respect to predefined set of classes. In this paper, we conduct a performance evaluation based on the noiseless and noisy EEG-based epileptic seizure data using various classification algorithms including BayesNet, DecisionTable, IBK, J48/C4.5, and VFI. The reconstructed and noisy EEG data are decomposed with discrete cosine transform into several sub-bands. In addition, some of statistical features are extracted from the wavelet coefficients to represent the whole EEG data inputs into the classifiers. Benchmark on widely used dataset is utilized for automatic epileptic seizure detection including both normal and epileptic EEG datasets. The classification accuracy results confirm that the selected classifiers have greater potentiality to identify the noisy epileptic disorders.**

Keywords- EEG; epileptic seizure; feature extraction; classifiers; classification accuracy.

I. INTRODUCTION

Electroencephalography (EEG) is a commonly used technique in detecting human healthcare for remote epileptic seizure detection. EEG is a technique of sensing electrical impulses produced by the brain neurons and detected by electrodes placed on the scalp [1]. The recording of the brain activity is achieved by placing electrodes on the scalp that measures the abnormal variations in the voltage impulses produced by the brain [2], and sending the acquired data wirelessly for remote analysis over sensor network. Seizure is one of the healthcare applications where tonic-clonic and epileptic are the most common types of seizure. The tonic-clonic seizure is manifested by sudden, repeated, and frequent rhythmic muscle movements and is often without a warning. While, epileptic seizures are repeated seizures are not related to acute illness or brain injury and they affect millions of people. Extremely long seizure can lead to neuronal damage, coma or death [3]. EEG can deliver valuable perception into disorders of the brain activity, and express the brain electrical activity. In this context, the seizure-free interval from epilepsy patients has been considered as an important component for the prediction process and diagnosis [4, 5 and 6]. Accordingly, clinicians can evaluate the disorders of a patient's brain from EEG and perform further diagnosis. Therefore, the recognition and analysis of the EEG signals is very important task. This could be difficult, because the size and form of these signals may change eventually due to different types of noise e.g. hardware, or communication noise. Many tools, methods and algorithms of signal processing theory have been proposed, explained and implemented [1]. A critical neurological disease emerging from abnormal discharges of the brain activity is represented by epilepsy. Epilepsy leads to the trembling and uncontrolled movements. Epilepsy is a neurological condition, which disturbs the nervous system due to brain injury. It has been reported that the epileptic disease affects about 1% of the world population and medication is of no help to about 30% of them [7]. However, the visual inspection of EEG signals can be very difficult and time consuming due to the lengthy inspection that will increase human's error from high level of concentration [8]. Therefore, machine intelligence techniques are proposed to enhance the process of epileptic seizure detection.

In this paper, we present the application of the reconstructed and noisy (uncertainty) EEG epileptic seizure dataset. Five different categories of classification algorithms have been used to analyze the dataset in order to find out which algorithm performs better in terms of its performance metrics: *Accuracy and Computation time* according to the *compression ratio* (*CR*). In this work, a set of classifiers has been experimented, namely BayesNet, DecisionTable, IBK, J48, and VFI. It is expected that these classifiers may have different strategies' regarding the current system state. Therefore, the contribution of each classifier in the overall classification process depends on how certain is the classifier is regarding the current hypothesis. This scheme provides an effective way of dealing with the uncertainty problem mentioned above. Extensive experimental work has been conducted. The results have demonstrated the effectiveness of the proposed technique as a 98.49% classification accuracy has been obtained. Finally, the analysis shows that some classification algorithms perform well over EEG epileptic seizure dataset.

The remaining parts of this paper are structured as follows. In Section II, related work is presented. Materials and Methods, which include description of EEG data, classification framework, feature extraction, and classifiers selections have been described in Section III. Experiments and results of the performance evaluation study are illustrated in Section IV, and the paper is concluded in Section V.

II. Related Work

Seizure signs can be categorized as clinical or electrical EEG. The clinical signs include physical behaviors such as continued open eyes with visual fixation, blinking repeated, and other slight facial appearances. Unlike adults, the children's clinical signs are minimized and therefore, it requires a constant attention of the medical staff to be detect[19]. Mirowski et al. [20] evaluated out-of-sample seizure predication performance in patients with epilepsy EEG, and then compared each combination of feature type and classifier. Classification methods and the success of pattern recognition have been given based on machine learning. The Freiburg dataset has been used to evaluate the prediction methods for classification of EEG signals in epilepsy. Support vector machines, logistic regression or convolutional neural networks are used as machine learning-based classifiers to discriminate interictal from preictal patterns of features. Results show that the proposed technique is outperforming previous seizure prediction methods on the Freiburg dataset. Another research [21] conducted performance analysis of EEG patters using Discrete wavelet Transform (DWT) and Independent Component Analysis (ICA). DWT & ICA have been used for feature extraction in the principle of time – frequency domain analysis. These features are used as input for the SVM and ANN for EEG classification. SVM and Neural Network algorithms have been implemented to detect epileptic seizure for the classification stage. The methods are then tested on only both data sets of EEG data (Sets H and S) for classification between normal and seizure signals of the same dataset. Orosco *et al.* [22] applied computational analysis and measures for quantifying and characterizing fractal behaviors and complexity such as the Hurst exponent, the scaling exponent, fractal dimension and various forms of entropy in epilepsy research to characterize epileptic seizure detection. The work focused on ANN and SVM as the seizure detection algorithms and presented a comparison between their performances. Chauhan *et al* [9] presented a comparative study for ten classification algorithms on NSL- Knowledge Discovery in Databases (KDD) intrusion detection dataset. J48, BayesNet and IBK classification algorithms were among these algorithms based on their performance metrics to find out the best suitable algorithm available. Cross validation of 10-fold has been used to measure the performance of the classification models. A data mining tool has been used to perform experiments and assessments of these methods using NSL-KDD intrusion detection dataset. The study shows that decision trees; J48, BayesNet; classifiers are the best classification of the intrusion. On the other hand, comparing the time consumed by the individual techniques reveals that the IBK is the fastest. Goyal and Mehta [10] conducted a comparison on the performance evaluation for Naïve Bayes and J48 classification algorithms. Both algorithms are used with the given Bank dataset from UC Irvine Machine Learning Repository with 300 instances. The classification accuracy and cost analysis are calculated. The results show that the classification accuracy of J48 classifier is 52.67%, which is better than the Naïve Bayes classifier, while Naïve Bayes is better than J48 classifier in terms of cost analysis of 155. This result is due to the simplicity of J48 technique which is based on building a decision tree. Another research work in [11] adopts various feature selection techniques in order to achieve optimal subset of features for student performance model for predictive accuracy of Educational Data Mining dataset. Six classifiers from different categories are selected as base classifiers and these are J48, IBK, Kmeans Clustering, NaiveBayes Updatable, ONER, and VFI Classifiers. The classifiers are used to rank 15 features of ASSISTments Platform dataset. In addition, the classifiers have been applied on the ranked features to get the optimal subset of features. The performance of these classification methods has been evaluated based on their predictive accuracy basis. The results show that the six classifiers give the best accuracy for only 3-7 features with 80% reduction in dataset size without sacrificing the performance and processing time. Authors in [12] develop a data acquisition model for railway using rule-based learning method. Rules have been produced with statistical analysis to identify a unique classifier for railway applications. Authors in this research used six classifiers, namely REPTree, J48, Decision Stump, IBK, PART and OneR. PART is moderately a new algorithm for producing "decision lists", ordered sets of rules. Authors applied twenty-five datasets considering track condition, wagon loaded and unloaded condition, data record etc. They generate the rules from the data matrix using the popular rule-based PART algorithm, which is built in machine learning tools. In another research [13], Salama et al., presented a comparison between several classifiers, namely decision tree (J48), Multi-Layer Perception (MLP), Naive Bayes (NB), Sequential Minimal Optimization (SMO), and Instance Based for K-Nearest neighbor (IBK). They used the Wisconsin Breast Cancer dataset and calculated classification accuracy and the confusion matrix based on the 10-fold cross validation method. The confusion matrix is used to display the relationship between predicted classes and outcomes. The results show that the comparison of accuracies of SMO, NB, MLP, J48 and IBK classifiers are (96.9957%), (95.9943%), (95.279%), (95.1359%), and 94.5637% respectively.

Classification techniques reported in the literature namely, BayesNet, DecisionTable, IBK, J48, and VFI have provided satisfactory performance given their work on different datasets. Based on our knowledge there is no reported work of these classifiers on the EEG data. This data is contaminated by different factors e.g. hardware and communication noise. The wireless EEG data is compressed before transmission, which means that on the receiver side some important information may get lost during the reconstruction process. In addition to that, a wireless channel may increase the transmission problem by adding some noise to the transmitted data. Hence, in order to be efficient, a prospective classification technique should take into consideration the uncertainty problem in the EEG data.

III. Materials and Methods

This part presents a benchmarking of EEG dataset for epileptic and non-epileptic seizure, followed by the classification system structure and introduces the selected classifiers that are using in this research.

A. Benchmark EEG dataset

The datasets used in this work originated from [14] being widely used for automatic epileptic seizure detection. Both normal and epileptic EEG datasets have been represented in this benchmark data. The EEG datasets have five sets termed A, B, C, D, and E. For sets A and B, the patient was relaxed and awake with eyes open and eyes closed, respectively. Sets A and B represent healthy subjects and therefore referred to as set A. Sets C and D both contain only the activity measured during seizure-free intervals. Hence, sets C and D represent unhealthy subjects in seizure-free intervals, referred to as set C. Finally, only set E contains seizure activity, which represents epilepsy subjects. Figure 1 illustrates the ideal raw EEG signals of sets A, C, and E used in this work.

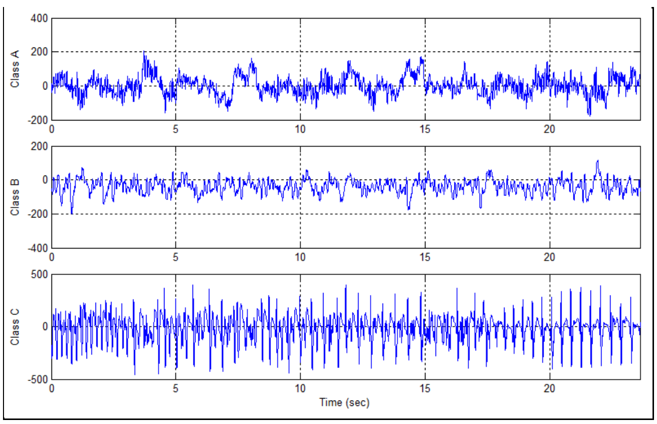

Figure 1: Example of three different classes of EEG signals taken from different subjects.

B. The Classification Framework

Figure 2 shows the classification framework, the original EEG signals x, the data, has been compressed using compressive sensing (CS) technique and discrete cosine transform (DCT) method with a measurement matrix Φ [15]. The main benefit of these techniques is to convert the signal to x_0 sparse signals. These methods have been used in order to compress the EEG data before sending it to the receiver. The spares signal x_0 is related to another compressed signal y. The compressed signal y is transmitted over the Additive White Gaussian Noise (AWGN) channel model [15]. The EEG data is transmitted as noiseless and noisy. In the noiseless case, we assume that the wireless channel is ideal, and in the noisy case we use different signal-to-noise-ratio (SNR) values. When the compressed data is received, at the receiver side, inverse DCT is used to reconstruct the EEG data back to its original size. After the reconstructed EEG epileptic seizure data is categorize into noiseless and noisy with 1 decibel (dB), 5 dB and 10 dB, the discrete wavelet transform (DWT) is used to extract statistical features from the data [15]. Five different categories of classifiers are used at the classification phase during after which the accuracy results are recorded.

N-dimensional representations of the original EEG signal *x* is considered to show the CS compression and reconstruction. Assume that this signal is represented by a projection on a different basis set Ψ:

$$= \sum_{i=1}^{N} x_{0_i} \Psi_i \quad \text{or} \quad x = \Psi\, x_0 \tag{1}$$

where $_0$ is an *N*1* functioning vector of bases, and Ψ is an *N*N* matrix of bases.

The basis (Ψ) in this case is DCT, it can also be Gabor, Fourier, Mexican hat, cubic spline, linear B-spline, or cubic B-spline function. In the projection above, it is assumed that x is related to another signal *y*:

$$y_{[M*1]} = \Phi_{[M*N]} \times \ _{[N*1]} \tag{2}$$

where Φ is a measurement/sensing/random matrix of dimensions *M*N*, and *y* is the compressive sensed version of *x*. Matrix *y* has dimensions *M*1*, and data compression is achieved if *M* < *N*. The compression ratio (CR) is then defined as follows:

$$CR = \left(1 - \frac{M}{N}\right) * 100 \tag{3}$$

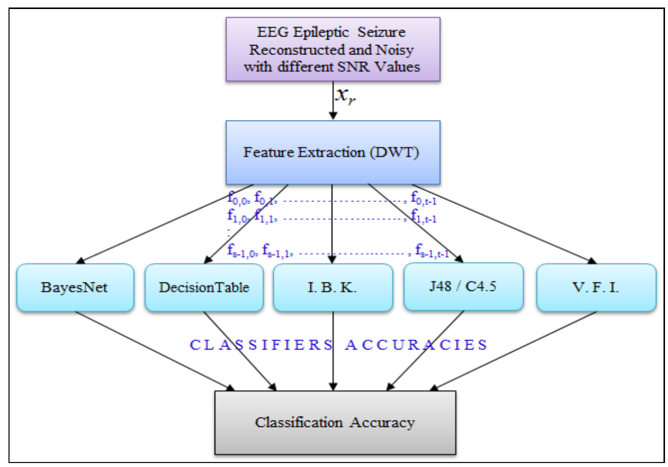

Fig. 2: Classification System Structure

In the noiseless EEG data, we assumed that the wireless channel is ideal during the transmitted compressed EEG data. While in the noisy EEG data, we considered the AWGN as wireless channel with different SNR values. The noise level was controlled by using the SNR=1 decibel (dB), SNR=5 dB and SNR=10 dB to demonstrate data imperfection. Other types of noise can also be incorporated similarly [15].

C. Feature Extraction

The discrete wavelet transform (DWT) method is widely used because EEG signals are generally time varying and space-varying non-stationary signals [15]. Based on the extensive experimental work for the reconstruction of noiseless and noisy EEG, Daubechies 6 with decomposition level 7 have been used, because they provide the optimum level in terms of classification accuracy and computational complexity of the EEG epileptic seizure category of data [15]. To extract statistical features namely, maximum, minimum, mean, and standard deviation from each wavelet sub-band, several implementations have been adopted including wavelet families and decomposition levels. The combined statistical features and wavelet sub-band are used to formulate a feature vector of 32 attributes used to represent an input to the classifiers.

D. Classifier Algorithms Selection

To obtain the accuracy results and the computation time for the classifiers, the classification accuracy has been averaged through ten trials. The classification accuracy (AC) is defined as [16]:

$$AC = \frac{TP+TN}{TP+TN+FP+FN} * 100 \quad (4)$$

where;

TP: is the true positive, the total number of correctly detected positive events;

TN: is the true negative, the total number of correctly detected negative events;

FP: is the false positive, the total number of erroneously positive detections (i.e., false alarms); and

FN: is the false negative, the total number of erroneously negative detections (i.e., missed detections).

In this research, five different classifiers are selected as the base classifiers and these are BayesNet, DecisionTable, IBK, J48, and VFI Classifiers. Each classifier belongs to a different family. BayesNet belongs to Naïve Bayes classifiers, DecisionTable is related to Rules, IBK belongs to Lazy (k-NN) classifiers, and J48 (C4.5) is related to Decision Trees and VFI to general classifiers. Each classifier has shown to be the best classifier in its family [10, 17]. Since each classifier using different classification strategy, it is expected that they may produce different classification results. Following is a brief description of each classifier.

BayesNet (BN) is a popular classification algorithm. The implementation of the BayesNet classifier is mainly a learning algorithm that uses different search algorithms and quality measures. The Base class for a Bayes Network classifier provides data structures based on probability distributions and facilities common to Bayes Network learning algorithms like K_2 and B [10]. According to Mitchell [18], Bayes theorem provides a way to calculate the probability of a hypothesis based on its prior probability, the probabilities of observing various data given the hypothesis and the observed data itself.

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \quad (5)$$

In Equation (5), *P(h)* is the prior probability and holds hypothesis *h*; *P(D)* stands for the prior probability that training data *D* will be observed, and *P(D|h)* is the probability of observing data *D* given some world in which hypothesis *h* holds. The posterior probability *P(h|D)* reflects the confidence that *h* holds after the training data *D* has been seen.

DecisionTable (DT) algorithm is proposed by Ron Kohavi 1995 for building and using decision table majority classifier in order to summarize the dataset in decision table, which contains the same number of attributes as the original dataset. It employs the wrapper method to find a good subset of attributes for insertion in the table. It decreases the probability of over-fitting and generates a smaller decision table by excluding attributes that slightly participated or not to a model of the dataset [11].

IBK (Instance-Based- K-nearest neighbor) is a simple instance-based learner K-nearest neighbor classifier. IBK is the operation of the *k*-nearest neighbor classifier. By simply using the stored dataset, it creates a model. Using a distance metric, it classifies the new data by comparing its items with the memorized data items. The new items are assigned to the category of the closest original data item. The number of nearest neighbors (*k*) can be set manually, or determined automatically using cross-validation [9, 11].

J48 classifier is a standard algorithm producing decision tree models and is widely used in practical machine learning. A decision tree is a tool to accomplish classification in three steps. 1) considers all the data instances as input, 2) generates the rules to carry out the classification task, and, 3) derive the class level. It works by forming pruned partial decision trees (built using C4.5's heuristics as the most popular tree classifier), and directly converting them into a corresponding rule [10].

VFI (Voting Feature Intervals) classification algorithm is based on the real-valued of voting. Feature intervals are then constructed for each feature dimension for each class. Among classes, each feature participates in the voting. The class that obtains the maximum number of votes is stated to be the predicted class. VFI represents each training sample as a vector of features with a label that represents the sample class [11].

$$ACC - VFI = \frac{vote\ [C_j]}{\sum_{i=1}^{k} vote\ [C_i]} \quad (6)$$

where C_j is the probability of the class.

IV. EPERIMENTS AND RESULTS

In this research work, the EEG-based epileptic seizure data has been classified into two different categories of data, noiseless and different values of SNR to show the effect of noisy data on classification. Two types of experiments with compressed EEG-epileptic seizure data are conducted. The first experiment is when the compressed EEG data is noiseless; we consider that the wireless channel is ideal. The second experiment considers the AWGN as a wireless channel with add noise of SNR= 1, 5, and 10 dB values. The performance results of the studied classifiers are reported, illustrated, and discussed.

A K-fold cross validation of the dataset is selected, for each of the K experiments using K-1 folds for training and the remaining ones for testing. The benefit of using K-Fold Cross validation is that all the instances in the dataset are ultimately used for both training and testing. The mechanism of cross validation divides/splits the data into K sets of size N/K, trains on K-1 datasets and tests on 1, then repeats K times and finally calculates the average to obtain the accuracy. In this research, cross validation of 2-fold has been used, which means that 50% of the dataset is used for training and the other 50% is used for testing in order to measure the performance of the classification algorithms reducing the computation time and the number of experiments.

In the first type of experiments, we study the results of the classifiers accuracy against CR in the case when the wireless channel is ideal. Figure 3 illustrates the classification accuracy against compression ratio (CR) using different classifiers, namely BayesNet, DT, IBK, J48, and VFI. The results show that accuracy decreases logarithmically with the increase of CR. We can divide the results into three main regions, at CR = 75%, and 80% respectively. While accuracy remains stable above 95% for all classifiers in the first region, IBK, BayesNet, and J48 seem to have significant accuracy of about 2% over VFI and DT classifiers. The decay in accuracy seems to be reasonable in the second region, showing IBK making the lead in high compression values, and then it starts to decay exponentially in the third region for all classifiers. While IBK is simple instance-based learner K-NN classifier, it outperforms the other four classifiers in most regions using distance metric. When it classifies the new data, it compares the items of the new data with the items of the memorized data.

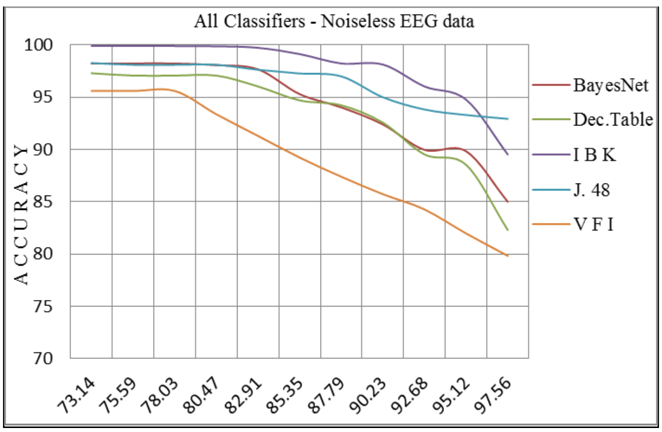

Figure 3: Classification accuracy against CR for noiseless data.

Figures 4-6 correspond to the second type of experiments considering the AWGN as a wireless channel with physical channel impairments using different SNR values at 1, 5, 10 dB respectively. The results of classification accuracy against CR for BayesNet, DT, IBK, J48, and VFI classifiers are reported.

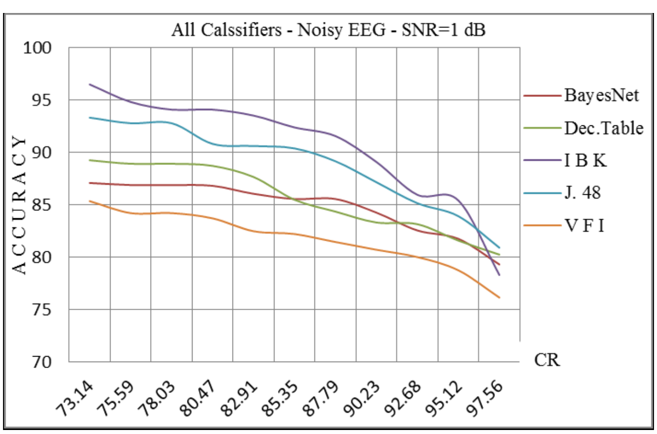

Figure 4: Classification accuracy against CR for SNR=1 dB

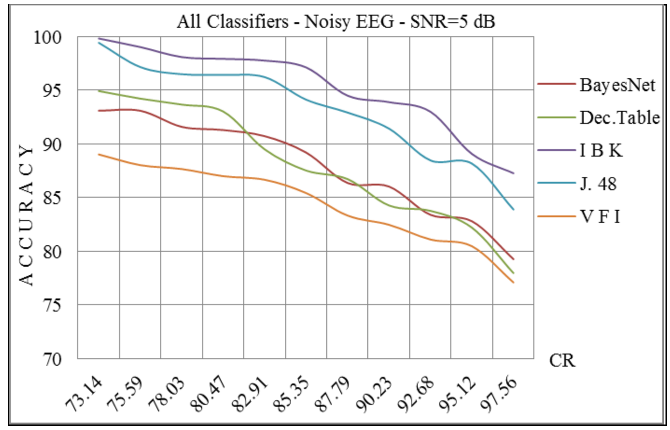

Figure 5: Classification accuracy against CR for SNR=5 dB

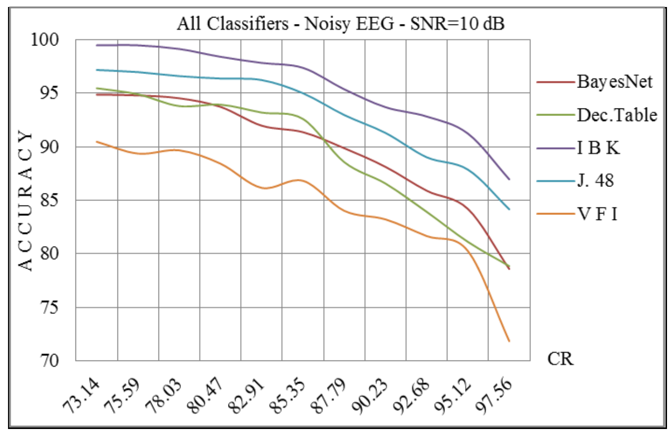

Figure 6: Classification accuracy against CR for SNR=10 dB

Figure 4 shows a slightly different behavior for all classifiers. At SNR=1 dB the noise percentage is too high in the compressed EEG data. While the classification accuracy starts to decay linearly after CR ≈ 78%, the effect of noisy communication is more evident, causing the decrease of all classifiers' accuracies. Nevertheless, each classifier is in its range of accuracy however with decrease in the accuracy values.

The classifiers accuracies decrease consistently when SNR=5 dB, while the exponential decay starts earlier at CR ≈ 80%. Figure 5 shows that the exponential decay starts at CR ≈ 85% while all accuracies are still over 85% for all classifiers. After this point the accuracy decay starts to decrease after CR=85.35%.

Eventually, Figure 6 shows steady decrease against CR, and SNR, nominating *IBK,* which is based k-NN, and *J48*, which produces decision tree models to be the best tolerable classifiers to wireless channel noise, and changes in CR.

From Figures 3-6, we see that the best compression ratio is at ≈ 85% giving the best classification accuracy for almost all classifiers. In addition, the IBK classifier gives a best accuracy around 99%, which is the highest accuracy, since it works based on distance that is the default parameters were used. J48 is the second best accuracy around 97.27% because it creates decision tree tool to achieve the classification accuracy by considering all data as input, generates rules, and finally obtain the class level. BN and DT are interfering in the accuracy at different CR because BN organizes the data based on the probability distributions and facilitate the common one, while DT creates a small decision table by excluding the attributes that are not slightly participated or not participated at all in the model of dataset. Finally, VFI is based on voting and belongs to the general classifiers, each feature contributes in the vote and the class that has a maximum number of votes becomes predicted class. Table 1 shows the classification accuracy at CR=85.35% for all classifiers.

Table 1: Accuracy at the CR = 85.35%

Data Types	BayesNet	Dec.Table	IBK	J.48	VFI
Noiseless	95.30	94.70	99.13	97.27	89.23
SNR=10	91.37	92.60	97.37	95.00	86.83
SNR=5	89.20	87.53	97.13	94.13	85.43
SNR=1	85.57	85.50	92.40	90.40	82.23

At the compression ratio of 85.35% the classification accuracy was around 85% accuracy. Compared with previous works, the classification accuracy of noiseless EEG data, 99% achieved for these algorithms, which is 9% higher than the work reported in [15 and 16] (90% accuracy), and 9.50% higher than the work done in [7] (89.50%) for the noiseless data. On the other hand, 12.40% higher than the research work in [15] for the noisy EEG data with SNR=1 dB considering the same dataset. The above table 1 shows that also at the same compression ratio of 85.35% the achieved classification accuracy is over the 85% for the noisy data with high number of SNR, which is 1 dB for all classifiers. Figure 7 shows that all the classifiers decay as a function of the SNR using the same way at the desired CR.

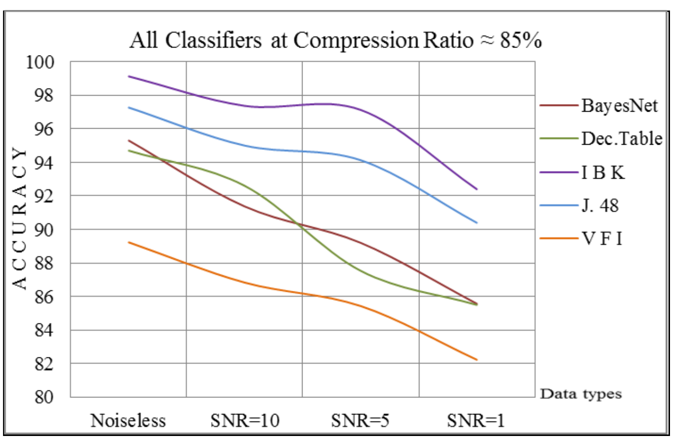

Figure 7: Classification accuracy for all classifiers at CR=85.35%

V. CONCLUSION

In this research work, reconstructed and noisy EEG data as well as five different classifiers have been utilized for epileptic seizure detection applications. The EEG data was compressed using CS and iDCT methods for low complexity for the reconstruction of the EEG data. Features were extracted from the reconstructed data using DWT. The proposed classification system structure has been applied on data sets A, C, and E of the EEG-based epileptic seizure application to measure the data accuracy. We have also investigated the impact of the wireless channel on the transmission of the compressed EEG data, showing the effect of wireless channel impairments. The results revealed that IBK with accuracy close to 99% outperforms the other four classifiers namely, BayesNet, DT, IBK, J48, and VFI when used with A, C, and E datasets. On the other hand, VFI has lowest accuracy among the other classifiers and therefore being impractical for tele-monitoring applications. VFI belongs to the general classifiers family and is only based on the majority vote. The results also show that J48 demonstrates the most stable accuracy as it tolerates imperfection of data due to channel noise, and high compression values; recall that J48 works based on the *k*-NN mechanism. For the analytical model that captures the classification accuracy as a function of compression, these results can be a basis or rule at the transmitter. This can be used to optimize the performance achievement and predict the classifier performance when the transmitter changes the compression to react to channel breakdown.

VI. CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

VII. ACKNOWLEDGEMENT

This work was made possible by; NPRP 09 - 310 - 1 - 058 from the Qatar National Research Fund, a member of Qatar Foundation. The statements made herein are solely the responsibility of the authors.

VIII. REFERENCES

[1] W. Xue-qing and Z. Shu-qin, “Research on Efficient Coverage Problem of Node in Wireless Sensor Networks”, in the Proceedings of International Conference on Industrial Mechatronics and Automation, IEEE ICIMA, PP.: 9-13, May 2009.

[2] Niedermeyer E, Lopes da Silva F. “Electroencephalography: Basic principles, clinical applications, and related fields”, Lippincot Williams & Wilkins; 2004.

[3] http://www.healthline.com/health/epilepsy, Atlanta, GA 30333, USA.

[4] Ogulata SN, Sahin C, Erol R. “Neural network-based computer aided diagnosis in classification of primary generalized epilepsy by EEG signals”, Journal of Medical Systems, Vol. 33; pp.107-112, 2009.

[5] Adeli H, Zhou Z, Dadmehr N. “Analysis of EEG records in an epileptic patient using wavelet transform”, Journal of Neuroscience Methods, Vol. 123; pp.69-87, 2003.

[6] Alkan A, Koklukaya E, Subasi A. “Automatic seizure detection in EEG using logistic regression and artificial neural network”, Journal of Neuroscience Methods, vol. 148 pp.:167-176, 2005.

[7] A. Mohamed, K. B. Shaban, and A. Mohamed, "Evidence Theory-based Approach for Epileptic Seizure Detection using EEG Signals", in the proceeding of 12th IEEE International Conference on Data Mining Workshops (ICDMW), IEEE 2013, Belgium.

[8] J. Wang, P. Liu, M. F.H. She, S. Nahavandi, A. Kouzani, “*Bag-of-words representation for biomedical time series classification*”, Journal of Biomedical Signal Processing and Control, Vol, 8 (634– 644), Elsevier 2013.

[9] H. Chauhan, V. Kumar, S. Pundir and E.S. Pilli, “A Comparative Study of Classification Techniques for Intrusion Detection”, in the proceeding of IEEE International Symposium on Computational and Business Intelligence, 2013.

[10] A. Goyal and R. Mehta, “Performance Comparison of Naïve Bayes and J48 Classification Algorithms”, International Journal of Applied Engineering Research, ISSN 0973-4562, Vol.7, No.11, 2012.

[11] H. M. Harb, M. A. Moustafa, “Selecting Optimal Subset of Features for Student Performance Model”, International Journal of Computer Science Issues (IJCSI), Vol. 9, Issue 5, No 1, September 2012

[12] Shafiullah, G. M., Shawkat, Ali A. B. M., Thompson, Adam and Wolfs, Peter J., "Rule-based classification approach for railway wagon health monitoring", in the Proceedings of the IEEE International Joint Conference on Neural Networks 2010, Barcelona, pp. 1-7, 2010.

[13] G. I. Salama, M. B. Abdelhalim, and M. A. Zeid, "Experimental Comparison of Classifiers for Breast Cancer Diagnosis", in the proceeding of 7th IEEE International Conference on Computer Engineering & Systems (ICCES), Cairo, November 27-29, 2012.

[14] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state", Journal of Physical Review E, American Physical Society , Volume 64, 061907, 2001.

[15] K. Abualsaud, M. Mahmuddin, R. Hussein, A. Mohamed, "Performance Evaluation for Compression-Accuracy trade-off using Compressive Sensing for EEG-based Epileptic Seizure Detection in Wireless Tele-monitoring", in the proceeding of IEEE - IWCMC, Italy, July 2013.

[16] S. F. Liang, H. C. Wang, and W. L. Chang, "Combination of EEG complexity and spectral analysis for epilepsy diagnosis and seizure detection", EURASIP Journal Advanced Signal Processing, vol. 2010, p. 853434, 2010.

[17] Ron Kohavi, "The Power of Decision Tables", in the proceeding of 8th European Conference on Machine Learning, 174-189, 1995.

[18] Tom Mitchell, Machine Learning textbook, McGraw Hill, 1997.

[19] M Roessgen, B Boashash, "A Comparison of Two Techniques for Detecting Seizure in Newborn EEG Data", Proceedings of the International Conference on Acoustics, Speech and Signal Processing 6, 3101–3104, Atlanta, Georgia, USA, 1996.

[20] Mirowski P., Madhavan D., LeCun Y., Kuzniecky R., "Classification of patterns of EEG synchronization for seizure prediction", Electroencephalography and Clinical Neurophysiology, 120, 1927–1940, 2009.

[21] M. Stella Mercy, "Performance Analysis of Epileptic Seizure Detection Using DWT & ICA with Neural Networks", International Journal Of Computational Engineering Research (ijceronline.com) Vol.: 2, 1109-1113, Issue: 4, 2012.

[22] Lorena Orosco, Agustina Garces Correa, and Eric Laciar, "Review: A Survey of Performance and Techniques for Automatic Epilepsy Detection", Journal of Medical and Biological Engineering, 33(6): 526-537, 2013.