ENSEMBLE APPROACH ON ENHANCED COMPRESSED NOISE EEG DATA SIGNAL IN WIRELESS BODY AREA SENSOR NETWORK

KHALID AHMED A. ABUALSAUD

DOCTOR OF PHILOSOPHY
UNIVERSITI UTARA MALAYSIA
2015

Permission to Use

In presenting this thesis in fulfillment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

Abstrak

Rangkaian penderiaan kawasan badan tanpa wayar (WBASN) digunakan untuk komunikasi di antara nod pengesan pada atau didalam badan manusia untuk memantau parameter dan pergerakan. Salah satu aplikasi penting WBASN adalah pemantauan kesihatan penyakit kronik seperti serangan epilepsi. Kebiasaannya, data serangan epilepsi elektroensefalograf (EEG) dikumpul dan dipadatkan untuk mengurangkan masa penghantarannya. Namun, pada masa yang sama, keadaan ini mencemarkan keseluruhan data dan merendahkan klasifikasi kejituaannya. Kajian terkini juga tidak mengambil kira data EEG yang besar. Akibatnya, data EEG adalah bersifat intensif jalur lebar. Dengan demikian, tujuan utama kajian ini adalah untuk mereka bentuk satu kesatuan rangka kerja mampatan dan klasifikasi data EEG sebagai menangani isu data bersaiz besar dengan memampatkannya sebelum penghantaran. Satu lagi matlamat adalah untuk menyusun semula data yang telah dimampat dan mengenal pasti kemudiannya. Kerana itu, teknik *Noise Signal Combination* (NSC) dicadangkan untuk pemampatan EEG data yang dihantar dan meningkatkan klasifikasi kejituan di sebelah penerimaan dalam keadaan data hingar dan tidak lengkap. Rangka kerja yang dicadangkan ini menggabungkan penderiaan mampatan dan transformasi konsinus diskret (DCT) untuk mengurangkan jumlah saiz penghantaran data. Tambahan itu, model hingar Gaussian juga digunakan dalam rangka kerja tersebut. Di sebelah penerima, NSC yang dicadangkan direka bentuk bersandarkan kepada wajaran undian menggunakan empat teknik klasifikasi iaitu rangkaian neural buatan, Naïve Bayes, k-Nearest Neighbour, dan Support Vector Machine sebagai input kepada NSC. Keputusan eksperimen telah menunjukkan bahawa teknik yang dicadangkan melangkaui kejituan tertinggi dari teknik konvensional untuk data besar kurang dan tanpa hingar. Tambahan lagi, rangka kerja tersebut berjaya melaksanakan peranan penting dalam mengurangkan saiz data dan pada masa sama meningkatan kejituan untuk kedua-dua data kurang dan tanpa hingar. Sumbangan utama kajian ini adalah kesatuan rangka kerja dan NSC. Keputusan menunjukkan keberkesanan rangka kerja yang dicadangkan dan menyediakan beberapa manfaat yang boleh dipercayai termasuk mudah dan meningkatkan ketepatan kejituan. Akhir sekali, kajian ini dapat menambah baik maklumat klinikal berkaitan bukan sahaja mengenai pesakit yang mengalami epilepsi, tetapi juga gangguan neurologi, masalah mental atau fisiologi.

Katakunci: Rangkaian Pengesan Kawasan Tubuh Tanpa Wayar, Kejituan Klasifikasi, Pengelas Gabungan, Data EEG.

Abstract

The Wireless Body Area Sensor Network (WBASN) is used for communication among sensor nodes operating on or inside the human body in order to monitor vital body parameters and movements. One of the important applications of WBASN is patients' healthcare monitoring of chronic diseases such as epileptic seizure. Normally, epileptic seizure data of the electroencephalograph (EEG) is captured and compressed in order to reduce its transmission time. However, at the same time, this contaminates the overall data and lowers classification accuracy. The current work also did not take into consideration that large size of collected EEG data. Consequently, EEG data is a bandwidth intensive. Hence, the main goal of this work is to design a unified compression and classification framework for delivery of EEG data in order to address its large size issue. EEG data is compressed in order to reduce its transmission time. However, at the same time, noise at the receiver side contaminates the overall data and lowers classification accuracy. Another goal is to reconstruct the compressed data and then recognize it. Therefore, a Noise Signal Combination (NSC) technique is proposed for the compression of the transmitted EEG data and enhancement of its classification accuracy at the receiving side in the presence of noise and incomplete data. The proposed framework combines compressive sensing and discrete cosine transform (DCT) in order to reduce the size of transmission data. Moreover, Gaussian noise model of the transmission channel is practically implemented to the framework. At the receiving side, the proposed NSC is designed based on weighted voting using four classification techniques. The accuracy of these techniques namely Artificial Neural Network, Naïve Bayes, k-Nearest Neighbour, and Support Victor Machine classifiers is fed to the proposed NSC. The experimental results showed that the proposed technique exceeds the conventional techniques by achieving the highest accuracy for noiseless and noisy data. Furthermore, the framework performs a significant role in reducing the size of data and classifying both noisy and noiseless data. The key contributions are the unified framework and proposed NSC, which improved accuracy of the noiseless and noisy EGG large data. The results have demonstrated the effectiveness of the proposed framework and provided several credible benefits including simplicity, and accuracy enhancement. Finally, the research improves clinical information about patients who not only suffer from epilepsy, but also neurological disorders, mental or physiological problems.

Keywords: Wireless Body Area Sensor Network, Classification Accuracy, Ensemble Classifier, Bio-signal.

Acknowledgement

This thesis concludes my Doctor Philosophy Degree in the Wireless Body Area Sensors Networks for Healthcare Applications submitted to the Awang Had Salleh Graduate School (AHSGS), College of Arts and Science (CAS), Universiti Utara Malaysia (UUM), Malaysia.

To begin with, I thank ALLAH. After that, I would like to express my sincere gratitude and appreciation to my supervisor Dr. Massudi Mahmuddin' for his excellent guidance, continuing inspiration, helpful advice, encouragement and support in critical review of the research. In addition, I extend thanks for School of Computing, UUM.

I would also like to thank Dr. Amr Mohamed for his instructing, help and support. I would like to thank Dr. Mohammed Saleh and Dr. Moutaz Saleh for their encouragement and support.

Many thanks to Prof. Qutaibah Malluhi the previous Head of Dept. of Computer Science & Engineering, College of Engineering Qatar University for his encouragement, and support. I also extend my thanks to Prof. Sebti Foufou the new Head of department for his continuing support.

Finally, I dedicate this thesis to my honourable, loving wife, children, family and friends for their patience, encouragement and fully support.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

Abbreviation	Meaning
AC_i	Classifier Accuracy
ANN	Artificial Neural Network
APSNS	Adaptive Particle Swarm Negative Selection
ASAp	Adaptive Sampling Approach
AWC	Application-specific WBASN Communication
BLDA	Bayesian Linear Discriminant Analysis
CCM	Classifier Confusion Matrix
CD(s)	Chronic Disease(s)
COPD	Chronic Obstructive Pulmonary Disease
CR	Compression Ratio
CS	Compressive Sensing
DAC	Detection And Classification
dB	decibel
DCT	Discrete Cosine Transform
DE	Differential Evolution
DF	Data Fusion
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
EEG	Electroencephalogram
EMD	Empirical Mode Decomposition
EMG	Electromyography
FFT	Fast Fourier Transformation

FN	False Negatives
FP	False Positives
HARMONI	Healthcare-oriented Adaptive Remote Monitoring
iDCT	inverse DCT
IMFs	Intrinsic Mode Functions
KF	Kalman Filter
k-NN	K-Nearest Neighbor
LR	Logistic Regression
LS-SVM	Least Square SVM
MAC	Media Access Control
ME	Mixture of Experts’
MLPNN	Multilayer Perceptron Neural Network
MV	Majority Voting
NB	Naïve Bayesian
NDE	Non-Destructive Evaluation
NSC	Noise-aware Signal Combination
nu-SVC	nu-Support Vector Classification
OS	Operating System
PDA	Personal Digital Assistant
PRD	Percentage Root-mean-square Difference
PR_i	Class Precision
PSO	Particle Swarm Optimization
RBFNs	Radial Basis Function Neural Networks
R-CSP	Regularized Common Spatial Patterns
RE_i	Class Recall

RF	Radio frequency
RM	Random Matrix
S	Samples
SAP	Sensing And Preprocessing
SCC	Sensor-based Compression and Classification
SNR	signal-to-noise-ratio
SSC	Signal Strength-based Combining
SSS	Small Sample Setting
SVM	Support Vector Machine
TN	True Negatives
TP	True Positives
WBASNs	Wireless Body Area Sensor Networks
WHO	World Health Organization
Wi-Fi	Wireless Fidelity
WSNs	Wireless Sensor Networks

CHAPTER ONE
INTRODUCTION

1.1 Overview

Networks of wireless sensor devices are being deployed to collectively monitor and disseminate information about a variety of phenomena of interest. A wireless sensor device is a battery-operated device, capable of sensing physical measurements. In addition to sensing, it is capable of wireless communication, data storage, and a limited size of computation and signal processing. Advances in integrated circuit design are continually reducing the size, weight and cost of sensor devices, while simultaneously improving their resolution and accuracy. A wireless sensor network (WSN) consists of a large number of wireless-capable sensor devices working collaboratively to achieve a common objective. A WSN has one or more base-stations, which collect data from all sensor devices. These base-stations are the interface through which the WSN interacts with the outside world [1].

A WSN is an infrastructure-less networks that consists of a number of self-configuring wireless devices capable of sensing vital signs for characterizing contemporary phenomena. A WSN consists of wireless nodes, which measure physical conditions using sensors, digitize it and keep or distribute the measured data over the network. Typical applications include, but are not limited to, data collection, monitoring, and medical telemetry [2]. Several applications have been intended for WSN. These range in scope from military applications, environmental monitoring and medical applications. For instance, WSNs can form a critical part of military command, control, communications, computing, intelligence, surveillance, reconnaissance, and targeting systems. In addition, it can be used in monitoring of friendly enemy forces; equipment and ammunition, monitoring; nuclear, and biological, as well as chemical attack detection.

By embedding a wireless sensor network within a natural environment, collection of long-term data on a previously unattainable scale and resolution becomes possible. It can be able to obtain localized, detailed measurements that are otherwise more difficult to collect. Some of these include habitat monitoring, animal tracking, forest-fire detection, precision farming, and disaster relief applications.

Potential health applications abound for WSNs. Imaginably, hospital patients could be equipped with wireless sensor nodes that monitor the patients' vital signs and track their location. Patients could move about more freely while still being under constant supervision. In case of an accident, the patient trips and falls the sensor could alert hospital workers as to the patient's location and condition. A doctor in close proximity, also equipped with a wireless sensor, could be automatically dispatched to respond to the emergency [2].

At the intersection of engineering and medicine stands a new borderline of advancement, wireless health, which seeks to make healthcare more personalized and proactive [3]. The field of sensor networks encompasses the sensing, signal processing, and communications disciplines and has become one of the most active areas of research in computing. Reduction of form factor also places physical constraints on battery size, which in turn limits the capability of long evaluation. Integrated circuit technology scaling has helped to create intelligent sensors that are progressively evolving into smaller, more capable, and less expensive platforms. The deployment of numerous networked, coordinated sensors enables wide field instrumentation for applications that require a distinct combination of temporal and spatial coverage [4]. Vital drivers of this vision are the novel wearable technologies, known as wireless body area sensor networks (WBASNs), which capture health data on-body, and address fundamental deficiencies in the state of health applications [5].

By definition, WBASNs consist of two or more interconnected nodes comprising sensing and communication capabilities, located on, near, or within a human body. WBASNs supporting peoples by providing healthcare services such as medical monitoring, enhancement memory control of home appliances, medical data access, and communication in emergencies situations. WBASNs instrument the human body and its immediate surroundings. It enables constant monitoring of the health conditions of people with chronic diseases (CDs) [5, 6]. This initiated a comprehensive research effort intended to develop inside a human body biosensors for continuous monitoring of multiple biological relevant diseases. In addition, the authors clarify the important role of WBASNs in medicine to minimize the need for caregivers and help the chronic illness and an independent elderly people live [5, 6]. Specifically, WBASNs consists of multiple on-body and ambient sensor nodes, capable of sampling, processing, and communicating one or more physiological signs (e.g. Heart activity, brain activity, movements, blood pressure and oxygen saturation) over an extended period. Despite, WBASNs making possible a variety of novel applications for healthcare, there existing new challenges and opportunities, which also require a unique research prototype that incorporates a comprehensive perspective of both engineering and medical realities [5, 7].

The sensor data readings are transmitted over a wireless communication channel to a base-station that gather raw data from all sensors then send it to a running application that analyzes and makes decisions based on these readings. Such readings are characterized into physiological signs and measured using different types of sensed signals such as the electrocardiogram (ECG) [8], and acceleration, as well as electroencephalogram (EEG) [9]. In addition, the base-station is used for communication among sensor nodes operating on, or inside the human body in order to monitor vital body parameters and movements as well as to enable its user with quality of life, assisted living, sports, or entertainment purposes [10].

The wireless transceiver of raw data consumes the majority of the measured power of WBASN applications. Simply transmit the minimum size of data is one way to reduce the sensor power consumption by decreasing the time the radio is active [11]. In addition, compression techniques focus on reducing the size of physical data traffic during communication by transmitting a shorter data stream that the sensor sends over the wireless channel. This reducing is to enhance the bandwidth utilization, ultimately reduc power consumption, and possibly speed of processing and memory space required by the application. Furthermore, physiological data detection and classification is very important to the timely diagnosis and analyze potentially fatal and chronic diseases proactively in clinical as well as various life settings [11, 12].

1.2 Research Motivation

The healthcare industry sector is one of the largest and fastest-growing industries in the world, as highlighted in the OnWorld [13]. Consuming more than 10 percent of gross domestic product of almost developed countries, healthcare can establish an enormous size of a country's economy [12]. Therefore, the need for high performance, cost-effective healthcare solutions is one of the crucial strengths for any developing country seeking sustainable future advancements. Using wireless sensors in the field of healthcare is one of the potential areas which is expected to save $25 billion worldwide in the current decade through leveraging cost-effective solutions and applications as highlighted in the OnWorld healthcare market report [13] illustrated in Figure 1.1.

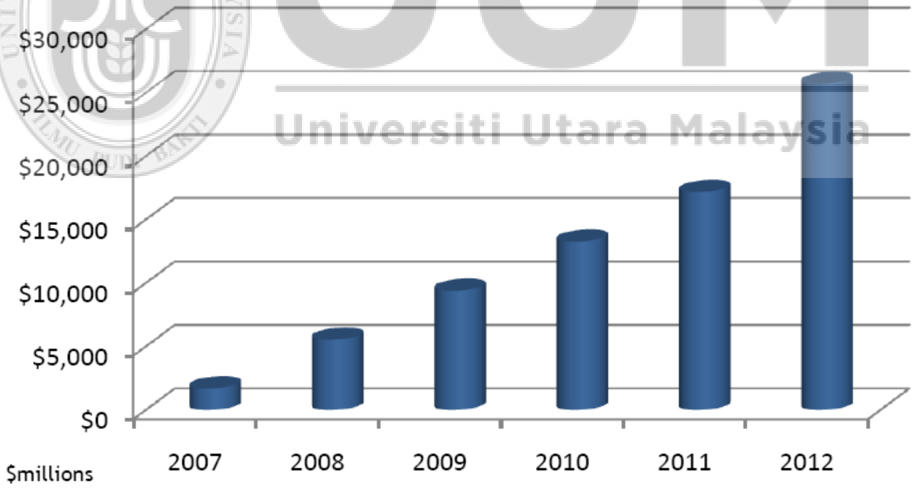


Figure 1.1. Annual savings, WSN-enabled healthcare solutions 2007–2012 [13]

In addition, according to the OnWorld mobile health and wellness report in 2013 [14], the investment in healthcare technologies towards long-term reduction in healthcare costs is dramatically increasing. Thus the mobile health and wellness WSN systems will increase by a 75% compound yearly growth rate and they will make up 57% of global shipments worldwide as shown in Figure 1.2.

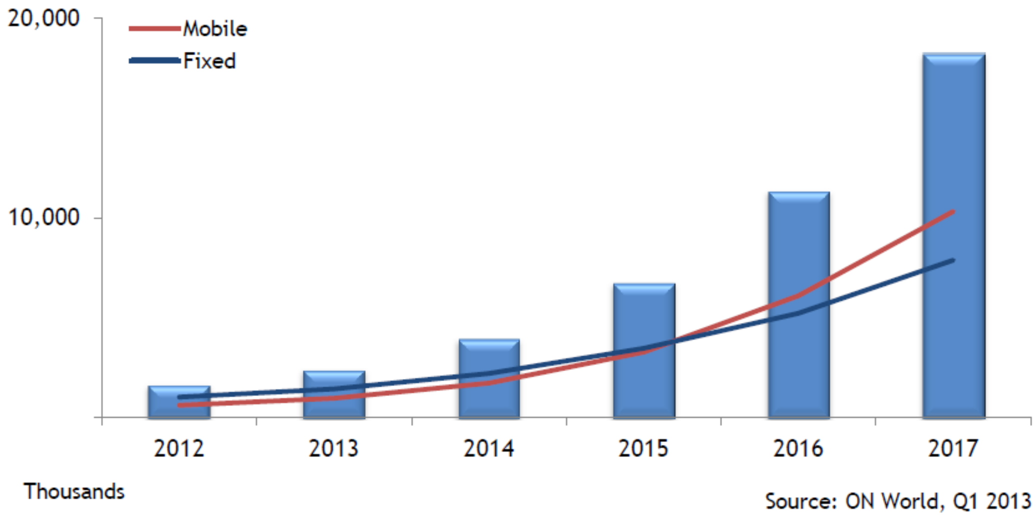


Figure 1.2. Global Health/Wellness WSN System Shipments by 2012-2017 [14]

In recent years, WSN has many advantages such as the flexibility, fault tolerance, high sensing, self-organization, low-cost and rapid deployment characteristics. These advantages of WSN have offered new opportunities in many domains. However, WSNs are mostly applied to those cases, which need a long-term surveillance, such as observing the humidity of the environment. WBASNs for patient monitoring applications have grown significantly and the sampling rate of data acquisition module of those nodes is low and constant [15].

Brain activities have three techniques, namely, functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG), and EEG. fMRI and MEG are neuroimaging procedures that measure brain activity by detecting associated changes in blood flow and recording magnetic fields produced by electrical currents occurring naturally in the brain, respectively [5]. The disadvantages of both of them are that the patient should sleep in a tunnel/scanner (motionless), more expensive, and cannot be applied on epileptic seizures. These restrict the application of these imaging techniques in young children and special populations. While EEG is abnormal impulses produced by the brain, which is utilized in epileptic seizure detection. In addition, exploiting EEG in epileptic seizure detection allows the patient to move freely as in his real daily life, and its process is not expensive [16]. Therefore, mobile EEG is more suitable for remote health monitoring. Moreover, typically, sensors (electrodes/channel) can sample up to 300samples/second, each sample is 2 Bytes (16-bits), which results in the need for an extensive bandwidth required for data transmission. For example, for 100 channels, this can lead to traffic of ~500kbps for one single user. Let alone, the energy consumed in transmitting such huge data. Therefore, reducing the size of data during the transmission using compression of EEG data is desirable [5, 6], while taking into consideration the effect of compression on the classification accuracy. This constitutes the motivation of this thesis.

1.3 Problem Statement

Generally, the 21st century healthcare systems aim to involve citizens and health professionals, especially citizens, to take over a higher level of responsibility for their own health status. Utilizing contemporary technologies such as internet, tablet, and mobile phones enable patients to active participate in treatment and rehabilitation. With good WSNs and data processing capabilities, they are a potential part of the future wireless health care system. The flexibility, fault tolerance, high sensing reliability, low cost, and rapid deployment characteristics of sensor networks create many new and exciting application areas. Indeed, wireless sensors have become an excellent tool for sensor-based personal health monitor [16].

In addition, one of the healthcare applications is seizure due to the abnormal impulses produced by the brain, which are sudden temporary physical movements, behavioral or sensory changes depending on the origin of the impulse. The most common seizures are tonic-clonic and epileptic. The tonic-clonic seizure presents sudden, repeated rhythmic muscle movements and many times without warning. While epileptic seizures are repeated seizures, they are not related to acute illness or brain injury and affect millions of people. Extremely long seizures can lead to neuronal damage, coma or death. Therefore, for those who have an epileptic seizure, it is important to wear a medical identification bracelet to help responders [17].

The most commonly used signal detection technique on human healthcare for epileptic seizure is the EEG which is a technique of recording electrical impulses produced by the brain neurons and detected by electrodes placed on the scalp. Accordingly, clinical clinicians can evaluate the conditions of patient's brain from EEG and perform initial diagnosis. Therefore, the recognition and analysis of the EEG signals is a very important task. This could be difficult, because the size and form of these signals may change eventually or be increased and can be noisy. Many tools, methods and algorithms from signal processing theory have been proposed, described and implemented [18].

Due to the compression process and channel impairments, EEG data will be distorted which may affect the classification accuracy. This thesis is concerned with the development of classification algorithm for compressed and noisy EEG data of epileptic patients [1, 6]. The obtained data was contaminated due to the reconstruction error and the receiver's added noise. This will affect the reconstructed data accuracy due to the loss of some data. Finally, the problem statement deals with the following:

1. The EEG is bandwidth intensive, time-varying and space-varying non-stationary data.
2. Current techniques did not estimate the accuracy in the cases of data recovery and noisy data.
3. Based on the conducted literature survey, most of the current research work did not consider:
    - Compressing the EEG data before transmission, rather most of the current techniques were applied on the raw data.
    - The seizure-free interval subjects (during data collection, the seizure was not active).
    - The effect of reconstruction error due to compression and noisy data

Therefore, the following research questions will be investigated as key research problems:

1. What is the compression ratio leveraging sampling theory that can be used to compress the EEG signal to enhance its classification accuracy without losing its important features?
2. Given the epileptic seizure application, the signal can be down sampled with a rate less than the sampling rate calculated above, such that can be performed with high accuracy in detection and classification within the context of the application of interest.
3. What are the best-selected techniques that are used in the detection and classification of EEG epileptic seizure?

4. Given wireless channel impairments, a noise-aware classification method has been proposed that could achieve high accuracy when reconstructing the EEG compressed data.

1.4 Research Objectives

The main objective of this research is to improve the compression and classification accuracy for noisy EEG-based epileptic seizure. This will be achieved by enhancing compression ratio and by introducing enhancements in both signal processing and classification techniques. Accordingly, this research will focus on reducing the EEG data size by utilizing a Compressive Sensing (CS) technique for sending a short data stream over the wireless channel, which ultimately increases the lifetime of batteries on the sensor devices as part of the continuance and energy-efficient network. Moreover, to achieve higher EEG data classification accuracy, an ensemble classification method is proposed.

Toward this main objective, the following sub-objectives are derived:

1. To propose a framework for compression and delivery of EEG data across wireless channel in order to address the problem of large data size of the WBASN’s EEG data during data transmission. Large data size requires longer transmission time, which leads to more power consumption.
2. To develop a noise-aware signal combination (NSC) method that enhances the classification accuracy of EEG-epileptic seizure noisy data.
3. To evaluate the effectiveness of such technique of achieving better EEG classification accuracy and compression ratio.

1.5 Scope of the Research

The scope of this research considers the following points:

1. This work is applied on WBASNs EEG-Based Epileptic Seizure detection. In WBASNs, the approaches of sensing and preprocessing as well as detection and classification have been investigated at the server side. For EEG-Based Epileptic Seizure framework, more emphasize is given to EEG raw data, CS and DCT for compression.
2. EEG data has large data size which requires high bandwidth to send to the server side. WBASNs' EEG will be sent over a channel that may affect the received data's accuracy. On the receiver side, the objective of this work is to retrieve the data with the minimum error or highest accuracy. Features of the received data are extracted in order to estimate the attributes required in the process of measuring the data accuracy. The main advantages of these features are their low computational complexity and computation time. They are used to reduce the dimensions of the cross-correlation sequences and as inputs into individual classifiers.
3. Different EEG classes represent different subjects. The objective is to find the accuracy percentage of each class. Therefore, the estimated attributes are inputted into accuracy classifiers in order to calculate these percentages. In order to improve the EEG data classification accuracy, combining different classifiers can be used.

1.6 Significance of Research

The significance of this research work is reducing the size of EEG data during the transmission by sending less size of data by applying a data compression technique. Moreover, the proposed NSC technique provides a better estimation accuracy than the classical classifiers, particularly, in the case of the reconstructed and noisy EEG data.

The evaluated results show that the proposed method provides several significant benefits such as simplicity, and the improvement of the overall classification accuracy.

1.7 Contributions of the Research

This research work focuses on the reduction of the size of EEG signals at the transmitter side with shorter time. In order to achieve this goal, our research work introduces the following main contributions:

- A unified WBASN framework for compression and classification of noisy EEG-Based epileptic seizure has been developed. This framework utilizes a combination of signal-processing techniques to reduce the size of the data transmission. Those techniques include compressive sensing (CS), discrete cosine transform (DCT) with random matrix, inverse DCT, additive white Gaussian noise (AWGN), and discrete wavelet transform (DWT).
- Based on the proposed framework, the effect of compression parameters on the classification accuracy for different classifiers and feature extraction methods have been investigated and analyzed.
- Develop an ensemble technique for compressed and noisy EEG-based epileptic seizure that improves classification accuracy by present Noise-aware Signal Combination (NSC) technique.
- Evaluate the efficiency of the method to achieve better EEG classification accuracy for imperfect data.

1.8 Thesis Outline

The thesis is organized into five chapters starting by introducing the literature review related to this work and ending by presenting the research conclusions and future work. The chapters are defined as the following:

Chapter 2 presents the literature review of the research work focused on introducing the concept of WBASN, sensing and preprocessing. The fundamentals of EEG and epileptic seizures, EEG signal classification techniques and its background knowledge are discussed. Finally, this chapter focuses on the concept of signal classification and ensemble, including its structure and methods.

Chapter 3 presents the research methodology adopted in carrying out the experiments. It’s start with the identification of the main research problems related to the compression and classification techniques. Then it defines the selected compression technique with the discrete cosine transform, the inverse of DCT as well as discrete wavelet transform methods. Finally, this chapter identifies the selected four legacy classifiers for EEG classification accuracy and their output will feed our proposed ensemble technique.

Chapter 4 explains in details the proposed compression and classification framework and the experiments. In addition, the chapter introduces the proposed technique of noise-aware signal combination method for the classification purpose. The aim of the proposed technique is to enhance the classification accuracy.

Chapter 5 measures and analyzes the compression and classification framework results. The chapter also explores the results between the desired compression ratio and accuracy in the case of noiseless and noisy EEG data.

Lastly, Chapter 6 concludes the findings of the work on the issues addressed by this research. In addition, this chapter proposes outlines of future research directions.

CHAPTER TWO

LITERATURE REVIEW

In this chapter, the state-of-the-art on Wireless Body Area Sensor Networks (WBASNs) is reviewed, discussing the major components of research in this area such as physiological sensing and preprocessing, which includes sampling, filtering, and compression. The worldwide market in 2010 was worth about $10 billion for home health monitoring. According to a report from Berg Insight, an analyst firm estimates that the market is growing about 10 percent annually [20]. Berg also mentioned that the market for home health monitoring of chronic disease (CD) was worth about $11 billion in 2008. Berg estimates that some 300 million people in the European Union and the United States have at least one or more chronic disease that may benefit from home health monitoring. Eventually, Berg listed cardiovascular irregularities, respiratory problems and diabetes as examples of conditions where home monitoring can become a treatment option [20].

2.1 Background

Wireless body area sensor networks (WBASNs) are still in the early development, faced with several technology challenges, of which low power consumption is of top priority. Saving energy is a very critical issue in WBASN because batteries typically power sensor nodes with a limited capacity. Since the radio is the main cause of power consumption in a sensor node, transmission of data size should be limited as much as possible. There are four basic components in a sensor network [21]:

- An assembly of distributed or localized sensors;
- An interconnecting wireless-based network ;

- A central point of information clustering;
- A set of computing resources at the central point e.g. PC/PDA or beyond in order to handle data correlation.

Power consumption is particularly critical for wireless sensor networks operating on limited power reserves, such as batteries or solar cells. Power consumption must be controlled due to the limited power supply for a sensor node with small capacity. Therefore, the reduction of data rate will affect the reduction of power consumed for wireless communication [22]. Hence, this research work will pursue a holistic approach to finally reduce the size of data by sending short stream of data, based on leveraging compressed sensing (CS) for data acquisition and detection and classification techniques.

The healthcare industry is one of the world's largest and fastest-growing industries. Consuming over 10 percent of gross domestic product of most developed countries, healthcare can form an enormous part of a country's economy [23]. Several factors lead to the increasing demand for revolutionary solutions in the healthcare industry, including:

- Increasing number of CD patients; currently more than 860 million [24], the World Health Organization (WHO) claims. It is estimated that global health expenditure on CDs exceeds 80% of whole health, national funds in the US and Europe [24].
- Increasing percentage of average of death caused by CDs, e.g. 87% in countries with high income per individual [25, 26].
- The percentage of elderly people over 60 is on the rise [25].

Due to these factors, traditional healthcare cannot provide the scalability required to cope with the growing number of elderly and CD patients, as it requires a physical one-to-one relationship between the caregiver and the patient [25]. Therefore, the need for high performance, cost-effective healthcare solutions is one of the critical strengths for any developing country seeking sustainable future advancements [13].

Mobile and wireless devices are growing rapidly and are estimated at 5 billion devices worldwide. A critical issue of using such devices is the energy that can be consumed by these devices. These devices operate on limited power reserves, as they are battery-operated, while sensing physical measures. Energy consumption is a major challenge in limiting the mobile device's form factor by reducing the size of data that will be sent over the wireless channels.

Remote monitoring using WBASN has recently emerged to provide real-time patient surveillance and to provide CD patients with more autonomy. The conditions most commonly treated by these remote monitoring services include diabetes, cardiac arrhythmia, sleep apnea, asthma, chronic obstructive pulmonary disease (COPD) and Epileptic Seizure through the EEG signal detection technique [20].

In recent years, the interest in the application of Wireless Body Area Sensor Network(s) (WBASNs) for patient monitoring applications has grown significantly [27]. The WBASN is a wireless network used for communication among sensor nodes operating on or inside the human body in order to monitor vital body parameters and movements. The WBASNs based on low cost wireless sensor network technologies could greatly benefit patient monitoring systems in hospitals, residential and work environments. Such a system allows easy inter-networking with other devices and networks, thus offering health care workers easy access to the patient's critical as well as non-critical data. The WBASN based monitoring system can be used to monitor athletes' performance to assist them in their training activities. This system type could be seen as a special purpose wireless sensor network with a number of system specific design requirements [28]. A WBASN is generally incorporate wearable and implantable nodes operating in two different frequencies. An implantable node is most likely to operate at 400 MHz using the Medical Implantable Communication Service band, whereas the wearable node could operate in an Instrumentation Scientific Medical/Ultra Wide Band (ISM or UWB) bands or any other specific bands [29].

WBASNs also enable constant monitoring of the health conditions of people with CDs. It consists of multiple on-body and ambient sensor nodes, capable of sampling, processing, and communicating one or more physiological signs (PSs) (such as Heart activity, oxygen saturation, movements, blood pressure and Brain activity) over an extended period. Such physiological signs are measured using different types of sensed signals such as the electrocardiogram (ECG) [8], acceleration, and EEG [9]. In addition, it is used for communication among sensor nodes operating on, or inside the human body in order to monitor vital body parameters and movements as well as to enable its user with quality of life, assisted living, sports, or entertainment purposes [10].

Brain status information is captured by physiological electroencephalogram (EEG) signals, which are extensively used for the study of different brain activities. They provide particularly important information pertaining to the epileptic seizure disease [30, 31, and 32]. Epilepsy is a neurological condition, which disturbs the nervous system due to brain damage. Also known as a seizure disorder, epilepsy can be diagnosed after one seizure if a person has a condition that places him at high risk of having another seizure attack. It is reported in [33] that this disease affects 1% of the world's population. However, visual inspection of EEG signals can be very difficult and time consuming. That is due to the difficulty of keeping a high level of concentration during a lengthy inspection, resulting in an increase in the false positives by the operator [34].

Compressive sensing (CS) as a compression technique with the discrete cosine transform (DCT) and its inverse (iDCT) for data reconstructed to the original size are utilized. Discrete wavelet transforms (DWT) are used to get the required feature extraction that will be needed for the decision making. Feature extraction, detection and classification of human related phenomena as well as ensemble classification method are utilized. A comprehensive study and comparisons of sensor technologies used in terms of applications, wireless radio technologies, and different detection and classification techniques is required to realize the end-to-end WBASNs framework for EEG-based epileptic seizure classification. The research framework consists of three parts: WBASNs to capture medical phenomena as raw data, data preprocessing such as compression technique, and communication medium as well as detection and classification at the doctor's side.

2.2 WBASNs General Framework

Recent improvements in signal processing and wireless communications have motivated great interest in application development of wireless technology in healthcare and biomedical research, including WBASNs. WBASNs General Framework (WGF) consists of three major components for real-time applications, namely sensing and preprocessing (SAP), an application-specific WBASN communication (AWC), and detection and classification (DAC) to the patient. SAP contains a number of sensors for capturing a raw data related to medical phenomena, including blood pressure, respiratory rate, ECG, and EEG. Analysis of raw data, including, detection and classification of medical anomalies will occur in the DAC component, providing strict and accurate criteria for the physician to make recommendations that may sometimes feed-back to the patient to provide proactive treatment. Figure 2.1 shows a WBASNs General Framework (WGF) in the real life.

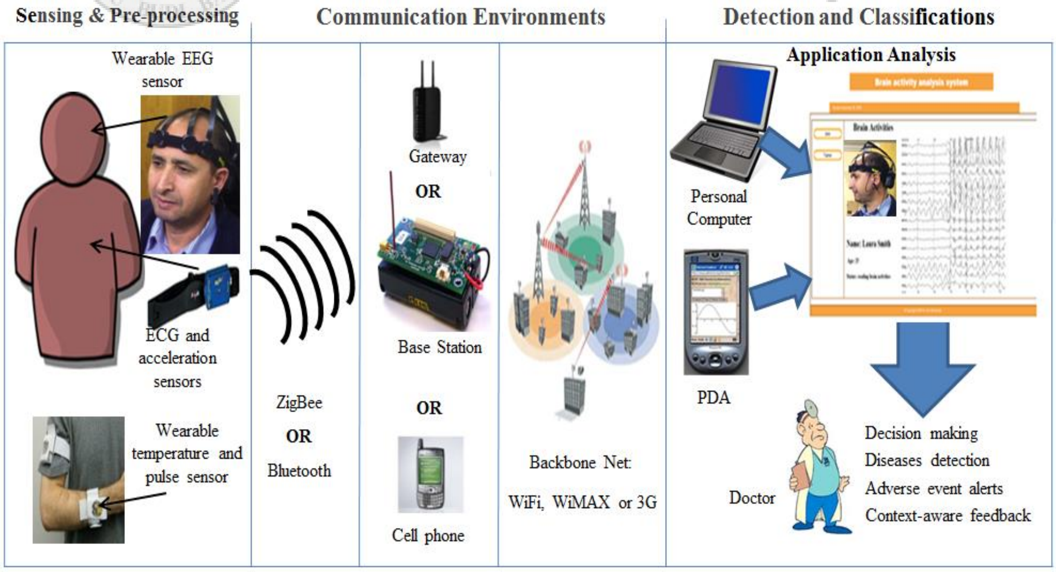


Figure 2.1. General WBASNs Framework [11]

Figure 2.2 shows the conceptual model that illustrates our WBASNs general framework in more details. In sensing and preprocessing stage, data are gathered and processed. This includes sampling, filtering, or compression. Next stage shows the communication environment, which transfers the processed data to server side. In the last stage, features are extracted from receiving data, which are used to detect and classify the seizure type.

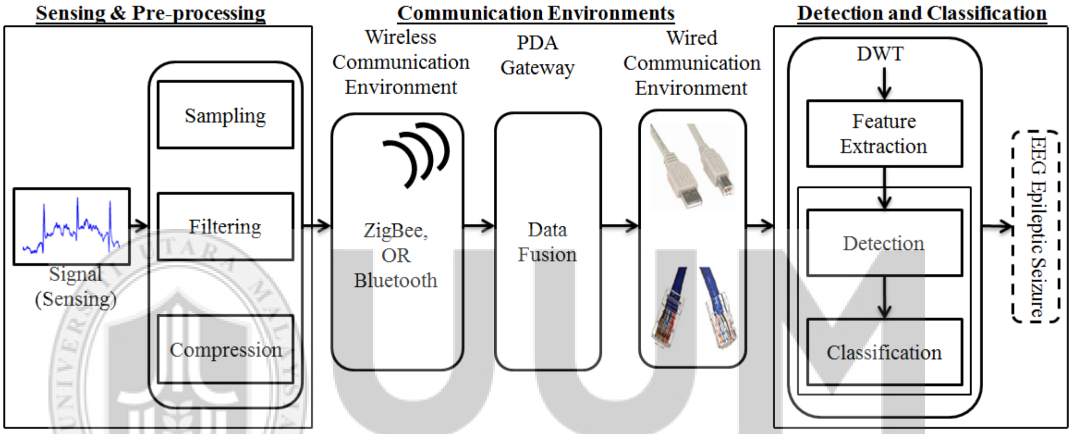


Figure 2.2. WGF Conceptual Environment

2.3 Sensing and Preprocessing

Sensor platform architecture typically consists of a sensing device, an operating system (OS), and a radio frequency environment and power management elements. Sensing is the detection of a physical presence of data and the transformation to a signal that can be read by an observer or instrument [35]. A well-designed WBASN provides doctors with precise real-time and historical information. It is important that the patient should be comfortable and accept the system technologies if they are used in daily life.

WBASNs can be categorized as single-sensor or multi-sensor systems. Single-sensor systems use a single unit placed on the human body with one channel of communication allocated to it. Multi-sensor systems use multiple units, either of the same or different types of sensors, on the patient. The readings from the sensors are usually processed together to extract more accurate readings. In this case, each sensor can have its own wireless channel or they can be combined all and synchronized into one channel [36].

The following types of sensors in WBASN, which are used in respiration application, may be single sensor or integrated multi-sensor platform. The sensors can be originated from [37]:

- ECG/EEG/Electromyography (EMG) sensors are used for monitoring heart activities, brain activity, and skeletal muscles.
- Pulse Oximetry monitors the quantity of oxygen that is being "carried" in a patient's blood.
- Blood glucose level sensor measures the patient’s glucose level in blood.
- Body temperature sensor is used to measure the temperature.

To better observe a human’s vital signals, a wide range of commercial sensor technologies are used to capture physical data such as, accelerometer, EMG, pulse oximetry, respiration rate, heart rate, blood pressure, blood sugar, temperature sensors, ECG electrodes, and EEG electrodes, will be deployed. ECG and EEG electrodes are manufactured in several types, including disposable electrodes, reusable electrodes, headbands or caps, and needle electrodes.

For example, an ECG electrode is a device attached to the skin on certain parts of the patient’s body, such as arms, legs, and chest to detect electrical impulses produced each time the heart beats. Electrode position for a 12-lead ECG is standard, with leads placed on both the left and right arm and leg. Another example is the EMG, which typically uses four electrodes to measure muscle tension as well as to monitor for an excessive size of leg movements during sleep [38].

In contrast, EEG Neuro-feedback, is a type of bio-feedback that uses real-time displays EEG to show brain activity, placing electrodes to gather EEG on approximately 20 different areas of the scalp [39]. The EEG electrodes are placed according to the international 10-20 system and EEG generally uses six "exploring" electrodes and two "reference" electrodes, except if a seizure disorder is suspicious, and in which event more electrodes will be applied to document the presence of seizure activity. The implantation of electrodes is accomplished to exactly localize the seizure generating area, which is described as the epileptogenic zone [40].

The following, Table 2.1, shows sensors employed in WBASNs systems and their typical data rate [41, 42].

Table 2.1:

Common Description of Sensors in WBANs

Sensor Type	**Signal Type**	**Measured data description**	**Compression needed**	**Sampling Rate**
Accelerometer	Body Move	Measure the three dimensional acceleration	Required for 3D	100 KHz
ECG/EEG/ EMG	Skin/ scalp Electrodes	Measure Electrical activity of the heart, brain activity, and skeletal muscles respectively	Recommend	250/250 2khz
Pulse Oximetry	Oxygen saturation	Measure the quantity of oxygen that is being "carried" in a patient's blood	Not required	<1 Hz
Heart rate	Pulse oximeter/	Frequency the cardiac cycle	Not required	60 kHz
Blood pressure	Arm cuff-based monitor	Used to measure the systolic and diastolic pressures.	Not required	<1 Hz
Blood glucose	Strip-base glucose meters	Measures the quantity of glucose in the blood. (type/source of sugar/energy)	Not required	<1 Hz
Temperature probe	Body and/or skin temperature	A measure of the body's ability to generate and get free of heat	Not required	< 1 Hz

The wireless sensor node OS plays a fundamental role in the overall capabilities and performance of the platform. Early research into the OS for sensor networks lead to the development of TinyOS by researchers at UC Berkeley [43]. TinyOS is used in the processing of signals captured. A transceiver communication unit allows the transmission and reception of data to other devices that connecting a wireless sensor node to a network. Power management is provided by the operating system to enforce an optimal way of utilizing energy. Conserving power involves accessing/controlling components of the sensor node. The components that expose power management interfaces are the processor, battery and radio. The components that can be controlled to conserve power are processor and radio. An increasing number of electrodes affect the mobility and convenience of the human subject. The more electrodes are used, the more accurate the data will be. However, this will affect the mobility of the user negatively.

Preprocessing manifested by the procedures performed on the raw data to be ready for analyzes and processing of the application. Preprocessing technique transforms the data into a format that will be more easily and processed effectively for the user purpose [44]. Kotsiantis *et al.* in [45] address issues of data pre-processing that can have a significant impact on the performance of the data analysis, including feature extraction, and classification. Data preprocessing includes sampling, data filtering, and compression. They present a well-known algorithm for each step of data preprocessing in case if there is much irrelevant and redundant information present or noisy and unreliable data. Xu in [46] presented a model of the data preprocessing to reduce the energy consumption attribute of communication between the nodes and enhance the effectiveness of data transmission of wireless sensor networks by means of utilizing independent and intelligent multi-agents. In addition, the model presents the algorithm to accomplish data preprocessing and avoids the error of the data collection.

Other research, Ahmad *et al.* in [47] present a software application that resides in the personal digital assistant (PDA) carried by each patient. This application periodically performs an initial assessment of the patient’s condition based on the available data to work reducing the level of data. If the parameter were within the normal range of health, the input data for that period will be decreased to a few representative values only. Otherwise, all the data will be transferred over the wireless network to the central database for further analysis. However, before transmitting the digitized EEG-based and epileptic seizure data, and as part of preprocessing inside the sensor, three major procedures are performed, namely data filtering, data sampling, and data compression.

One of the important healthcare areas is concerned with the epileptic seizure, which is a noninfectious disorder disease of the brain that affects people of all ages. Seizures are a symptom of something going on in the brain. Seizures seen in epilepsy are caused by disturbances in the electrical activity of the brain. Epileptic seizure refers to one of the most common brain disorders. Epilepsy is a neurological condition, which affects the nervous system. It is a medical disorder that produces seizures affecting a variety of brain and physical functions, also known as a seizure disorder.

2.3.1 EEG-Based Epileptic Seizure

The EEG is the most popular technique used to study brain functions and to diagnose neurological disorders by physicians and scientists, Adeli *et al.* [48]. According to the World Health Organization (WHO) Fact-sheet No. 999, October 2012, [49] around 50 million people over the world have epilepsy. EEG is potential measurement of that reflect the electrical activity of the human brain. EEG is an easy test available that provides a proof of how the brain functions over time. Brain status information is capture by physiological EEG signals, extensively used for the study of different brain activities. One of these states is the epileptic seizure detection. Epilepsy or epileptic seizure refers to the one of the most common brain disorders usually caused by brain injury. Approximately, one in every 100 persons is expected to experience a seizure disorder in their lifetime [33]. The diagnosis of epilepsy is clinical, however, the scalp EEG is the most widely accepted test for the diagnosis of epilepsy [50, 51].

The EEG is the representation of the electrical activity happening on the surface of the brain. This activity will show on the EEG machine screen as waveforms of altering frequency and amplitude computed in micro-voltages. Generally the EEG waveforms are classified by the frequency, amplitude, and shape in addition to the locations on the scalp that are recorded. The most well-known classification techniques are used to classify the EEG waveform frequency, such as alpha, beta, theta, and delta [52]. Normally, EEG waveforms are defined and depicted using their frequency, amplitude, and location [19]. A key characteristic used to describe normal or abnormal EEG rhythms is frequency (Hertz-Hz). For the EEG of a wakeful adult, waves of 8 Hz and higher frequencies are normal, while waves with seven and less are classified as abnormal. In certain situations, EEG waveforms of suitable frequency for age and state of wakefulness are considered abnormal because they occur at an unsuitable scalp location or determine inequalities in rhythmicity or amplitude [53].

Generally, the challenge with WBASNs is the power consumption, as they should, work continuously in order to keep the monitoring patient’s status. One of the ways to ultimately save power is to reduce the size of EEG data sent over wireless channels by using data compression techniques. Due to compression, and during reconstruction at the receiver side, some important information may get lost. Additionally, transmitted signals might be influenced by distortion, interference, and noise. In order to minimize EEG data size, several techniques have been utilized: sampling, filtering, and compression.

2.3.2 Sampling

Sampling is a principle that engineers follow in the digitization of analog signals. For analog-to-digital conversion that leads to the accurate reproduction of the signal, samples of the analog signal must take repeatedly and the number of samples per second is called the sampling rate or sampling frequency. Sampling theory, known as the Nyquist theorem [54], provides a prescription for the minimal sampling interval required to avoid aliasing. It states that a signal must be sampled at least twice its highest analog frequency in order to extract all of the information from the bandwidth and accurately represent the original acoustic energy. This can be represented mathematically in [54] as:

$$f_s \geq 2f_c \tag{2.1}$$

where *fs* is the sampling frequency (samples are taken per unit of time or space), and *fc* is the highest frequency contained in the signal. Hence, a simple way to avoid aliasing is always have enough samples to capture the spatial or temporal variations in the signal [54].

Sampling techniques in the field of sensor networks have been extensively studied in the literature. For instance, [55] proposed a novel adaptive sampling technique based on Kalman Filter estimation error to adaptive adjust the sensors' sampling rate within a given range; where the sampling rate at each sensor adapts to the streaming-data characteristics. This adaptive sampling was found to be desirable not only to preserve resources, but also to improve the overall quality of results. To minimize Kalman Filter estimation error over all active streaming sensors, the server allocates new sampling rates under the constraint of available resources. The system results show performance upgrade 3 to 4 times compared to the uniform sampling when input parameters are chosen carefully.

An adaptive sampling approach (ASAp) to energy-efficient periodic data collection in sensor networks also has been developed in [56]. In this work, a dynamically changing subset of the nodes used as samplers such that the sensor readings of the sampler nodes are directly collected. To predict the values of the non-sampler nodes, probabilistic models that locally and periodically constructed are used. ASAp approach mechanisms are used to minimize the number of messages used to extract data from the network and increase the network lifetime while keeping the quality of the collected data high in the scenario. This scenario, like a certain size of data quality, can be traded off in order to decrease the power consumption of the network. The simulation-based experimental results and study present the effectiveness of ASAp under different system settings.

Nodes in wireless sensor networks often suffer from failure, limited storage capacity, computing ability, and battery power. Hence, focusing on inaccuracy data and power limitation, [22] proposed a sampling frequency control algorithm and a data compression algorithm. These algorithms are combined based on features of sensed data. The sampling frequency control algorithm adjusts the sampling frequency of the sensed data dynamically while the data compression algorithm is adopted to reduce the size of transmitting data to save energy of sensors when the signal frequency cannot be controlled. The results of experiments and analysis show that the proposed algorithms can decrease sampling times, reduce the size of transmitting data, save energy of sensors, and improve the query efficiency.

In nature, usually WSN communications are multi-hop and multipath; data from a source to a sink node is relayed and overheard using several intermediate nodes. Due to this multi-hop and multipath communication advantage, each user query may leave a trace of collecting data among the communication nodes, which can then be used to answer future global aggregation requirements at node level without further communication. This mode of data communication by proposing an opportunistic sampling method for accumulating global knowledge at the node level can be seen in [57]. This global knowledge can greatly improve numerous WSN applications when used in data validation, event detection, and query optimization.

2.3.3 Data Filtering

Filtering refers to the process of defining, detecting and correcting errors in sensed data, in order to minimize the impact of errors generated by external noisy data. Filter scans the data for multiples/clones/duplicates of records and inconsistent data to be excluded. Whenever unnecessary data occurs, one representative record is chosen or constructed. The rest are deleted and data is reduced in size as a result. The output will finally be refined data, but still potentially contains errors. An effective energy-saving filtering mechanism is proposed in [58] to enhance the energy-efficiency of data gathering. This proposed filtering framework mainly puts emphasis on reducing the production of redundant loads at gathering stage to greatly reduced energy cost using a self-adaptive filtering scheme. This framework aims to achieve energy saving by reducing the redundant communication loads in networks, which is the trade-off between data precision and energy-efficiency. Compared to classical data gathering approaches, framework in [58] performs better regarding energy-saving effect.

To reduce the sensor data volume that arises from the use of continually transmitting sensors (e.g., EEG, ECG or EMG), [59] proposed a context-aware filtering technique in which the relaying mobile device dynamically modifies its processing logic based on changes in the user's context. They implemented the Healthcare-oriented Adaptive Remote Monitoring (HARMONI) middleware in order to evaluate this technique, on a mobile device, and used it to collect real sensor data from users. HARMONI includes a lightweight event engine that runs on the mobile device, and processes incoming sensor data streams using rules that are appropriate for the current context. Their experiments demonstrate that context-aware filtering can reduce the uplink bandwidth requirements of the system by up to 72%. Eventually, given brief information about the filter technique, which is one of the sensing preprocessing in the above conceptual model, however, the filter technique is out of this research scope.

2.3.4 Data Compression

In contrast to filtering, which potentially focuses on reducing errors and unnecessary data (such as duplicate data), compression focuses on reducing the size of physical data traffic that the sensor sends over the wireless channel, to improve bandwidth utilization, power consumption, and possibly speed of processing and memory space required by the application. Data compression can be categorized into two methods: lossless, and lossy. The lossless method promotes the reconstruction of the original signal after compression with no loss of any type. While preserving all the signal characteristics, this method may require excessive bandwidth for communication [60]. In addition, depending on the application, signals may not be required in full to be able to detect the patient's anomalies.

In contrast, the lossy method will not get the original signal accurately after compression but may be more bandwidth efficient, while minimizing the effect on the application process. A lossy method used two major criteria: The CR that is representing the ratio between the original signal and compressed signal; the percentage root-mean-square difference (PRD) which is defined as the error criterion in estimating signal rebuilt for lossy compression [60]. The error criterion for lossy compression techniques to estimate the distortion of the signal rebuilt with respect to the original one is very important, especially for an EEG signal, where a slight loss or change of information can lead to wrong diagnostics. The controlled transmission quality measure PRD for EEG compression is described in Equation 2.2 as [60, 61]:

$$PRD = \sqrt{\frac{\sum_{i=1}^{N}(x(i) - \hat{x}(i))^2}{\sum_{i=1}^{N}(x(i) - \mu)^2} \times 100\ \%} \tag{2.2}$$

where $x(i)$ and $\hat{x}$ (i) are the i^{th} samples of original and reconstructed EEG signals of length N respectively, and μ is the signal mean value.

In [62], designed and implemented two lossless data compression algorithms, namely entropy encoded codebook compression and pipelined codebook compression. Both of them are built over the codebook compression techniques. After deployment, they assumed that each sensor node is stationary and capable of getting its location information using GPS using a camera sensor network. The benefit of this technique is to reduce the size of memory occupied by the compression and short processing time needed. This leads to a reduction of the overall delay of the data packet transmission. Other research in [63] proposed quad level vector for ECG signal processing to achieve a better performance for both compression flow and classification flow, considering low-computational complexity. The classification algorithm has been employed for methods of the heartbeat segmentation and the R-peak detection. The results show that with the proposed compression techniques the overall power consumption is reduced by 45.3%.

In [64] presented a lossless data compression algorithm to perform compression using two code options. The data sequence to be compressed is divided into blocks; the optimal compression scheme has been applied to each block. The results achieved compression performance up to 74.02%.

In fact, data compression reduces the power consumed during communication by transmitting a shorter data stream due to compression. Transmitting single bits of real data is equivalent to executing 300 lines of code on a typical processor [29]. One of the famous data compression techniques is compressive sensing (CS), it is considered an alternative sampling theory, which is, employed to compress the EEG data with different values using discrete cosine transform (DCT) technique. Thus, to reconstruct the compressed data back at the receiver side, the inverse DCT (iDCT) technique was utilized; the following sub-sections will give more details about the CS, DCT and DWT.

2.3.5 Compressive Sensing

Compressive sensing (CS) is a new approach of reconstructing a sparse signal much below the significant Nyquist rate of sampling [65]. Thus, CS has been considered for efficient EEG acquisition and compression in several application contexts since it is acquire measurements fast [66, 67]. CS shows that certain signals can be recovered from far fewer samples than Nyquist sampling uses [68]. The term sparse indicates that most of the signal sets are of zero values. However, spikes in the signal with indeterminate distances represent nonzero values. So sparse means that magnitude arranged transform coefficients decay quickly, at that the signal is compressed [61, 69].

As the compression process required in the wearable devices, CS has a low computational complexity that is potentially suitable for use in wearable computing systems as the compression process. The research works in [66, 67] have focused on the sparse modeling of EEG signals and evaluating the efficiency of CS-based compression in terms of signal reconstruction errors. The work in [61] has tried to estimate the low-power potential of CS for portable EEG systems using datasheet-extracted power consumption figures for the various components. It also estimates the required size of processing and wireless transmission.

The research works of [70, 71, 72, and 73] demonstrated that a finite-dimensional signal having a sparse or compressible representation could be reconstructed from a small set of linear, non-adaptive measurements.

As the sensor is battery operated, the power in wireless body sensor networks (WBASNs) is limited. In [74], proposed a compressed sensing framework to efficiently compressed EEG signals in WBASNs to minimize the size of data transmitted by the sensor. For energy efficiency, they provide a non-adaptive compression scheme to compress EEG signals in WBASNs at the sensor node. On the other hand, a CS-based reconstruction process has been implemented. This achieves a better quality and more energy efficient as compared to the energy-hungry JPEG2000 [74] compression framework in terms of computation and wireless communication.

In the tele-medicine applications, the potential of CS in EEG signal compression was found in [75, 76]. The research in [75] provides a practical performance of the compressive sensing in several implementations when applied to scalp EEG signals. They focused on reviewing existing data sparsifying dictionaries and data reconstruction algorithms. This has been done for 18 different implementations of the CS theory. Consequently, authors in [75] introduced performance results from testing different groupings of these elements to determine which one produced the best results. The results show the applicability of single-channel CS for EEG signals based on the proposed application and acceptable reconstruction error.

After two years, the same authors in other work [76], focused on the reconstruction algorithms, surveying existing sparsifying dictionaries, in order to determine the best results, they tested a different combination of these elements. The results show that the limited applicability as a compression technique when applied to single-channel CS for EEG signals, authors noticed that this mostly depends on the requirements of the application and more precisely on the reconstruction error that is acceptable.

In the last decade, the Discrete Cosine Transform (DCT) has appeared as the effective data transform in most visual systems due to its low complexity at the transmitter side, which is very important. DCT is useful since it has been extensively used in image and audio compression by modern video coding standards. DCT has a possibility of fixed basis data and fast implementations. It involves the use of just cosine functions and real coefficients as well as simpler to calculate. DCT is used as the basis to sparsity the EEG signal as part of the CS framework. It is similar to a Fourier-related transform like the discrete Fourier transform making use of both Sins and Cosines and requires the use of complex numbers; however, DCT uses only real numbers and low computational complexity at the transmitter side [19]. More detail discussion about discrete cosine transform is given in Chapter 4.

The following section, feature extraction, detection and classification will be discussed, which are on the server side.

2.4 Detection and Classification

On the other side, in order to analyze EEG data that should be decompressed, reconstructed and then extract its features, and finally, the EEG-based epileptic seizure classified.

2.4.1 Signal Reconstruction

There are many reconstruction techniques; the well-known traditional approach of reconstructing signals or images from compressed data is a Shannon Nyquest theorem [42], which indicates, “The sampling rate must be twice the highest frequency”.

Compressive sensing (CS), compressive sampling or sparse recovery provides a new fundamental approach to compress data [70, 73]. CS is a promising technique that performs compression below Shannon Nyquest theorem, while obtaining the desired reconstruction accuracy. It predicts that a specific signal or images can be recovered from what was previously believed to be highly incomplete information. Certainly, in order to compress signal x_0 simply one may store the largest *k* entries. The non-stored entries are set to zero for reconstructing x_r from its compressed version, and reconstruction error. At this point, it is highlighted that the procedure of achieving the compressed version of x_0 is adaptive and nonlinear since it needs the search of the largest entries of *x* in absolute value. Especially, the location of non-zero is a nonlinear type of information.

On the receiver side, a reconstruction/decompression executes the reverse operations to those made by compression to get the original source data. The challenge in this process is to retrieve original data, which might be noised or lost.

2.4.2 Feature Extraction

EEG Feature Extraction plays a significant role in diagnosing most of the brain diseases. Obtaining useful and discriminant features depends largely on the feature extraction method.

The Discrete Wavelet Transform (DWT) method is widely used [51] due to the nature of EEG signals, which are time varying and space varying. It captures both frequency and time location information [77, 78, 79 and 80]. The key feature of wavelets is the time-frequency localization. It means that most of the energy of the wavelet is restricted to a finite time interval [32]. Using multi-resolution wavelet analysis, basically, DWT decomposes the EEG signals into different frequency bands. The Discrete Wavelet Transform (DWT) is an implementation of the wavelet transform using a discrete set of wavelet measures and translations follow some defined rules. In other words, this transform decomposes the signal into different orthogonal set of wavelets, which is the main difference from the continuous wavelet transforming, or its implementation for the discrete time series [79, 80].

The reconstruction of the signal is concerned, in many cases as far as the information was highly redundant. This redundancy requires a significant volume of computational time and resources. The DWT provides sufficient information for both analysis and assembly of the original signal, with a significant reduction in the computation of time. Eventually, DWT is the most popular technique, because the EEG data is time-frequency domain, this method is capable of capturing valuable time and frequency information simultaneously [81].

Since EEG signals are time-varying and space-varying non-stationary signals, the DWT method is widely used [51, 77, and 78]. The research work in [82] proposed a method for seizure detection in Intracranial EEG by using lacunarity with Bayesian linear discriminant analysis (BLDA) classifier. They used DWT as an analysis tool for EEG feature extraction. Lacunarity is a measure of homogeneity for a fractal. The wavelet transform with the statistical features of lacunarity extracted from the sub-band are decomposed with EEG signals. These features are inputs for BLDA classifier for training and classification. In addition, they used post-processing operations on the BLDA outputs to achieve accurate EEG classification and reduce false positive. This post-processing includes smoothing, threshold judgment, multichannel integration, and collar technique. The experimental results on Freiburg EEG datasets show that the proposed method can achieve a better sensitivity and low false detection rate [82]. Based on the features used, different techniques can be generally categorized into time-domain or frequency-domain-based techniques.

Time-domain features are easily computed and usually their time complexity is manageable. The authors in [83] proposed Time Domain Parameter (TDP) based feature extraction. It is a generalized form of the Hjorth parameter and can be computed efficiently. The TDP feature is then fed to a Linear Discriminant Analysis classifier that is utilized in a Brain Computer Interface application. Five time-domain features; namely summation, average, standard deviation, zero crossing and energy are proposed in [84]. Subsequently, they are used by a set of classifiers for the purpose of epileptic seizure detection. The output of the classifiers was then combined, using the Dempster's rule of combination, for a final system decision. 89.5% of classification accuracy was achieved.

Mirowski *et al.* in [85] evaluated out-of-sample seizure prediction 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.

SVM logistic regression or convolutional neural networks have been used as machine learning-based classifiers to discriminate interictal from preictal patterns of features. Results show that the proposed technique is outperformed by the previous seizure prediction methods on the Freiburg dataset. However, it only targets two datasets that discriminate two interictal from preictal seizure; also, they did not address the compressed and noisy EEG data. Other work in [51] has shown some preliminary investigations on Best Basis-based wavelet packet entropy as a feature extraction of EEG signals for epileptic seizure detection. This work did not consider the time for feature extraction, which represents a significant factor of EEG-based detection of epileptic seizure, the fast feature extraction is important in a real-time setting. However, fast feature extraction might affect identification accuracy.

Other research work in [86] presents an overview of appropriate signal processing techniques requested to investigate sleep EEG signals in both pediatric and adult populations. Three key stages needed for the investigation of sleep EEG namely, preprocessing, feature extraction, and feature classification. Preprocessing describes the signal processing technique that will deal with the preparation of sleep EEG prior to further analysis. Feature extraction and classification focus on most commonly used signal investigation methods used for characterizing and classifying sleep EEGs.

Frequency-domain features are usually obtained by transforming the EEG signals into their basic frequency components. The characteristics of these components fall primarily within four frequency bands [87]. In [88], a classification system uses a one-second time window to extract relevant features. The fast Fourier transformation (FFT) is used to transform the data within the window into the frequency domains. In order to distinguish between several brain states, frequency components from nine to 28 Hz were studied and presented to a modified version of Khonen’s learning vector quantization classifier.

Other research in [51] has proposed an EEG classification system for epileptic seizure detection. It consists of three main stages, namely, 1) the best basis-based wavelet packet entropy method, which is used to represent EEG signals by wavelet packet coefficients, 2) *K*-NN classifier with the cross validation method in the training stage of Hierarchical Knowledge Base (HKB) construction. Lastly, 3) the top-ranked discriminate rules from the HKB will be used in the testing stage to compute classification accuracy and rejection rate. They reported a classification accuracy of close to 100%, however, their experiments considered only healthy subject which is class A and epileptic seizure active subject which is class E data and never considered seizure-free intervals which are class C or class D. Trivially, neglecting such classes eliminated the main source of difficulty in this classification process.

Other reported techniques utilize a mix of time and frequency domain features, such as in [53]. Using the EEG amplitudes, the first, second, third, and fourth statistical moments (i.e., mean, variance, skewness, and kurtosis) were extracted. Along with these time-domain features, energy and other frequency domain features were extracted. A support vector machine (SVM) was then applied to the obtained features for seizure classification. The following paragraphs provide a brief description of these classifiers. More details are given in chapter 3.

Song and Lio in [89], proposed an approach for epileptic seizure detection based on sample-entropy (SampEn) together with an extreme learning machine (ELM) to get features for EEG classification. This approach is used to classify different subjects as normal subjects, subjects not having an epileptic seizure, or subjects having an epileptic seizure. The results show that the proposed approach gives a better performance with the accuracy and fast learning speed. However, the value of the SampEn falls rapidly during an epileptic seizure.

Guo *et al.* in [90] proposed a new feature extraction method based on line length feature of the sub-band signals and combined this with an artificial neural network (ANN) to detect epileptic EEG signals. This combination has achieved the following success: the traditional updating for the back propagation algorithm of ANN, and manage to decrease into a local minimum. However, the running speed is very slow to fulfill the requirements of clinical and real-time applications.

Another research in [91] conducted performance analysis of EEG patterns using discrete wavelet transform (DWT) and Independent Component Analysis (ICA). DWT & ICA have been utilized 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 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 with the same dataset. However, they tested only two datasets of EEG that easily discriminate between both of them and did not address the noisy EEG data problem.

2.4.3 Classification Techniques

EEG detection and classification plays an essential role in the timely diagnosis and analyzes potentially fatal and chronic diseases proactively in clinical as well as various life settings [11]. Liang *et al.* in [92], proposed a systematic evaluation on EEG by combining both complexity analysis and spectral analysis for epilepsy diagnosis and seizure detection. Around 60% of the features extracted from the dataset were used for training, while the remaining ones were used for testing the performance of the classification procedure of randomly selected EEG signals [92].

The method proposed in [93] uses the features namely, average EEG amplitude and average EEG duration, coefficient of variation, dominant frequency, and average power spectrum as feature inputs to an adaptive structured neural network. The method proposed by Pradhan *et al.* in [94] uses raw EEG signal as input to a learning vector quantization (LVQ) network. The authors in [95] have proposed a new neural network model called LAMSTAR network and two time-domain attributes of EEG, namely, relative spike amplitude and spike rhythmicity. They have been used as inputs for the purpose of epilepsy detection.

The research work in [34], proposed bag-of-words model used for biomedical time series represented as a histogram of the code-words, each entry appeared in the time series such as EEG and ECG signals. In this model, both local and global structural information is well utilized to capture high-level information. DWT was employed to extract a feature vector from each local segment to characterize the local segment. The results of the proposed method are insensitive to the parameters of the bag-of-words model and robust to noise.

Other research in [96] presents numerous classifier systems for classification of EEG signals. In the robust manner, they apply several methods to classify motor imageries originating from the brain. At each level, DWT decomposition was used as feature extraction and reduced the dimension of the feature space by extracting mean, min, max and standard deviation parameters of the brain signal. Also, they utilize several classic classifiers in order to apply them in the multiple classifier system, such as k-NN, multilayer perceptron, Naïve Bayes, linear discrimination analysis (LDA), and SVM, which outperforms the results of the other proposed methods on the dataset. For ensemble, to reduce the classification error, the Bagging, Majority voting, weighted majority voting, and Adaboost well-known combination methods were evaluated and compared. The results refer to that the ensemble method can work to boost EEG classification accuracy.

Since the EEG signals are non-stationary, the traditional method of frequency study is not well effective in diagnostic classification. The research work in [97] provides efficient automatic scheme to support a physician in the detecting process. These schemes facilitate the detecting of epilepsy and improve the administration of long-term EEG recordings. Using wavelet transforms to analyze the EEG signals, and for classification, they were using multilayer perceptron neural network (MLPNN) and logistic regression (LR). In addition, since the early days of automatic analysis of epileptic EEG signals, demonstrations based on the Fourier transform and parametric methods have been applied. These methods are based on previous observations that the epileptic seizures give growth to changes in certain frequency bands, such as δ (0.4–4 Hz), θ (4–8 Hz), α (8–12 Hz), and β (12–30 Hz) bands. In each sub-band, certain conventional statistics over the vectors were used to reduce the interval of these vectors, such as average power, mean, entropy and standard deviation of the wavelet coefficients. The results show that the MLPNN based classifier was more accurate than the LR based classifier.

Two years later, the author in [32] extended his research to present a mixture of experts' (ME) network method in addition to MLPNN to develop classifiers for detecting epileptic EEG signals. DWT coefficients of EEG signals were supplying a modular neural network structure. The results of the proposed model ME network has some potential in the epileptic seizure detection, and achieve a higher accuracy rate than the standalone neural network model. However, these methods are not appropriate for the frequency decomposition of these signals because epileptic EEG signals are non-stationary and multi-component.

A collection of methods based on ANN has been utilized in the epileptic seizure detection and EEG signal classification [98]. This research in [98] categorizes the learning technique for neural networks into two types, supervised learning and unsupervised learning. Supervised learning requires previous knowledge of the analyzed data. The unsupervised learning paradigm has fewer demands for the preceding knowledge of the data. The authors in [98] presented an ANN based classification model in the epilepsy treatment as a diagnostic decision support mechanism. In addition, the authors used the most popular technique DWT to decompose EEG signals into frequency sub-bands to evaluate the proposed classification model by implementing different experiments on different mixes of seizure data. Results show the proposed model in satisfactory classification accuracy rates.

Other work in [99] proposed a seizure detection method based on using the higher-order statistical moments of EEG signals calculated in the empirical mode decomposition (EMD) domain, such as variance, skewness, and kurtosis. The purpose of these moments is differentiating the EEG signals through an extensive analysis in the EMD domain. These moments were used as input features in an ANN classifier, where the results demonstrate a higher classification accuracy, sensitivity, and specificity values. These values were achieved using different approaches based on wavelets, time-frequency analysis, and higher-order statistical analysis in the EMD domain.

The authors in [100] presented Daubechies-4 wavelet transform to decompose the EEG into three sub-signals for automatic seizure detection with high sensitivity. The three sub-signals were within the range of 16 – 32 Hz (d3), 8 – 16 Hz (d4), and 4 – 8 Hz (d5). In this method, they utilized DWT for detection of seizures from the long-term intracranial EEG (iEEG) signals. The proposed sub-band signal method was delivered through a feature extraction block to calculate four main features of relative energy, relative amplitude, and coefficient of variation and fluctuation index from wavelet coefficients. Consequently, these features, construct a feature matrix as input for SVM to classify both seizure and non-seizure activities. Results show a sensitivity of 94.46% and a specificity of 95.26% with a false detection rate of 0.58/h by using this method on a large dataset of 509hrs from 21 epileptic patients in long-term EEG. However, in some previous experiments [100] training data and test data from EEG signals almost are obtained from the same cases, which may affect the reliability of the clinical classifiers.

The identification of the epilepsy from EEG signals is difficult, very time consuming, and costly [101], taking into consideration the large number of epileptic patients registered to hospitals and the huge size of data needed to be recorded. In this regard, Ba-Karait *et al.* in [102] introduced an adaptive particle swarm negative selection (APSNS) method based on particle swarm optimization (PSO) algorithm to automate the detection of epileptic seizure process in EEG signals. This algorithm belongs to random optimization algorithms and is used to find an optimal solution to numerical and qualitative problems. In addition, it used DWT to analyze the EEG signal to extract some features that are required for decision making using the proposed APSNS. These features have been used to investigate the performance of the proposed APSNS algorithm in classifying the EEG signals. The results show that this method outperforms other techniques in the literature in terms of classification accuracy.

The research work in [103] presented a topographic brain mapping with wavelet transform–neural network method for the classification of epilepsies of EEGs signals. They used brain mapping for finding the epilepsy location in the brain. The preprocessing was included to remove blinking artifacts, electrode movement and used DWT for eyeball movement, in order to enhance the speed and accuracy of the processing stage. They categorized EEGs signals to normal, petit mal and clonic epilepsy using an expert neurologist, then confirmed by Fast Fourier Transform analysis. Counties Wavelet Transform of EEG recording are used to extract features that will be the classifier input. Results show that accuracy of the proposed classifier in experimental clinical data was achieved by 80%.

Other research work in [52] proposed a classification technique for seizure and non-seizure of EEG signals using empirical mode decomposition (EMD) method. They used least square SVM (LS-SVM) technique to classify seizure and non-seizure EEG signals. The amplitude and frequency modulation bandwidth of intrinsic mode functions (IMFs) are using a feature extraction to feed the LS-SVM classifier. Another research, Sharma *et al.* in [104] presented a classification method of two focal and non-focal EEG signals. Data from five epilepsy patients who had longstanding drug resistance has been used to test the method. The only base classifier used was the LS-SVM classifier. Average sample entropy and average variance of the intrinsic mode functions (IMFs) was obtained based on EMD of EEG signals. The results show that the proposed method gives a classification accuracy of 85%.

The second-order difference plot method of IMF [105] has been used as a feature for epileptic seizure classification. The computed area from the diagnostic signal demonstrates that the IMF detection is found to be a significant parameter for analysis of both healthy and unhealthy subjects [106]. The mean frequency feature of the IMFs has come up as a feature to identify variance between ictal and seizure-free EEG signals [107]. Wavelet and multi-wavelet transformations have been included in analysis and classification of EEG time-frequency of epileptic seizure [108]. However, these methods used noiseless data, while in this research both noiseless and noisy data were used. Compared with our methods, these datasets are only using the LS-SVM as a base classifier, while in this research four different classifiers were used. Compared to [103], the proposed technique in [52] achieves better accuracy and the result is the best among the LS-SVM classifiers.

Recently, numerous research and techniques have been developed for analyzing the EEG-based epileptic seizure classification. In addition, feature extraction methods have been combined with different individual classifier categories. There are different methods for data classification such as Decision Trees (DT) and fuzzy logic [31,87,109], artificial neural network (ANN) [97,32,93,110], multilayer perceptron neural network (MLPNN) [97,32], Naïve Bayesian (NB), k-nearest neighbor (*k*-NN) [83,19], support vector machines [30,19,109], least square support vector machine (LS-SVM) [52], and other signal analysis techniques. The comparison of the classifiers and using the most predictive classifier is very important. Each of the classification methods shows different efficiency and accuracy based on the kind of datasets [51, 53]. In addition, there are various evaluation metrics for comparing the classification methods that each of them could be useful depending on the kind of the problem.

Extensive experiments have been conducted to test some classifier on the EEG benchmark data to nominate and select the best and most popular of four classifiers. The following sub-sections depicting common techniques have been used for data classification, such as, ANN, NB, k-NN, and SVM classifiers.

2.4.3.1 Artificial Neural Network

Artificial neural networks (ANNs) are mathematical models that are based on structure and functional aspects of biological neural networks. ANNs are highly correlated managing elements, ones or neurons that theoretically emulate the structure and operation of the biological system. Learning in ANNs is carried out via specific developed training algorithms; it is found that the rules of learning are supposed to simulate the learning procedures of biological systems. There are several types and architectures of neural networks mixed fundamentally in the way they learn. ANN is able to estimate the posterior probabilities in order to establish classification rules and perform statistical analysis [85, and 97].

2.4.3.2 Naïve Bayesian

Naive Bayesian (NB) is a type of statistical classifier. It has been demonstrated to be effective in many practical applications, including text classification and performance management systems. NB is a simple probabilistic classifier that applies the Bayes' theorem. It uses a strong (naïve) assumption of the features that depict the classified objects, which are statistically independent of each other. The assumption of the NB method makes the calculation of the NB classifiers more efficient than the exponential complexity.

Despite this strong assumption, NB proved very effective in many real world-applications. The NB approach’s main advantage comes from the learning procedure computational efficiency, which shows a linear computational complexity [84, 111]. Simply, it works by considering that the presence of certain features of a class is irrelevant to any other features. Based on the presence of the other features, the NB classifier considers all features individually contribute in the probability.

2.4.3.3 K-Nearest Neighbor (k-NN)

This method, amongst the simplest algorithm of all machine-learning algorithms, classifies the example by the majority vote of its neighbors. The *k*-NN classifier is based on learning by comparing a given test class with training classes that are similar. This algorithm combines two-steps: the first step, find the *k* training samples that are closest to the invisible sample. Secondly, take the commonly occurring classification for these *k* samples, and find the average of these *k* values in the regression. It can be defined by a distance metric called normalized Euclidean distance between two points as equation (2.3), given two points Y_1= (y_{11}, y_{12},…, y_{1n}) and Y_2= (y_{21}, y_{22},…, y_{2n}) [19, 112].

$$dist(Y_1, Y_2) = \sqrt{\sum_{i=1}^{n} (y_{1i} - y_{2i})^2} \tag{2.3}$$

2.4.3.4 Support Vector Machine

SVM uses nonlinear mapping to transform the original data into a higher dimension to realize a linear optimal separated hyperplane. SVM is a classification algorithm of both linear and nonlinear data. An SVM learner is a strong classifier based on the statistical learning theory. It constructs an ideal hyperplane in order to separate the data into two different classes to minimize the risks (“that is, a “decision boundary”).

It takes a set of input data and predicts for each given input, where two possible classes involves might be the input. It is found that SVM presents a compact description of the learned model that can be used for prediction as well as a highly accurate classification [85, 111]. Moreover, SVM applies decision combination to practice multiclass classification since it is a binary classification. SVM is an integrated software and powerful method for both classification and regression as well as distribution estimation. SVM operator supports types *C-SVC* and *nu-SVC* for classification tasks, *epsilon-SVR* and *nu-SVR* types for regression tasks. Finally, the *one-class* type is for distribution estimation [19].

Classification techniques reported above provide satisfactory performance given that the EEG data are not contaminated by different factors. Although the raw EEG datasets (free of artifacts) were used, the lossy compression will introduce signal distortion, which will affect the reconstructed data. Therefore, wireless EEG data often are compressed before transmission, which means that some important information may get lost during the reconstruction process on the receiver side. Moreover, a wireless channel may augment the transmission problem by adding noise artifacts to the transmitted data. Therefore, a prospective classification technique should take into consideration the uncertainty problem contained in the EEG data to be efficient.

2.4.1 Classification parameters

The overall accuracy of the classifier represents the degree of closeness of measurement results to the true value. There are two factors that affect accuracy classification which are specificity and sensitivity, which are defined as a function of the true and false positives and negatives. First, False Positives (FP) refers to the condition in which the results are perceived as positive, when there is no definite disease or severe illness. On the other hand, True Positives (TP) is a test that shows correct behavior by detecting definite disease or severe illness. Similarly, True Negatives (TN) is defined as the correct behavior to detect the normal patient condition with no severe illness, while False Negatives (FN) are the incorrect detection of normal condition, where the subject suffers a severe disease or illness [113].

Specificity in diagnostic laboratory refers to the ability of an assessor to measure one particular organism or substance [60]. Specificity, also known as a class precision, is a medical term defined as a percentage ratio of true negative tests for the total number of infected patients tested. Moreover, sensitivity also known as a class recall, in diagnostic laboratory testing represents the smallest size of a substance in a sample that can be accurately measured by an assessor. Sensitivity is defined as a percentage ratio of true positive tests for the total number of affected (positive) patients tested. Therefore, specificity, sensitivity, and overall accuracy of the classifier can be defined as follows [11]:

$$Specificity = \frac{TN}{TN + FP} \tag{2.4}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{2.5}$$

$$Accuracy = \frac{TP + TN}{TN + FP + TP + FN} \times 100 \tag{2.6}$$

Where TP, FP, TN, FN are true positive, false positive, true negative, and false negative, respectively. Both positive and negative terms are denoted the classifier’s prediction or expectation, and true and false are referring to whether that prediction corresponds to the external judgment/observation or not. Consequently, these terms have compared the results under the test of the classifier with trusted external judgments.

These equations show that a test with high specificity has few FPs; where as a test with high sensitivity has few FNs.

A classifier performance depends significantly on the data characteristics need to be classified. Several experimental tests have been conducted to find the data characteristics that affect the classifier accuracy. The accuracy measures and confusion matrices are very popular used to evaluate the quality of classification schemes.

2.5 Ensemble Classification Techniques

Fundamentally, ensemble technique plays a significant role in the problem of EEG signal classification. An ensemble classifier is a set of classifiers whose individual decisions combined to classify EEG signals. Ensemble approaches utilize the inputs of multiple techniques instead of a single technique. The ensemble method is able in dealing with small sample size and high dimensionality [114]. Over the past ten years, there is a significant theoretical and experimental research in this area which led to many methods such as bagging and boosting in order to solve many real problems. The ensemble methods idea is to build a predictive classifier approach by combining different classifiers in order to enhance classification prediction accuracy. Combining predictions of an ensemble are almost more accurate than the individual classifiers that make them up. Generally, it is true that combining several classifiers leads to increase in the classification accuracy [115].

Recently, ensemble (combination) classification methods have attracted growing attention from both academia and industry. The research work in [116], evaluated the performance of three popular ensemble bagging, boosting and random subspace ensemble methods. These methods were evaluated based on EEG signal classification. The dataset used was recorded from normal subjects during three kinds of mental imagery tasks such as imagination of repetitive left hand movements, right hand movements and generation of different words starting with the same random space letter. The experiment was conducted on three subjects for underlying applications with K-NN, decision tree, and support vector machine as base classifiers. The work was mainly based on the spatially filtered by means of surface Laplacian. The authors claim that the capability of ensemble methods is subject to the type of base classifiers, particularly the setting and parameters used for each individual classifier. However, they did not find the best parameter configuration and adaptive method for each subject to improve the classification performance.

Generally, it is known that the ensemble classifiers perform better than each individual classifier of which they contain. There is an extensive set of several synthesis rules that could be used; the simplest one is taking a majority vote of the classifiers used. This technique outperforms individual classifiers; however, a more sophisticated approach was developed in [117]. Bostrom *et al.* in [117] use ensemble classifier, which is composed of a set of classifiers; the ensemble output is dependent on the classifiers outputs'. Contrary to most approaches, which implemented a Dempster,'s rule of combination, the authors in [117] applied Shafer's theory of evidence for the classifier fusion. The results of this research show that among the evidentiary rules, the combined ones appear to have better accuracy than the non- combined. However, in general, the evidential combination rule does not perform better than the voting rules for this particular ensemble design because the choice of combination rule can have a significant impact on the performance for a single dataset.

In other research, [118] presented an ensemble of Radial Basis Function Neural Networks (RBFNs) method to identify the epileptic seizure, by optimizing the Differential Evolution (DE) and analyses of the EEG signal. This method was based on the bagging approach and using the DE-RBFNs as the base classifier. Discrete wavelet transform (DWT) was utilized to decompose EEG into different sub-bands. The DWT uses multi-resolution filter banks and special wavelet filters for the analysis and reconstruction of EEG signals. They extract some basic statistical features to apply them as inputs for the ensemble method. They did three different experiments to get the performance of the presented ensemble method in the classifications of normal and seizure segments. Results have reported a promising performance and the authors assured that the proposed ensemble method is better than the individual classifier.

Diversity classifiers, which in turn make a final combining instruction, play a critical role such as learning scenarios. He and Cao in [110] proposed a signal strength-based combining (SSC) method to combine the outputs of multiple classifiers to support decision-making in classification. They assumed that each classifier ensemble is accompanied with a decision profile $P_d(Y_i|x_t)$, which is described as the voting probability from each hypothesis h_j for each testing instance x_t through all possible class identity labels. The elements in $P_d(Y_i|x_t)$ can be obtained directly from each hypothesis output, from the confusion matrix based on the training data or cross-validation method, depending on the different base learning algorithms. This method integrates different classifiers in an ensemble system. Comparisons have been made for the proposed method with nine major existing combining rules. In addition, the authors discussed the relationship of the proposed method with respect to margin-based classifiers including the boosting methods (AdaBoost.M1 and AdaBoost.M2) and SVMs by margin analysis. The results show that the proposed method is competitive when compared to the existing classifiers.

Other research in [119] presented regularized common spatial patterns (R-CSP) and aggregation techniques for EEG signal classification in a small sample setting (SSS). An ensemble-based solution was given through an aggregated number of R-CSPs in order to tackle the problem of regularization parameter determination. The regularized CSP uses two parameters, to lower the estimation variance, and to reduce the estimation bias, where the rule of generic learning is applied in the regularization process. The cross-validation method was employed to determine the regularization parameters of the R-CSP for the EEG signal classification in SSS. The classification accuracy rates were promising. The overall performance accuracy was 83.9%.

As seen in the literature above, further research is required on detection and classification, especially on compressed and noisy EEG data with different levels of

SNR. In addition, the current techniques did not estimate the classification accuracy in the case of noisy data and the effect of reconstruction error due to compression. Therefore, a new method is desirable for classifying the compressed and noisy EEG data.

The research gap can be concluded from view of the above literature survey. It has been found that there is no previous work that focusses on the compression and reconstruction of noisy EEG especially when using the AWGN as the channel model. Furthermore, there exists no work on the EEG classification problem due to the noise or the reconstruction error. Furthermore, most of the above research works focus on two classes representing healthy and unhealthy subjects; they did not work with the seizure-free intervals subjects. Therefore, this research will investigate the effect of reconstruction error due to compression (e. g, using combination of discrete cosine transform (DCT) and random matrix) on the classification accuracy. In addition, four statistical features from discrete wavelet sub-bands are extracted to collect 32 attributes of the EEG signal. The classification accuracy of noisy compressed EEG data has been evaluated using different individual classifiers. Finally, an ensemble classifier is proposed to enhance the classification accuracy of the noisy compressed EEG data. The following table 2.2 shows that the current works compare with the proposed work.

Table 2.2:

Comparisons between current work and the proposed work

No.	Current Works (same dataset)	Proposed work
1	EEG raw data classes [51, 53, 66, 67, 83, 84, 92, 94, 116, 117, 118]	EEG compressed and noisy
2	Using only two classes of healthy and unhealthy [51-54, 89-94, 100-106, 110, 118]	Using three classes of healthy, healthy with seizure free interval, and unhealthy
3	Did not address the compressed and noise, working only on undistorted EEG data [32, 40, 52, 75, 78, 79,84, 86, 92, and many others]	Address the compressed data and noisy using AWGN as the channel model
4	Current techniques did not estimate the classification accuracy in the case of noisy data [64, 74, 82-84, 88, 102, 103, 108, and many others]	Propose an ensemble classifier to measure the accuracy of noisy data
5	Did not address the effect of reconstructed error due to data compression [32, 76, 82-84, 88, 102, 103, 108, and many others]	Address the effect of reconstruction error due to compression

2.6 Summary

This chapter introduced and extensively analyzed the WBASNs signal processing and compression framework. The framework consists of two major components for EEG-based epileptic seizure, namely sensing and preprocessing (SAP), detection and classification (DAC) of EEG-based epileptic seizure.

In SAP, EEG raw data was used that are related to medical phenomena. To do so, three preprocessing techniques have been surveyed: 1) Sampling to provide a formula for the minimal sampling interval required to avoid aliasing based on Nyquist theorem. 2) Filtering to clean and revise the duplicate records and inconsistent data acquisition, which ultimately achieve a reduction in the redundancy data. 3) Compression technique was used to reduce the size of physiological data traffic that sensors send over the wireless channel using lossless and lossy methods. While EEG-based epileptic seizure data is representing the brain disorder, which is bandwidth intensive. The compression is mainly our interest of this research work.

In DAC, analysis of raw data, including, possibly, detection and classification of medical anomalies will occur to get strict and accurate criteria for the physician to make recommendations that are usually feedback to the patient to provide proactive treatment. In fact, data feature extraction, detection and classification providing efficient tools for enhancing the diagnosis of illnesses in various clinical and life settings. Finally, features extraction was achieved by DWT, and classification was by four individual classifiers. Due to the impairment channels and EEG data which resulting in noise or lost some information a Noise-aware signal combination method has been proposed as an ensemble classifier to enhance the classification accuracy.

CHAPTER THREE
RESEARCH METHODOLOGY

This chapter describes the research methodology used in our work. First, it started with a literature review of the research problem. Then, two different frameworks for this problem have been designed and developed. Afterwards, a mathematical formula is formulated for the computation of the hypothesis with the highest probability, and proposes a new ensemble method for non-noised and noised EEG data. To obtain an efficient solution, the key significant features of two frameworks applied to numerous and various inputs were used.

The proposed solution through extensive experimental and simulation benchmarks has been validated and refined. For compressed and reconstructed EEG data, a simulation program to elaborate the proposed framework is developed, in order to find better classification accuracy and compression ratio. In addition, a simulation program has been developed for the mathematical formula underlying the proposed classification technique in order to demonstrate the proposed ensemble method. Finally, the framework has been evaluated in terms of compression ratio and classification accuracy, the proposed ensemble technique in terms of noisy data, and compared the results of the proposed solution to existing methods, in order to verify its efficiency.

3.1 Background

Figure 3.1 illustrates the general research procedure and shows the different steps of this procedure for EEG-based compression and classification.

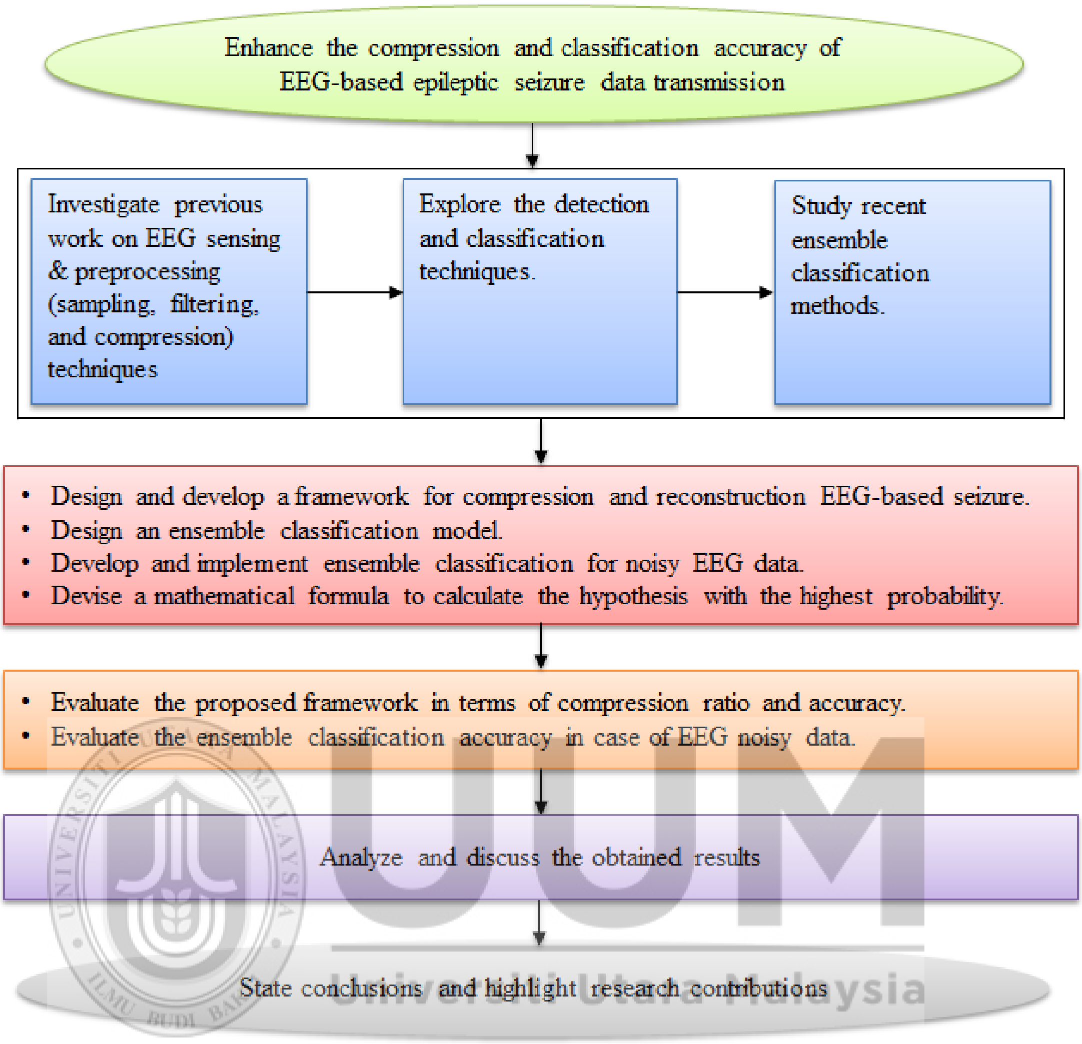


Figure 3.1. General Research Procedure

This section presents the developed compression and classification framework to efficiently compress and reconstruct EEG signals using combined different methods. It describes the compression and classification framework that employs and combines many methods, namely; CS and DCT techniques, with a measurement random matrix and using AWGN as a channel model for EEG data compression. DWT is used for feature extraction, and four classifiers are employed for EEG signal classification. In addition, noise-aware signal combination (NSC) technique is proposed for noise EEG signal classification. The presented below are different blocks and algorithms that make up our proposed framework. Where discusses the compression, the transmission channel model, the reconstruction, the feature extraction, explores four classifiers, and briefly reviews the proposed method. Figure 3.2 depicts the block diagram of the proposed framework. It shows the different steps of the proposed framework for EEG-based compression and classification.

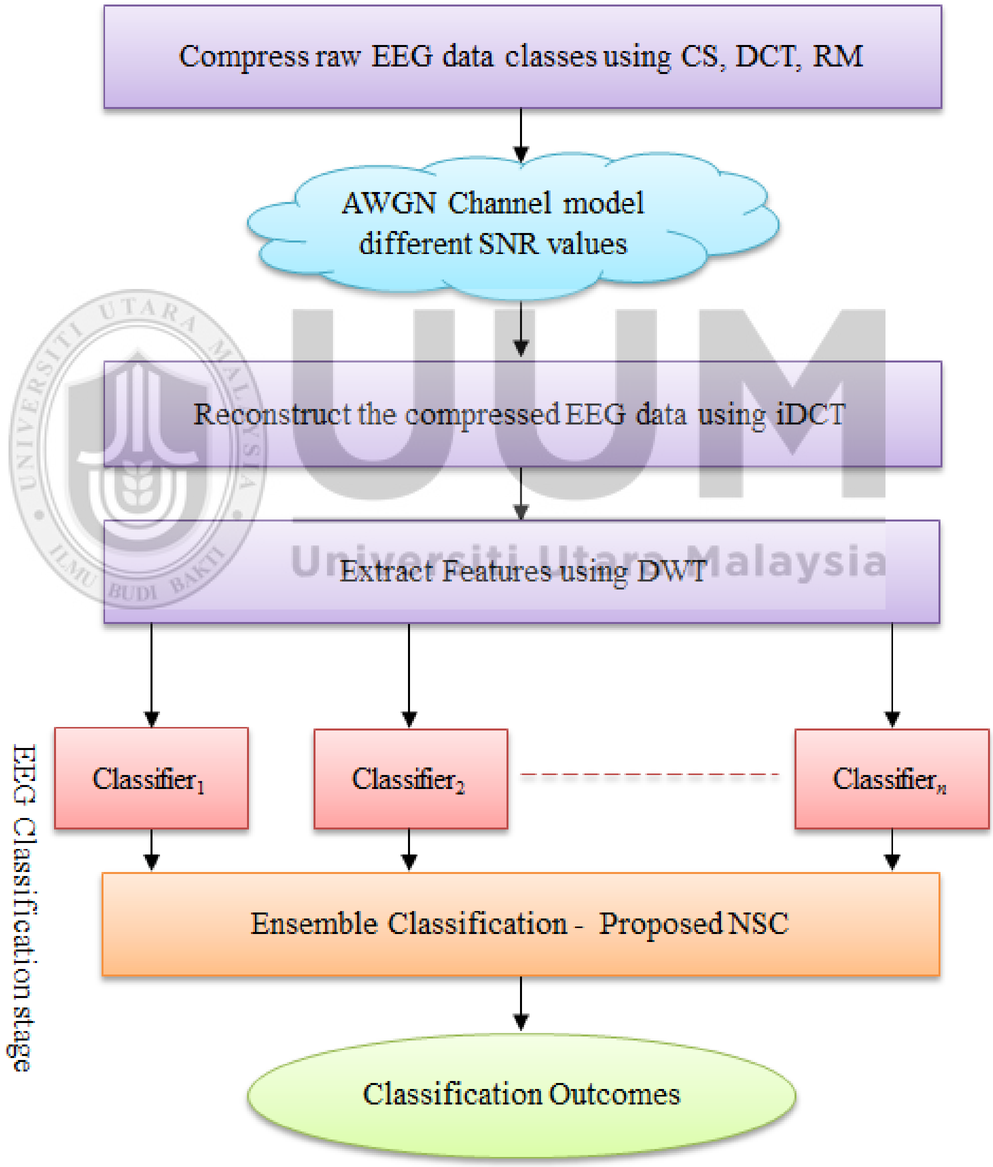


Figure 3.2. Proposed framework methods

The proposed method of this research is mainly consists of six parts:

1) Compressive sensing (CS) and DCT methods in addition to the measurement random matrix (RM) have been used for EEG data compression.
2) AWGN a channel model is utilized to simulate the real environment to study the effect of noise on the compressed data.
3) iDCT for data reconstruction to the original EEG data.
4) DWT is also used in order to extract statistical feature, such as, min, max, mean and standard deviation.
5) Four individual classifiers have been employed for EEG data classification.
6) Ensemble classification method and how the classification result is enhanced by our proposed NSC technique.

In comparison with traditional communication technologies, it is necessary for WBASN in health application context to keep working continuously and gradually for a long time. Therefore, effectively decreasing the transmitted size of EEG data during the data transmission, this will reduce the transmission time which ultimately reduce the power, is a major problem to be addressed.

Many research works like [120, 69, and 121] strived to reduce the sampling rate below the Nyquist rate, without causing significant aliasing. This is obvious, especially when the signal is sparse, i.e. has frequency spectrum holes. A sparse signal is a signal that contains a small number of dispersed frequency components, higher than zero, in some transform domain [69]. This is to motivate the investigation of calculating the optimal sampling rate below the Nyquist rate, so that the signal characteristics are preserved, the best technique known for this purpose is compressive sensing.

3.2 Compressive Sensing Technique

In the last few years, compressive sensing (CS) is rapidly growing and has attracted much attention in the area of signal processing and computer science, by proposing that it may be achievable to exceed the limits of traditional sampling theory. In this field, researchers have dedicated hundreds of conferences, special sessions, and workshops. In addition, thousands of papers were published in this growing area [70]. CS developed as a new context for signal acquisition and sensor design to influence the transform coding concept. In an appropriate basis, CS constructs upon the fact that using a few non-zero coefficients (sparse) yields to represent several signals.

In the sampling and computation costs for sensing signals, CS enables a significant reduction when they have a sparse or compressible representation [71]. While the Nyquist-Shannon sampling theorem declares that a certain minimum number of samples is required to perfectly capture an arbitrary band-limited signal, the number of measurements that need to be stored can be reduced when the signal is sparse in a known basis [72]. Accordingly, when sensing signals are sparse, the proposed framework is able to perform better than that proposed by traditional methods.

The basic idea of CS is to find techniques to directly sense the data in a compressed form instead of first sampling at a high rate and afterward compressing the sampled data, i.e., at a lower sampling rate [73]. The CS goal is to design a stable measurement matrix, Φ, and allow the reconstruction algorithm to recover a signal x of length N from $M < N$ measurements. Compressible signals M size are well approximated by *k*-sparse representations.

Consider a general linear measurement process that computes $M < N$ inner products between x and a collection of vectors $\{\phi_j\}_{j=1}^{M}$ as in $\langle x, \phi_j \rangle$. The measurements y_j in an $M * 1$ vector y and the measurement vectors ϕ_j^T as rows in an $M * N$ matrix ϕ, the measurements have:

$$y = \Phi x = \Phi \Psi s = \Theta s$$

where $\Theta = \Phi\Psi$ is an $M * N$ matrix.

For example, the following is given a compressed measurement y at the receiver, the signal x can be reconstructed by solving one of the following optimization problems.

$$\text{Minimum } \| x_0 \|_2 \text{ Subject to } y_i = < \Phi_{\mathrm{i}} \Psi x_{0\mathrm{i}} > \tag{3.1}$$

Use a trick of basis Pursuit in order to find the vector x_0 with the lowest L_2 norm that satisfies the observations made. For N-dimensional EEG signal x:

$$x = \Psi \alpha \tag{3.2}$$

where Ψ is a DCT basis and α is the wavelet, both are in domain coefficients. While CS is used to reduce the size of data required to send from transmitter to receiver, it has been considered for efficient EEG acquisition and compression in several application contexts [66]. Since the signal acquisition is the critical part in most applications, where the acquisition time or the computational resources are limited, CS technique has the main advantage that it offloads processing from data acquisition until data reconstruction. CS acquires measurements faster due to fewer sizes of samples. CS performs the time consuming processing in the recovery, where you mostly have more processing time and higher computational capacity.

An *N*-dimensional 4096 samples raw EEG signals *x* is considered to illustrate the CS compression and reconstruction. Assume that this signal is represented by a projection on to a different basis set Ψ:

$$x_0 = \sum_{i=1}^{N} x_i \Psi_i \quad \text{or} \quad x = \Psi x_0 \tag{3.3}$$

where *x* is the original signal, x_0 is the sparse of representation of *x,* and Ψ is an $N * N$ bases matrix.

The sparse vector x_{0i} can be calculated from the inner product of *x* and Ψ:

$$x_{0i} = < x, \Psi_i > \tag{3.4}$$

The basis (Ψ) can be Gabor, Fourier, or DCT, Mexican hat, Linear Spline, Cubic Spline, Linear B-spline, or Cubic B-spline basis. In compressive sensing, Ψ is chosen such that x_0 is sparse. The vector x_0 is k-sparse if it has k non-zero entries and the remaining (*N*–*k*) entries are all zeroes.

In addition to the projection above, it is assumed that *x* can be related to another signal y:

$$y_{[M*1]} = \Phi_{[M*N]} \times x_{[N*1]} \tag{3.5}$$

where Φ is a measurement matrix (also called sensing matrix) that depends upon the sparsity k level of the signal, and *y* is the compressive sensed version of *x*. The measurement matrix is of dimensions *M* * *N*, where *N* represents the original raw EEG data and *M* represents the different compression values. Matrix *y* has dimensions *M*1* and if *M* < *N*, then data compression is achieved. The measurement vectors length is also determined by the *K* sparsity in the signal *x*. CS can sample the signal by much lesser measurements than those required by the Nyquist sampling theorem. Two major concepts are Sparsity and Incoherence, which includes sparse representation, random measurements taken and signal recovery via $\ell 1$ minimization. It can be shown that this technique is possible if Φ and Ψ are incoherent. To satisfy this condition, Φ is chosen as a random matrix. The Compression Ratio (CR) is then defined as [11]:

$$CR = \left(1 - \frac{M}{N}\right) * 100 \tag{3.6}$$

DCT is used as the basis to sparsity of the EEG signal as part of the CS framework. It is a Fourier-related transform like the discrete Fourier transform (DFT). However, it is using only real numbers, and has a low computational complexity [19]. In order to obtain the signal $x(n)$ in the DCT domain, that will lead to the definition of the $(N+1) * (N+1)$ DCT transform matrix, whose elements are given by:

$$[C]_{mn} = \sqrt{\frac{2}{N}} \left\{k_m k_n COS\left(\frac{mn\pi}{N}\right)\right\}, m, n = 0, 1, \dots, N \tag{3.7}$$

$k_i = 1$ for $i \neq 0$ or N

= $1/\sqrt{2}$ for $i = 0$ or N.

This matrix is unitary and when it is applied to a data vector x of length $N+1$, it produces a vector called $\mathrm{X_c}$, where $\mathrm{X_c} = [\mathrm{C}] * \mathrm{x}$, whose elements are given by,

$$\mathrm{X_c}(m) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N} k_m k_n COS\left(\frac{mn\pi}{N}\right) \mathrm{x}(n) \tag{3.8}$$

The following section gives brief information about the channel medium that has been considered in this research work.

3.3 Additive White Gaussian Noise (AWGN) Channel Model

Several applications containing wireless sensor networks, smart homes and machines, automated factories, and remote tele-medicine were developed from research ideas to solid systems. When communicate over the wireless channel, there is an impairment interfering with the signal from the medium. Several sources of noise can alter the data, including, wireless channel fading, path loss, thermal noise at the receiver, etc. In this research, without loss of generality, the thermal noise using the AWGN model at the receiving side has been considered as the most widely used model for representing thermal noise [74, 122, 61 and 88]. The noise level using the signal-to-noise-ratio (SNR) has been controlled to demonstrate data imperfection, also, to study the behavior of the different classification techniques the presence of such noise. The transmitted data over the communication channels are impacted by attenuation, interference, and thermal noise.

Additive White Gaussian Noise (AWGN) is one of the models that represents noise on the transmission channel that occur in real life situations. Noise effects are formed by AWGN with a power spectral density that is based on the channel signal-to-noise ratio (SNR). Using AWGN channel, noise has been added to the transmitted compressed data based on a specific SNR, therefore, the received data can be introduced as: $r(t) = s(t) + n(t)$, where $s(t)$ is the transmitted data and $n(t)$ is a noise.

To use the efficiency of the brainwaves for a particular reason that is built on the performance of the wireless system. Certainly, wireless communication provides us with more possibilities to employ EEG signals transmission [88]. However, due to the

AWGN, the channel capacity could be affected according to the Shannon-Hartley theorem, which is expressed as follows [122]:

$$C = Blog_2\left(1 + \frac{S}{N}\right) \tag{3.9}$$

where C is channel capacity of maximum error-free and measured in bits per second; B is the bandwidth of the channel and measured in hertz; S is the average received signal power over the bandwidth; N is the average noise or interference power over the bandwidth, measured in watts or v^2. Also, the signal-to-noise-ratio is defined as the ratio of signal power to the noise power, often measured in decibels. The decibel is a logarithmic unit used to express the ratio between two values of a physical quantity, often power or density. Most of the metrics used to evaluate the performance of wireless communication algorithm are computed for different SNR levels. The SNR is the ratio between the received signal power and the noise power:

Signal-to-Noise-Ratio (SNR) is given by:

$$SNR = S/N \tag{3.10}$$

And SNR in dBs

Signal-to-Noise-Ratio (SNR_{dB}) is given by:

$$SNR_{dB} = 10\ log_{10}\ (S/\ N)\ dB \tag{3.11}$$

Note that as SNR increases, if the noise power is assumed to be constant, this indicates that the received signal power increases and vice versa. An SNR_{dB}=1 dB means that the received signal power (S) equals 1.2589 times the noise power (N). An SNR_{dB}=5 dB means that the received signal power (S) equals 3.1623 times the noise power. An SNR_{dB}=10 dB means that the received signal power (S) equals 10 times the noise power [68, 74, and 129].

In other words, the relation between SNR and S is proportional, i.e., as SNR increases, the effect of noise in the received signal diminishes [74]. Moreover, the relation between SNR and N is inverse proportional, i.e., as SNR decreases, the received signal will be more contaminated with noise [75]. We choose to simulate the performance of our proposed algorithm at the SNR levels presented earlier, which represents three levels of SNR, low: SNR = 1dB, moderate SNR =5 dB and acceptable SNR =10dB [75, 76].

The motivation of using AWGN is to study the effect of the wireless channel on the proposed framework. In addition, it is to study the proposed NSC with different SNR levels. In order to study the noise effect on the compressed EEG data, the AWGN channel model as a transmission channel has been used. At the receiver side, a data reconstructed algorithm has been utilized for getting some features and then doing detection and classification. SNR values are selected based on extensive experiments to represent high, moderate, and low noise cases [131, 114]. However, the communication environments are outside of our research scope.

3.4 Signal Reconstruction

The well-known traditional approach of reconstructing signals or images from measured data is Shannon Nyquest sampling theorem [124], which states that the sampling rate must be the double the highest frequency. A fundamental new approach has been provided by compressed sensing (CS), or sparse recovery, to data acquisition which overcomes this common problem. CS predicts that specific signals or images can be recovered from what was previously believed to be highly incomplete measurements.

Certainly, in order to compress a signal *x*, one may simply store the k largest entries. The non-stored entries are set to zero for reconstructing *x* from its compressed version, and the reconstruction error. At this point, it is highlighted that the procedure of achieving the compressed version of *x* is adaptive and nonlinear since it needs the search of the largest entries of *x* in absolute value. Especially, the location of non-zero entries is a nonlinear type of information.

On the receiver side, the basis of the inverse discrete cosine transform (iDCT) has been utilized in the CS obtain the reconstructed signal (x_r) [125], $x_r = iDCT\ (y)$ returns the inverse DCT of y, i.e., the original signal as follows [114]:

$$x_r(a) = \sum_{k=1}^{N} w(k)\ y(k)\ cos\left[\frac{\pi(2a+1)k}{2N}\right] \tag{3.12}$$

where N is the length of both time series and cosine transform signals, a is the time series index $(a = 1, 2, \ldots N)$, k is the cosine transform index $(k = 1,2, \ldots N)$, and the window function $W(k)$ is defined as,

$$w(k) = \begin{cases} \frac{1}{\sqrt{N}} & k = 1 \\ \sqrt{\frac{2}{N}} & 2 \leq k \leq N \end{cases}$$

After obtaining the contaminated reconstructed signal (x_r), then DWT is used as a feature extraction and selection techniques.

3.5 Feature Extraction

EEG Feature Extraction plays a significant role in diagnosing most of the brain diseases. Obtaining useful and discriminant features depend largely on the used feature extraction method. The feature extraction stage must reduce the original data to a lower dimension that contains most of the useful information included in the original vector. It is therefore necessary to find out the key features that represent the whole dataset, depending on the characteristics of the dataset. These features are calculated from each cross-correlation sequence to create feature vector sets. Since EEG signals are time-varying and space-varying non-stationary signals, the DWT method is widely used [51]. It captures both frequency and time location information. Using multi-resolution wavelet analysis, DWT basically decomposes the EEG signals into different frequency bands.

EEG data is generally a non-stationary signal, which is heavily dependent on the subject condition. The DWT Daubechies 6 was employed, where the data was sampled at a rate of 173.61 Hz. This means that the EEG data frequency is 86.81 Hz, so the filter length is long as well; frequency wavelet sub-band is the same as the fundamental component of EEG. Hence, the decomposition level 7 is determined based on the EEG frequency. In addition, based on extensive experimental work on the reconstruction accuracy for different wavelet families and filter lengths and decomposition levels, in this research, Daubechies 6 with different decomposition levels from 1-8 have been used. The interesting thing that was found is that the Daubechies 6 with decomposition level 7 is the optimum level in terms of the classification accuracy and computational complexity of the EEG epileptic seizure category of data. Given the EEG signal $f(x)$, the Wavelet series expansion is depicted in [122] and computed as follows:

$$f(x) = \sum_{k} c_{j0}(k)\varphi_{j0,k}(x) + \sum_{j=j_0}^{\infty}\sum_{k} d_j(k)\, \psi_{j,k}(x) \tag{3.13}$$

where $f(x) \in L^2(R)$ and $L^2(R)$ are relative to wavelet $\psi(x)$ and scaling function $\varphi(x)$, c_{j0} is the approximation coefficients.

In the first sum, the approximation coefficients c_{j0} can be represented as the outcome of the inner product process between the original signal $f(x)$ and the approximation function $\varphi_{j0,k}(x)$ as expressed [122]:

$$c_{j_{\mathbf{0}}}(k) = \langle f(x), \varphi_{jo,k}(x) \rangle \tag{3.14}$$

In the second sum, a finer resolution is added to the approximation to provide increasing details. The function $d_j(k)$ represents the details coefficients and it can be obtained by the inner product between the original signal $f(x)$ and the wavelet function $\psi_{j,k}(x)$ calculated as:

$$d_j(k) = \langle f(x), \psi_{j,k}(x) \rangle \tag{3.15}$$

where *j* and *k* represent the scaling and time shifting parameters.

Generally, combined time-frequency domain features outperform, in terms of classification accuracy for frequency domain-based and time domain-based independently features [122]. Different implementation choices, including different wavelet families; filter lengths, and decomposition levels have been experimented for the feature extraction purpose. Accordingly, four conventional statistical features minimum, maximum, mean, and standard deviation were extracted from each wavelet sub-band composed of 32 attributes. These features represent the statistical characteristics of the EEG signal and are recommended for real-time applications. They also have been found the most representative parameters that are able to distinguish different signals.

The main advantage of these four features is their low computational complexity and computation time. They are the most representative values to describe the distribution of the EEG signals. They are used to reduce the dimensions of the cross-correlation sequences and as inputs into individual classifiers. The statistical features extraction rules that have been implemented on the wavelet sub-band [19, 84, and 122] are:

Maximum feature: x_k such that $x_k > x_i$ for all $i \neq k,$ i=1,…….,n.

$$d_i(x) = \underset{\mathrm{i=1,\ldots,k}}{\mathrm{Max}}\{d_i(x)\} \tag{3.16}$$

Minimum feature: x_k such that $x_k < x_i$ for all $i \neq k,$ i=1,…….,n.

$$d_{\boldsymbol{i}}(x) = \underset{\mathrm{i=1,\ldots,k}}{\mathrm{Min}}\{d_i(x)\} \tag{3.17}$$

Mean can be calculated by:

$$\hat{x} = \frac{1}{N}\sum_{i=1}^{n} x_i \tag{3.18}$$

Standard Deviation: is a statistical feature, which indicates the distribution of the data with respect to the mean.

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N} (x_i - \hat{x})^2 \tag{3.19}$$

The original EEG signal was analyzed for the wavelet sub-bands A7 and D7-D1. Eventually, four conventional statistical features are selected from each wavelet sub-band individually. As a consequence, 32 attributes are obtained from the whole sub-bands to be fed to the classifiers. So these features maximum, minimum, mean, and standard deviation are contributed to the classification accuracy in this research. In this research, it has been found that these features are robust with the dynamic environment of the wireless channel [19, 114]. Meanwhile, these features have low computational complexity.

3.6 Classification Methods

EEG detection and classification play an essential role in the timely diagnosis by analyzing potentially fatal and chronic diseases proactively in clinical as well as various life settings.

In this work, four of the common and predictive classification methods have been used namely: ANN, NB, *k*-NN, and SVM. Based on the kind of datasets, each method shows separate efficiency and accuracy. Each one of them has special advantage as follows:

- ANN provides a better result in complex domain because its effectiveness and non-parametric and testing is very fast.
- NB is easy for implementation and computation, it estimates the probability of each class using Bayes rule.
- k-NN is simple, effective and easy to implement.
- SVM is compact description of the learned model, more capable to solve multi-label classification.

Initially, the classifiers have been implemented to work individually for performance comparison purposes. Each classifier belongs to a different family of classifiers, and has been shown to be the best classifier in its family. However, since they are using a different classification strategy, it is expected that these classifiers may yield different classification results. An ensemble method for combining the output of all classifiers has been developed in order to reduce the effect of data imperfection, while maximizing the classification accuracy. For more information regarding the individual classifiers refer to Chapter 2. The following section will present the ensemble-based classification techniques. Then, it will include our proposed noise-aware signal combination method to enhance the classification accuracy.

It is worth noting that other classifiers (BayesNet, DecisionTable, IBK, J48, and VFI) have been experimented. They also provide a good accuracy for the same data are used in the noiseless data and noisy data with SNR= 1, 5, and 10 dB. However, the other classifiers are effective and easy to implement.

3.7 Ensemble-Based Classification

Several combination techniques have been introduced in the literature and each one may offer certain advantages while suffering from certain limitations. One of these well-known combination techniques is the majority vote. The majority voting (MV) rule technique collects the votes of all classifiers and investigates the class name that is mostly reported by the classifiers. It then chooses that class as a final decision [96]. However, MV is based on the idea that the classifiers participating in the voting process have the same weight, and it then completely ignores the inconsistency that may arise among the classifiers. Thus, the performance of the classification can be deteriorated. For this research, a probability-based voting scheme has been adopted, in which a combination method should assign a probability value (p) that reflects the classifier confidence in the viewpoint of the combination method.

For instance, a weight (p) of 0.70 has been assigned to the first classifier, while assigning only 0.30 as a weight to the second classifier.

Let T be the set of classifiers,

$$T = \{C_1, C_2, \ldots, C_n\}, \tag{3.20}$$

and C be the set of classes,

$$\{p_1, p_2, \ldots, p_n\},$$

Let $d_{i,j}$ be the decision of classifier i defined as follows:

$$d_{\boldsymbol{i},\boldsymbol{j}} \in \{0,1\}, \tag{3.21}$$

where $i = 1, \ldots, T$, and $j = 1, \ldots, C$.

Let p_i represent the weight of a classifier i. The probability-based voting decision is calculated as:

$$\sum_{i=1}^{|T|} p_i d_{i,j} = \max_{j=1}^{C} \sum_{r=1}^{|T|} p_r \, d_{r,j}, \tag{3.22}$$

Considering the weight for each classifier, Equation (3.22) counts the votes coming from the participating classifiers. The following sub-section shows the proposed Noise-aware Signal Combination (NSC) method.

3.7.1 The NSC Proposed Ensemble Method

A number of classifiers (*n)* built on various hypotheses $H = \{h_0, h_1, \dots, h_{n-1}\}$, are fed with input data. Each classifier *k* built on hypothesis h_k, is trained on the data in order to predict the label representing the class c_j that best describes a given set of features $(f_{i,0}, \ f_{i,1}, \dots, f_{i,l})$, corresponding to observation, o_i.

At the end of the training of each classifier, a set of multiclass classification performance measurements of interest is recorded. More details and information about the proposed noise-aware signal combination method and its performance measurements are provided in Chapter 4.

Briefly, the proposed NSC method is fed with the output of the, *n* trained classifiers. The classification decision of a testing sample is obtained by gathering the decisions from the corresponding n classifiers at each layer using Noise-aware Signal Combination method. A subset of the performance measures of each classifier together with the predicted class label $c \in C$ for an observation $o \in O$ provided by each classifier with hypothesis $h \in H$, are used to construct the confusion matrix for each classifier. These confusion matrices form the input to the hypothesis used by this combined classifier. More information about the confusion matrix is available in Chapter 4.

3.7.2 Classifiers settings

- The ANN has several parameters, in this research the configurations of ANN are training cycles=500, learning rate=0.3, and momentum decay=0.2 were used.
- NB; Laplace correction to prevent high effect of zero probabilities is used as the default configurations.
- K-NN; in this research, the default configurations are value of *k*=10 was used, and the *mixed measures* was selected as the measure type, which makes the *Mixed Euclidean Distance* the only available option.
- SVM, have several parameters, in this research, SVM configurations are *nu*-SVC and *radial basis function* kernel were used for classification technique.

3.8 Classification Measurements

To measure and evaluate the accuracy of this framework, the following subsections will be considered. The following section will describe in details the EEG benchmark dataset that was used in this research work.

3.8.1 EEG Datasets Descriptions

The dataset used in this work is one of the most comprehensive dataset, it is very widely used which indicate the correctness of the results; it fits the application that is being targeted. The dataset was created by Andrzejak *et al.* [126], which are widely used for automatic epileptic seizure detection. It contains a benchmark data including both normal and epileptic EEG datasets. EEG datasets were collected from five patients. The patients passed through a complete seizure control after resection of one of the hippocampal formations (identified as an epileptic zone). The datasets consisted of five sets and indicated as A, B, C, D, and E. Each set was composed of 100 single channel EEG segments of duration of 23.6 seconds.

Both sets A and B were relaxed in a conscious state with eyes open and eyes closed, respectively. Segments of sets A and B were taken from surface EEG recordings that were carried out using a standardized electrode placement scheme, performed on five healthy subjects. Segments in set C were recorded from the hippocampal formation of the opposite hemisphere of the brain. Segments in set D were recorded from within the epileptogenic zone. While sets C and D contained only brain activities measured during seizure free intervals. Only set E contained seizure activity. All EEG signals were recorded with the same 128- channel amplifier system, as stated by Andrzejak *et al.* in [126]: “neglecting electrodes that having strong eye movement artifacts (A and B) or pathological activity (C, D, and E)”.

The data was written continuously to the disk of a data acquisition computer system, at a sampling rate of 173.61 Hz. Kumar [127] reported that when the performance of set A and E was compared with set B and set E, and it was concluded that set A and set E were more efficient [127]. In addition, set A and set B are similar in feature properties that are hard for the classifier to distinguish between both sets represent healthy patients. It is worth noting that during performance evaluation many experiments using different groups of classes (i.e. one group was all five classes, another group was A, C, and E, etc.) have been conducted, and the best results were evident for the class group of A, C, and E.

Therefore, this research uses only three sets: set A to represent healthy subjects, set C to represent unhealthy with seizure-free interval subjects, and set E to represent the epileptic seizure active subjects. In this case, 300 EEG segments used, each class 100 segments. Figure 3.3 illustrates the ideal raw EEG signals of sets A, C, and E, respectively.

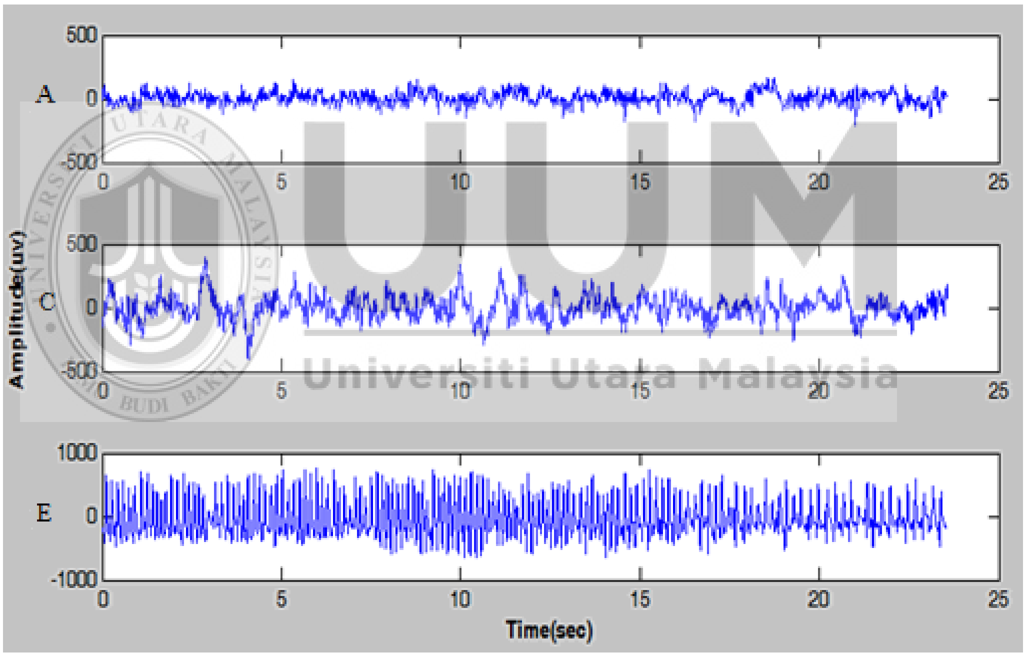


Figure 3.3. EEG signals of three classes taken from different subjects [126]

3.8.2 Data compression and reconstruction

For measuring the effect of compression, a compression ratio metric has been used for measuring the loss of data information. This can be measured by dividing the compressed data size *M* on the original data size *N*, and then, computing the difference from 1 and multiply by 100. Therefore, the compression ratio (CR) is defined in Equation (3.23), as expressed:

$$CR = \left(1 - \frac{M}{N}\right) * 100 \tag{3.23}$$

where CR is the compression ratio, *M* is the compressed data value at a time, *N* is the original data value at a time.

3.8.3 Detection and classification

The classification accuracy is categorized into specificity and sensitivity, which are defined as a function of the true and false of both positives and negatives. Accuracy is the degree of proximity of a measurement quantity to that quantity is true [113]. Specificity in diagnostic refers to the ability of an assessor to measure one particular material. Specificity is defined as a percentage ratio of true negative tests for the total number of infected patients tested. It is also, known, as a class precision. Moreover, sensitivity in diagnostic testing represents the smallest quantity of material in a sample that can be accurately measured by an assessor. Sensitivity is defined as a percentage ratio of true positive tests for the total number of affected (positive) patients being tested. It is also known as a of class recall.

Finally, classification accuracy refers to the range to which a testing tool is able to strictly classify the certainty and uncertainty data [60]. Therefore, specificity, sensitivity, and accuracy can be defined as follows [11]:

$$Specificity = \frac{TN}{TN + FP} \times 100 \tag{3.24}$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \tag{3.25}$$

$$Accuracy = \frac{TP + TN}{TN + FP + TP + FN} \times 100 \tag{3.26}$$

where TP, FP, TN, FN refer to true positive, false positive, true negative, and false negative, respectively. Both positive and negative terms are denoted the prediction of classifiers', and true and false is indicated whether that prediction is reacting with the external judgment. These equations show that a test with high specificity has few FPs; where as a test with high sensitivity has few FNs. For more information about these metrics, please refer to Chapter 2.

3.9 Summary

Since the WBASNs are operating on batteries or solar cells, this point is critical. Using the most popular compression techniques for data acquisition, compressive sensing can reduce the transmitted data size. This chapter presents in details the research methodology adopt for the general framework that employs and combines many methods. The framework incorporates CS-based energy-efficient compression, and noisy wireless channel to study the effect on the application accuracy. This ultimately effects the power consumption. In addition, the research method model has been highlighted which illustrates and summarizes the proposed framework. This model shows several steps of the proposed framework for EEG-based compression and classification. It discussed the compressive sensing and showed that the goal of CS is to design a constant measurement matrix Φ, and to allow the reconstruction of a signal

x of length N from $M<N$ measurements. In addition it covered the DCT as the basis used to handle the sparsity of the EEG signal as part of the CS framework. AWGN was used as a channel model that represents noise on the wireless channel, occurring in real life. iDCT was used to reconstruct the compressed data to the original size at the receiver side. DWT was used to extract the number of features necessary for data classification. Finally, due to the compressed and decompressed EEG data, usually the information is accompanied by noise artifacts, which may hamper the performance of EEG-based classification techniques. Moreover, data is compressed in order to reduce transmission time and cost. Consequently, some important information may get lost. Therefore, it proposed a noise-aware signal combination method as an ensemble classifier.

CHAPTER FOUR

COMPRESSION AND CLASSIFICATION MODEL

This chapter introduces a new compression and classification framework, in addition to a new Noise-aware Signal Combination (NSC) method based on four individual classifiers, for the classification of three classes of EEG signals. This study aims to establish a method to determine an optimal classification scheme and to deduce the signs of the extracted features.

In this chapter, a new algorithm has been proposed which distinguishes EEG segments from different pairs of three-class EEG signals in a multiclass EEG dataset. In addition, diagnosis differences between neurophysiologists may occur for the same EEG recording due to the individual description of the analysis.

In other words, on the data acquisition side a highly accomplished framework that presenting CS technique with high compression ratio has been developed [19]. On the detection and classification at the server side, a new combination method for detection and classification has been developed [128]. The developed method operates on imperfect data and provides acceptable accuracy given EEG-based epileptic seizure data compared with previous work in the literature review.

4.1 EEG Data

Despite the fact that EEG signals offer a significant deal of brain activity information, classification and evaluation of them are still a major problem of research [129]. Nowadays, EEG signal is almost inspected manually by experts, classification techniques will help in distinguishing the EEG healthy and non-healthy subjects. Therefore, machine intelligence techniques are proposed to diagnose, evaluate, and enhance the process of the epileptic seizure detection. The big challenge here is that EEG compressed data is noisy and needs to reconstruct it at the server side, also using correct classification algorithms to classify properly and efficiently different EEG signals of different SNR values. Using the CS framework, measure the classification accuracy for an ensemble combination method, in order to reduce the size of data at the transmitter.

At the server side, a detection and classification technique has been developed for the recovered EEG data, while keeping the overall classification accuracy for each classifier at a satisfactory minimum level of 80% [19]. This chapter focuses on the design of an efficient CS-based framework for raw EEG signal acquisition and reconstruction.

The trade-off between CR and classification accuracy is addressed in the chapter. At the same time, the major components of research in this area are exposed, including physiological sensing and data preprocessing using CS technique. More details are provided in Chapter 2. Noisy wireless communication by adding different SNR values on the compressed data, DCT, feature extraction using DWT, all detailed are in Chapter 3.

Finally, classification accuracy of the EEG-based epileptic seizure application is also being considered. Accordingly, this research work is intended to compress the raw EEG-epileptic seizure of 4096 samples using CS and DCT methods before to sending/transmitting it to the other side. The compression process has been developed at different values starting from 100s/s up to 1100s/s to investigate at which compressed value can get a better accuracy. Extensive experiments have made on each compressed value at a certain measure of compression ratio. It has been found that the best classification accuracy is reached over 85% at CR = 85.35%, which is equivalent to 600 s/s, where the best compression ratio is get a better accuracy over 80% [74, 75]. Finally, the proposed framework compresses the raw EEG data, transmits it over the wireless channel, showing the effect of channel impairments on the compression requirements to achieve target application accuracies.

Ensemble methods consist of a group of individually trained classifiers with combined prediction when classifying different instances. Most research reported in the literature [130, 41, 85, and 87] has shown that an ensemble classifier is more accurate than any individual classifier in the ensemble. Therefore, the proposed ensemble method, which is Noise-aware Signal Combination method, will also be addressed.

In this research, considering that, the EEG signal is in nature bandwidth hungry, several works have considered in-network processing for either compressing EEG data [131] or transferring EEG features instead of delivering the raw uncompressed signal [114]. Another reason considering that the sensor is battery operated, if the data is transmitted without compression, the battery power will be consumed faster.

Therefore, we propose unified framework where the EEG data is compressed using compressive sensing (CS) and sent using two different types of channels. In the first, it was sent over a noiseless channel while the second was sent over the Additive White Gaussian Noise (AWGN) wireless channel in three different cases where SNR=1, 5, 10 dB, more information in Chapter 3. On the other hand, the compressed data was reconstructed and statistical features were extracted. To address this scenario, a unified framework has been designed, which presents a compressive sensing-based technique to send compressed EEG data over AWGN wireless channel, reconstruction, and feature extraction using time-frequency domain analysis in preparation of data classification. Such framework makes this work more practical because it performs classification considering data imperfection due to compression and wireless channel transmission.

4.2 EEG-Epileptic Seizure Framework

The system model consists of two main parts, the transmitter and receiver. The following section presents the background knowledge related to EEG signals and introduces some terminologies and related information about its characteristics.

4.2.1 Overall System Model

The proposed system model has four phases, namely 1) compressing the raw EEG, 2) reconstructing the compressed EEG, 3) cross-validation and training stage for process, the obtained optimal values, and 4) testing stage, respectively.

There are two main parts representing the system model, the transmitter and the receiver. Typically, transmitters are mobile devices, equipped with battery sources; hence, huge size of data at the transmitter is the most critical. Therefore, using EEG-epileptic seizure raw data, CS and DCT methods have been utilized to reduce the size of EEG data before transmission, since CS has low complexity for down sampling at the transmitter based on the CR, which comes with the cost of higher complexity at the receiver side [19].

Figure 4.1 shows the framework that has compressive sensing and data reconstruction as well as the classification processes for EEG-based epileptic seizure [19, 27].

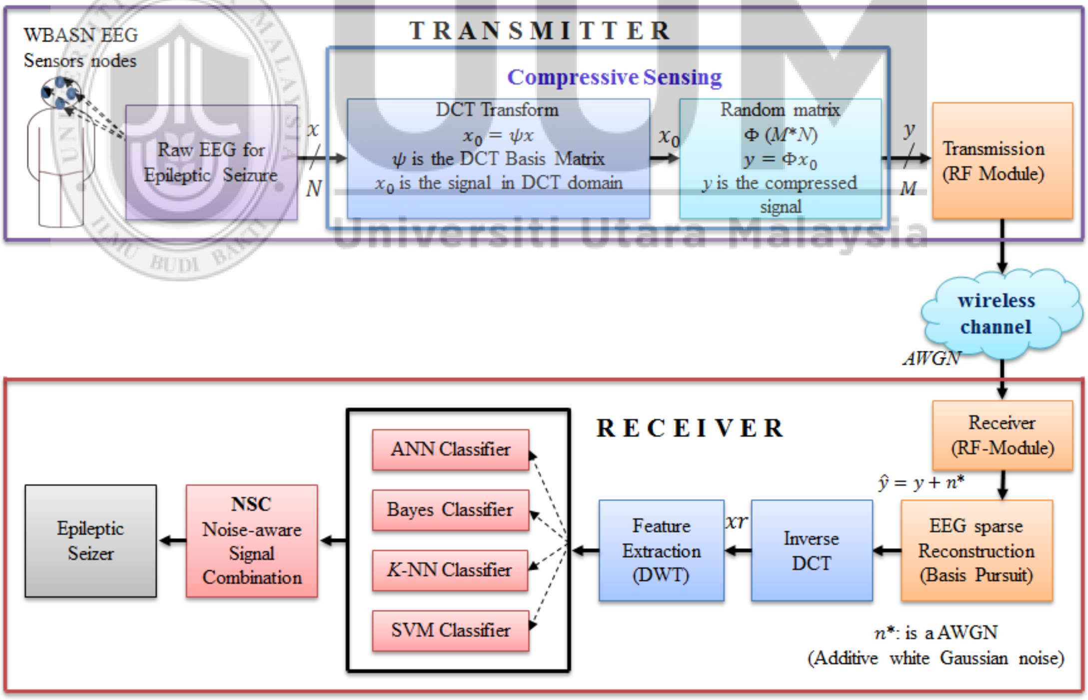


Figure 4.1. EEG-based epileptic seizure compression framework

The above system model consists of two main parts, the transmitter and receiver. The transmitter has *N*=4096 samples raw EEG represented by (*x*), and uses a CS technique to down sample the data based on sparse measurement matrix. In this framework, DCT and basis ψ have been used for different quantities of *M,* to get the compressed data $\hat{x}$ that will be transmitted over noiseless and noisy channels (i.e., radio frequency or Bluetooth). The measurement matrix size is *M* * *N* with random content and it can be represented by [128]:

$$y_{[M*1]} = \Phi_{[M*N]} \times x_{[N*1]} \tag{4.1}$$

The CS base uses DCT, which is one of the wavelet families. On the other hand, for transmitting the same data on noisy wireless channel, an Additive White Gaussian Noise (AWGN) to enforce SNR with different values has been added, such as 1 dB, 5 dB and 10 dB. Extensive experiments have been conducted to find SNR range in order to be added to the compressed EEG data that will be used to evaluate the proposed framework and ensemble NSC technique [128]. Compress the raw EEG data of size *N* samples to a size *M* < N samples using a CS technique to compose the transmitted data denoted by *y*. This data is transmitted over an AWGN channel.

At the receiver side, the received data have been reconstructed using iDCT. Then, statistical features of the data have been extracted using DWT. Next, the extracted features have been fed to four different classifiers, namely, ANN, NB, *k*-NN, and SVM, to evaluate the classification accuracy. The calculation of the classification accuracy is done independently for each classifier.

While the receiver, which receives the compressed signal M size, reconstructs back the EEG data using inverse DCT (iDCT) and basis pursuit to obtain the reconstructed signal (x_r). The iDCT reconstruction algorithm is for the DCT or an optimization problem with certain constraints is solved for the CS [122, 132, and 133].

4.2.2 Classification Methods

EEG detection and classification plays an essential role in the timely diagnosis and analyzes potentially fatal and chronic diseases proactively in clinical as well as various life settings [11]. In this research work, four different classification methods have been used, namely; ANN, NB, *k*-NN, and SVM. Initially, the classifiers have been developed to work individually for performance comparison. However, a data fusion method has been developed for combining the output of all classifiers in order to reduce the effect of data imperfection, while maximizing classification accuracy. Each classifier belongs to a different family of classifiers and has been shown to be the best classifier in its family. However, since each classifier uses a different classification strategy, it is expected that each one may yield different classification results [112, 111, and 97].

ANNs are a mathematical model that is motivated by the structure and functional aspects of biological neural networks. Learning in ANNs is carried out via specific training algorithms that are supposed to simulate the learning procedures of biological systems. To establish a classification rule and perform statistical analysis, ANN is able to estimate the posterior probabilities [97].

NB classifier is a statistical classifier. It has been demonstrated to be effective in many practical applications, including text classification and performance management systems. The NB approach main advantage comes from the learning procedure computational efficiency, which shows a linear computational complexity [111].

This method implies an intensive work when given large training sets and widely used when computer power is important [78]. The *k*-NN classifiers are based on learning by comparing a given test class with training classes that are similar to it. It can be defined by a distance metric called normalized Euclidean distance between two vectors as given two points Y_1 and Y_2 [111], more details in Chapter 2.

SVM uses nonlinear mapping to transform the original data into a higher dimension to construct a linear optimal separating hyperplane. For each given input, SVM takes a set of input data and predicts for two possible classes involved the input. SVM found that a compact description of learning model could be used for prediction to get high classification accuracy [111]. However, SVM applies decision combination to practice multi-classes classification since it is a binary classification.

The classification techniques reported above provide sufficient performance given that the EEG data are not contaminated by different factors. However, wireless EEG data are often compressed before transmission, which means that on the receiver side, some important information may get lost during the reconstruction process.

Moreover, a wireless channel may augment the transmission problem by adding noise artifacts to the transmitted data. Hence, a prospective classification technique should take into consideration the uncertainty problem contained in the EEG data to be efficient.

4.3 Ensemble Classification Methods

Several combination techniques have been introduced in the literature and each one may offer certain advantages and suffer from certain limitations. However, given several classifiers, the combination method has to deal with two critical issues: the dependency among the potentially combined classifiers, and the consistency of information each classifier provides.

For the first issue, since each classifier consider as a source of information, the classifiers have to be independent. This means that each classifier simply works on the input feature set independently, meaning in parallel, while the classification is based on combining the outcomes of all classifiers simultaneously. In this work, the classifiers to be independent will be considered.

For the second issue, as different classifiers are expected to pose different viewpoints of the current system state, classifiers may have conflicting decisions. To compromise this anticipated conflict, an effective mechanism that is capable of quantifying the assurance of each classifier in their decision is desirable.

One of these well-known combination techniques is the majority vote. The MV rule technique collects the votes of all classifiers and investigates the class name that is mostly reported by the classifiers. It then chooses that class as a final decision [96]. However, MV is based on the idea that the classifiers participating in the voting process have the same weight. It completely ignores the inconsistency that may arise among the classifiers. This of course can cause less capable classifiers to override classifiers that are more capable. Thus, the performance of the classification system can be deteriorated. Since the classifier models proposed in this work are expected to have different discriminant weight, the MV technique is not suitable as a combination method, because each classifier has the same weight and the weak classifier will affect the powerful classifier.

On the other hand, in the probability-based voting schemes, the weighted combination method should assign a probability value (p) that reflects the classifier confidence in its viewpoint. One of these schemes can be based on an accumulated experience. For instance, a given classifier is correct in identifying a certain hypothesis 75% percent of the time, while another classifier can correctly identify a different hypothesis 30% of the time. These values can actually be interpreted as probability assignments.

If the classifiers happen to yield these different and conflicting hypotheses as an explanation to the current system state, then classifiers should not be treated equally at the classification stage. Clearly, the first classifier is more confident in its decision than the first one. This valuable information should be incorporated in the combination process.

Assign a weight (p) of 0.75 to the first classifier, while assigning only 0.30 as a weight to the second classifier.

Let T be the set of classifiers,

$$T= \{C_1, C_2, \ldots ,C_n\}, \tag{4.1}$$

and C be the set of classes,

$$\{p_1 , p_2 , \ldots, p_n\},$$

And let $d_{i,j}$ be the decision of classifier i, and defined as follows:

$$d_{\boldsymbol{i,j}} \in \{0,1\}, \tag{4.3}$$

where $i = 1, \ldots , T$, and $j = 1, \ldots , C$.

Let p_i represent the weight of a classifier i, then the probability-based voting decision is calculated as:

$$\sum_{i=1}^{|T|} p_i d_{i,j} = \max_{j=1}^{C} \sum_{r=1}^{|T|} p_r \, d_{r,j}, \tag{4.4}$$

Considering the weight for each classifier, Equation (4.4) counts the votes coming from the participating classifiers.

4.4 Proposed Ensemble System Model

The proposed model consists of three stages for detecting electroencephalogram seizures namely statistical feature extraction, classifier prediction, and the proposed noise-aware signal combination (NSC) method. Statistical features extractions were discussed in Chapter 2, Section 2.5.2. As for classifier prediction, four popular classifiers are utilized in this model. These classification methods are trained using the most popular data mining tools, which is an industry standard and widely used for research [134]. The training process is conducted on similar data adhering to various combinations of SNRs and down sampling rates.

After exhaustive iterated experiments, the trained models are saved and their averaged performances in different scenarios report to the NSC. The NSC is our proposed ensemble method using combinations of probability estimates. Eventually, the ultimate classification accuracy is obtained through the epileptic seizure detection.

Having s tabular observations $O = \{o_0, o_1, \dots, o_{s-1}\}$ where, each o_i is a t_tuple of readings $R_i = (r_0, r_1, \dots, r_{t-1})$.

These observations fall into m different categories of classes $= \{c_0, c_1, \dots, c_{m-1}\}$ with each of size $\frac{s}{m}$.

The DWT is applied to the set of observations O to obtain l_tuple of features $F_i = (f_0, f_1, \dots, f_{l-1})$ for each $o_i \in O$. In other words, $DWT: O \rightarrow F$ such that, $DWT(o_k) = (f_{k,0},\ f_{k,1}, \dots, f_{k,l})$ where, $f_{k,j}$ is an l_tuple extracted features for the observation o_k obtained by DWT. Figure 4.3, illustrates the feature extraction by DWT.

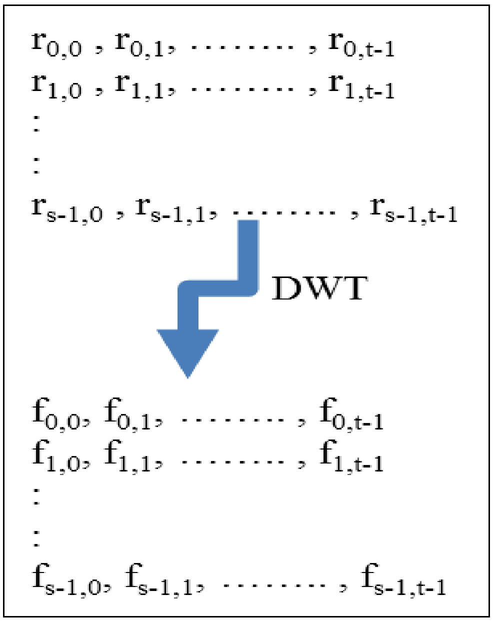


Figure 4.2. Feature Extraction using DWT

Hence, $DWT(O) = \{ (f_{i,0},\ f_{i,1}, \dots, f_{i,l}) | i = 0, 1, \dots, s-1 \}$ is the training and testing tabular, l_tuple format representing the input data for the classification model in this research work. The individual classifiers are supplied with extracted features using DWT. Each classifier uses its trained model to predict a class label for each set of features obtained from an individual observation in the testing data. The output of the classifier is a classifier confusion matrix (CCM) having important performance measures of interest such as class prediction label $PL_{i,j}$, class precision $PR_{i,j}$, class recall $RE_{i,j}$, and accuracy AC_j. These confusion matrices are supplied to the NSC and then transformed into normalized forms. The ensemble classifier will then use its own hypothesis to determine the predicted classes for each observation. The normalization of CCM and the hypothesis used by the NSC ensemble classifier are described in details in Section 4.4.2, while Figure 4.3 shows this general concept of the proposed model of NSC.

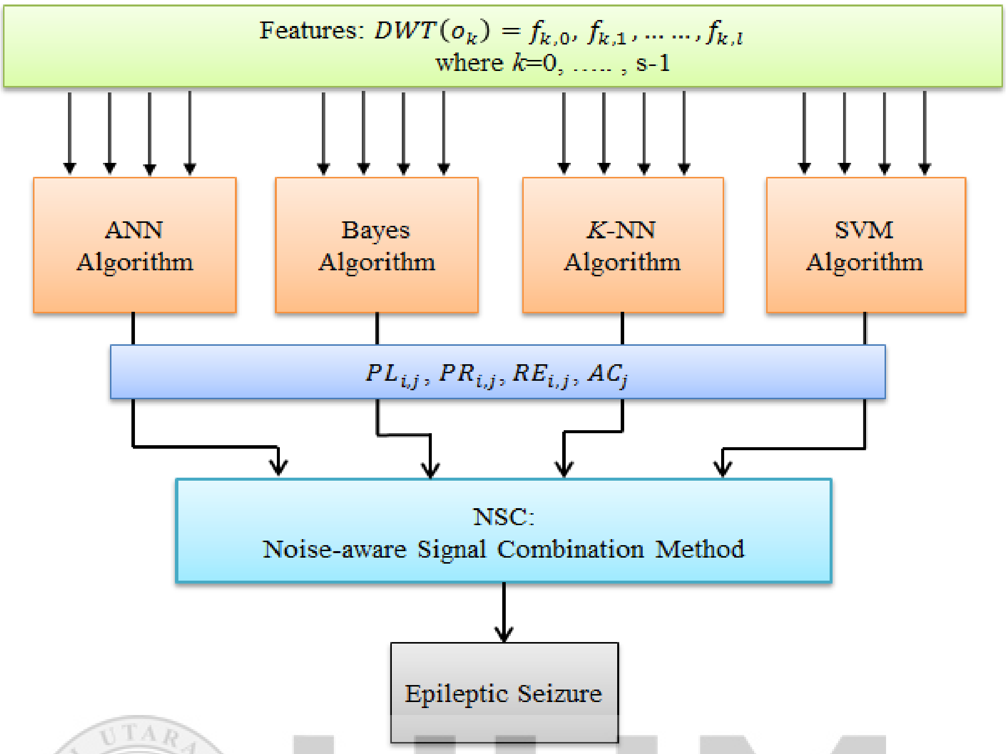


Figure 4.3. Proposed Ensemble Model

The following section describes the proposed Noise-aware Signal Combination ensemble method.

4.4.1 The NSC Proposed Ensemble Method

A number of classifiers (n) built on various hypotheses $H = \{h_0, h_1, \dots, h_{n-1}\}$, are fed with input data. The input data, $DWT(O)$, is represented in a tabular l_tuple format as discussed above. Each classifier k built on hypothesis h_k, is trained on the data aiming at predicting the label representing the class c_j, that best describes a given set of features $(f_{i,0},\ f_{i,1}, \dots, f_{i,l})$, corresponding to observation, o_i.

At the end of the training of each classifier, a set of multiclass classification performance measurements of interest is recorded. Table 4.1, represents some classifiers performance measurements. The trained model will then be saved to apply it on various categories of testing data. This process is replicated and repeated to represent an output that can be averaged to describe model behavior on long run times.

Table 4.1:

Classifier's Performance Measurements

Measure	Description
$PL_{i,j}$	Predicted label of o_i using hypothesis h_j.
$PC_{i,j}$	Confidence value predicting c_i using hypothesis h_j.
$PR_{i,j}$	Precision value of c_i using hypothesis h_j.
$RE_{i,j}$	Class recall value of c_i using hypothesis h_j.
AC_j	Accuracy value of applying hypothesis h_j.

The proposed ensemble classification method is fed with the output of the n trained classifiers. The training data is bundled first in two parts and is used to train the n classifiers with the testing part of the bundle.

Finally, the classification decision of a testing sample is obtained by combining the decisions from the corresponding n classifiers at each layer using the Noise-aware Signal Combination method. A subset of the performance measures of each classifier together with the predicted class label $c \in C$ for an observation $o \in O$ provided by each classifier with hypothesis $h \in H$ are used to construct the confusion matrix for each classifier. These confusion matrices form the input to the hypothesis used by this combined classifier.

4.4.2 The Classifier Confusion Matrix

The classifier confusion matrix (CCM) represents multiclass classification performance, which is represented the outcomes of each classifier. It has information about the actual and predicted classification accuracy reported by the classifier. CCM is a common tool used to estimate the goodness approaches of a classifier in the unbalanced classes' scenario. A raw confusion matrix is a square matrix that represents the count of a classifier's class prediction with respect to the actual outcome on some labeled learning set [135].

Classifier performance is usually evaluated by data in the matrix, table view for each hypothesis h_k, based on the reported performance results of the trained hypothesis $\mathrm{h_k}$, is calculated using the algorithm shown in Figure 4.6. An entry $\mathrm{CCM}[i][j]$ in the matrix of reported performance results for hypothesis h_k, represents the frequency of predicting class j as of being class i. Therefore, CCM $[i][j]$ represents the frequency of correct predictions being in class i, while $\sum_{j\neq i}^{m} \mathrm{CCM}\ [i][j]$ is the frequency of wrong predictions of other classes as of being class i.

Hence,

- PR_i represents the class precision of class i calculated by the equation:

$$\frac{\mathrm{CCM}[i][i]}{\sum_{j\neq i}^{m} \mathrm{CCM}[i][j]} * 100\% \qquad (4.5)$$

- RE_i, represents the class recall of class j calculated by the equation:

$$\frac{\mathrm{CCM}[i][i]}{\sum_{i\neq j}^{m} \mathrm{CCM}[i][j]} * 100\% \qquad (4.6)$$

Finally;

- *AC_i* represents the overall classifier accuracy, using hypothesis h_k, calculated by the averaged classes' precisions given by:

$$\frac{\sum_i^m PR_i}{m} * 100\% \tag{4.7}$$

For example, it is assumed that there are 300 observations of three different types of EEG classes, each EEG class has 100 observations, and the classes are labeled by 1, 2, and 3 respectively. Table 4.2: illustrates the confusion matrix or multiclass classification performance or performance vector for a classifier of three classes.

Table 4.2:

Multiclass classification performance/Performance vector

Accuracy: 83.00%

	true 1	true 2	true 3	class precision
pred. 1	89	39	0	69.53%
pred. 2	11	60	0	84.51%
pred. 3	0	1	100	99.01%
class recall	89.00%	60.00%	100.00%	

To read the above confusion matrix table (classifier outcomes), let us start with column 2, that is labeled by (true 1) with the three predictions rows (pred.1, pred.2, and pred.3). The classifier predicted class 1, by 89 times out of 100 which is called true positive (TP), and this prediction are true by 89 times out of 100, also, the same classifier predicted the rest 11 to be class 2, which is called false positive (FP), this 11 out of 100 is false. Finally, the same classifier predicted 0 of class 3, this was false, and this is true. Similarly, the values for class 2 and class 3 have the same meaning.

On the other hand, from the above table there are three values namely; class precision which is known as specificity, class recall also known as a sensitivity, and last one is accuracy of classifier for the three classes. A high-class recall means the classifier returned most of the important results, while a high-class precision means the classifier returned significantly more important results than non-important. A detailed information about the calculation and equations of the three values are illustrated in Section 2.6.2.

The three values are calculated as:

- Class precision (PR_i) /Specificity $= \frac{TP}{TP+FP}\% = \frac{89}{89+39+0}\% = 69.53\%$=0.89%
- Class recall (RE_i) /Sensitivity $= \frac{TP}{TP+FN}\% = \frac{89}{89+11+0}\% = 89.00\%$
- Accuracy (AC_i) = $\frac{TP + TN}{TP + TN + FP + FN}$

  $= \frac{(89+60+100)+(0+0)}{(89+60+100)+(0+0)+(39+1)+(11+0)}$

  $= \frac{249}{300} \times 100 = 0.83 \times 100 = 83.00\%$

Similarly, the calculation of the other two classes is done in the same manner.

A normalized form of CCM across all hypotheses is obtained to be used by the ensemble classifier, namely, $\overline{CCM}$. In this normalized confusion matrix, an entry $\overline{\mathrm{CCM}}[i][j]$ is the normalized entry of $\mathrm{CCM}\ [i][j]$ on class *i* recall with respect to all hypotheses. That is:

$$\overline{CCM}[i][j] = \frac{CCM\ [i][i]}{\sum_i^n CCM\ [i][j]} \tag{4.8}$$

Therefore, $\sum_i^n \overline{\mathrm{CCM}}[i][j] = 1$.

Hence,

- The normalized PR_i, across the set of hypotheses *H* is given by:

$$\overline{PR}_i = \frac{PR_i}{\sum_j^n PR_i} \qquad \text{therefore,} \quad \sum_i^n \overline{PR}_i = 1 \tag{4.9}$$

- The normalized RE_i and the normalized *AC_i* across the set of hypotheses *H,* are,

$$\overline{RE}_i = \frac{RE_i}{\sum_j^n RE_i}, \qquad \text{therefore,} \quad \sum_i^n \overline{RE}_i = 1 \tag{4.10}$$

$$\overline{AC}_i = \frac{AC_i}{\sum_j^n AC_i}, \qquad \text{therefore,} \quad \sum_i^n \overline{AC}_i = 1 \tag{4.11}$$

The following example demonstrates how to calculate the normalized CCM from the raw classifier confusion matrix across all hypotheses at the target compression ratio of 85.35%, which resulted from extensive experiments, which is equal to 600s/s with noisy data of SNR=1 dB. To calculate the normalization of class precision (PR_i), class recall (RE_i), and classifier accuracy (AC_i), a raw confusion matrix is the following:

True class label

Predicted class label

	true 1	true 2	true 3	PR
pred. 1	80	38	4	0.66
pred. 2	20	61	2	0.73
pred. 3	0	1	94	0.99
RE	0.80	0.61	0.94	

ANN Accuracy 0.78

True class label

Predicted class label

	true 1	true 2	true 3	PR
pred. 1	80	40	0	0.67
pred. 2	20	59	0	0.75
pred. 3	0	1	96	0.99
RE	0.80	0.59	0.96	

Bayes Accuracy 0.78

True class label

Predicted class label

	true 1	true 2	true 3	PR
pred. 1	92	46	0	0.67
pred. 2	8	53	7	0.78
pred. 3	0	1	93	0.99
RE	0.92	0.53	0.93	

k -NN Accuracy 0.79

True class label

Predicted class label

	true 1	true 2	true 3	PR
pred. 1	81	30	5	0.70
pred. 2	19	70	13	0.69
pred. 3	0	0	82	1.00
RE	0.81	0.70	0.82	

SVM Accuracy 0.78

Figure 4.4. Raw Confusion Matrix for CR=85.35% at SNR=1 dB

Calculate each normalized class precision (*PR$_i$*) by dividing it on the sum of all *PR$_i$* for all hypotheses, for each corresponding class.

- Class 1 (c_0) = 0.66 + 0.67 + 0.67 + 0.70 = 2.70.
- Class 2 (c_1) = 0.73 + 0.75 + 0.78 + 0.69 = 2.95
- Class 3 (c_2) = 0.99 + 0.99 + 0.99 + 1.0 = 3.97

Class recall (*RE$_i$*) calculation is done in the same manners of (*PR$_i$*),:

- Class 1 (c_0) = 0.80 + 0.80 + 0.92 + 0.81 = 3.33.
- Class 2 (c_1) = 0.61 + 0.59 + 0.53 + 0.70 = 2.43
- Class 3 (c_2) = 0.94 + 0.96 + 0.93 + 0.82 = 3.65

Finally, calculate the of overall classifier accuracy (*AC$_i$*) for all classifiers:

$$\sum AC_i = 0.78 + 0.78 + 0.79 + 0.78 = 3.13$$

The normalization for each parameter has been calculated below:

- Normalized (*PR$_i$*) for all classes, namely, c_0, c_1, c_2, as:

Class 1 ($c_{0)}$ = 0.66/2.70= 0.244 for ANN, 0.67/2.70 = 0.248 for NB,

= 0.67/2.70= 0.248 for k-NN, and 0.70/2.70= 0.259 for SVM.

Simply, the same normalized calculation for the other classes to calculate resting of the *PR$_i$* has been done, and to calculate the all *RE$_i$*, as well as *AC$_i$*.

True class label

Predicted class label

	c$_0$	c$_1$	c$_2$	$\overline{PR}$
c$_0$	80	38	4	0.245
c$_1$	20	61	2	0.248
c$_2$	0	1	94	0.249
$\overline{RE}$	0.240	0.251	0.258	

h_0 =ANN $\overline{AC_0}$ 0.249

True class label

Predicted class label

	c$_0$	c$_1$	c$_2$	$\overline{PR}$
c$_0$	80	40	0	0.248
c$_1$	20	59	0	0.254
c$_2$	0	1	96	0.249
$\overline{RE}$	0.240	0.243	0.263	

h_1 =NB $\overline{AC_0}$ 0.249

True class label

Predicted class label

	c$_0$	c$_1$	c$_2$	$\overline{PR}$
c$_0$	92	46	0	0.248
c$_1$	8	53	7	0.265
c$_2$	0	1	93	0.249
$\overline{RE}$	0.276	0.218	0.255	

h_2 =k-NN $\overline{AC_0}$ 0.253

True class label

Predicted class label

	c$_0$	c$_1$	c$_2$	$\overline{PR}$
c$_0$	81	30	5	0.259
c$_1$	19	70	13	0.233
c$_2$	0	0	82	0.252
$\overline{RE}$	0.243	0.288	0.225	

h_3 =SVM $\overline{AC_0}$ 0.248

Figure 4.5. Normalized Confusion Matrix ($\overline{CCM}$), for CR=85.35% at SNR=1 dB

The normalized forms of important performance measures such as, PR_i, RE_i, and AC_i describe the goodness of these measures with respect to corresponding measures in other classifiers. Being normalized that way; such values represent a probability space for each performance measure.

The ensemble classifier NSC considers the probabilities corresponding to these performance measures for each hypothesis and uses it to calculate a probability of selecting a hypothesis reporting a class *k*. Therefore, the reported class *k* by a hypothesis is a fourth parameter considered in NSC decision. Since the calculated values of class precision and class recall are dependent, they have been grouped as one term in the equation leaving the overall accuracy as second term. The two terms have direct proportion to the goodness of selecting a classifier predicting class *k*. However, their effect is not necessary to be equal. Hence, a weight $0 < \sigma < 1$ to control this effect has been introduced. This weight is a long run tuned with Equation 4.12 for optimal performance on EEG-based epileptic seizure data.

The prediction of the NSC combination classifier is calculated by the following hypothesis with the highest probability defined as:

$$P(h_j) = \frac{\sigma \times \left(\overline{PR}_{k,j} + \overline{RE}_{k,j}\right) + (1-\sigma)\,\overline{AC}_j}{\sum_{i=0}^{n-1} \sigma \times \left(\overline{PR}_{k,i} + \overline{RE}_{k,i}\right) + (1-\sigma)\,\overline{AC}_i} \tag{4.12}$$

where, *k* is the label of predicted class and $\sum_{j=0}^{n-1} P(h_j) = 1$

Equation 4.12 is used by NSC to calculate the goodness probability of a hypothesis *j* predicting class *k* by dividing the summation of the *σ* weighted terms, described above, by the total of goodness for all hypotheses if they were to predict the same class *k*. The resultant value reported the goodness of hypothesis *j* reporting class *k* with respect to other hypotheses reporting class *k*.

NSC calculates the probabilities describing the goodness of each hypothesis corresponding to its predicted class *k* in the same manner. Finally, it confirms the predicted class of the hypothesis having the highest probability.

```
Pseudo code: NSC
Preliminaries
  Let $O = \{o_0, o_1, \dots, o_{s-1}\}$ be the set of observations
  Let $C = \{c_0, c_1, \dots, c_{m-1}\}$ be the set of class labels
  Let $H = \{h_0, h_1, \dots, h_{n-1}\}$ be the set of hypotheses
  Let $h_n$ be the hypothesis of the combined classifier
  Let $PL_{i,j}$, $PR_{i,j}$, $RE_{i,j}$, $AC_j$ be the $j$ predicted class label, class
  precision, class recall, and accuracy of $h_i$

PROCESS
  $\forall h_i | i = 0, \dots, n-1$
   $\overline{PR}_{i,j} = \frac{PR_{i,j}}{\sum_{j=0}^{n-1} PR_{i,j}}$,   $\overline{RE}_{i,j} = \frac{RE_{i,j}}{\sum_{j=0}^{n-1} RE_{i,j}}$,   $\overline{AC}_j = \frac{AC_j}{\sum_{j=0}^{n-1} AC_j}$
  $if \exists O$ /** on receiving a periodical batch of observations*/
    $\forall o_i | \ i = 0, \dots, s-1$
      $W = 0$
    $\forall h_j | j = 0, \dots, m-1$
      $k = PL_{i,j} | k \in C$
      $w(h_j) = \sigma(\overline{PR}_{k,j} + \overline{RE}_{k,j}) + (1 - \sigma) \overline{AC}_j$
      $W += w(h_j)$
    $\forall h_i | i = 0, \dots, n-1$
      $P(h_i) = \frac{w(h_i)}{W}$
      $PL_{i,n} = Max(P(h_i)|_{i=0}^{n-1})$ // Record prediction
  Calculate $h_n$ Performance measurements of interest
```

Figure 4.6. Noise-aware Signal Combination Pseudo code

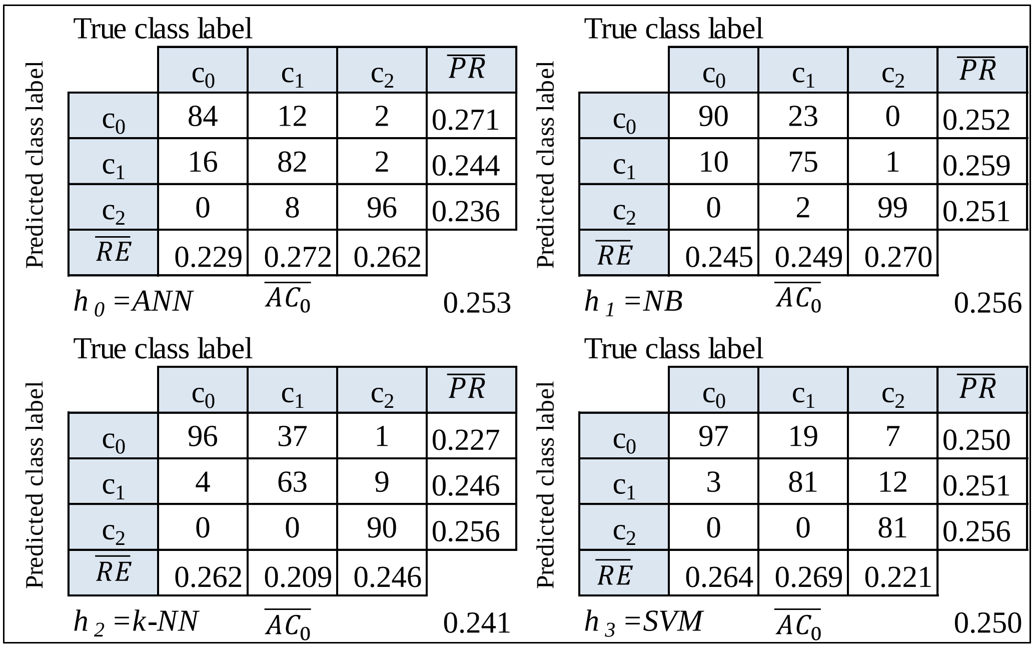
The classification results of the individual classifier are displayed in a confusion matrix. In the confusion matrix, each cell consists of numbers of vectors classified for the corresponding combinations of (true class & predicted class labels) outputs.

Figures 4.7-4.10 show the confusion matrices for the four classifiers, which have been selected before. These matrices represent the finalized weighted performance of the trained classifiers based on noiseless data and three different levels of data noisy, SNR= 1 dB, 5 dB, and 10 dB for EEG-based epileptic seizure at $M = 600$, down sample value. Also, these figures show that c_0, c_1, and c_2 are representing class A, class C, and class E, respectively.

For example, Figure 4.7 represents EEG noiseless data, classes' c_0, c_1, and c_2 in vertical line are representing the predicted class label; on the other hand, in the horizontal line, the true class label has shown. The normalized precision $\overline{PR}$, of class A in the first row of the four matrices is 0.271, 0.252, 0.227, and 0.250 for ANN, NB, k-NN, and SVM respectively. The normalized class recall $\overline{RE}$ of class A in first four matrices is 0.229, 0.245, 0.262, and 0.264 for ANN, NB, k-NN, and SVM, respectively. Furthermore, the normalized overall accuracy $\overline{AC}$, is 0.253, 0.256, 0.241, and 0.250 for the same set of classifiers respectively.

Also, shown from these figures, is the confusion matrix the normalization of class precision, class recall, and the overall classification accuracy for each classifier. Also, in Figure 4.10, you can read one table as rows by saying that the classifier prediction for class 0 (c_0) is 88%, while classifier prediction for class1 (c_1) is 12% and classifier prediction for class2 (c_2) is 0%. This means that the 88% is the overall classification accuracy for c_0, and so on for the other classes.

According to the confusion matrix, there are many mis-classifications - occurred using the utilized four classifiers and the proposed method.



True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	84	12	2	0.271
c_1	16	82	2	0.244
c_2	0	8	96	0.236
$\overline{RE}$	0.229	0.272	0.262	

h_0 =ANN $\overline{AC}_0$ 0.253

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	90	23	0	0.252
c_1	10	75	1	0.259
c_2	0	2	99	0.251
$\overline{RE}$	0.245	0.249	0.270	

h_1 =NB $\overline{AC}_0$ 0.256

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	96	37	1	0.227
c_1	4	63	9	0.246
c_2	0	0	90	0.256
$\overline{RE}$	0.262	0.209	0.246	

h_2 =k-NN $\overline{AC}_0$ 0.241

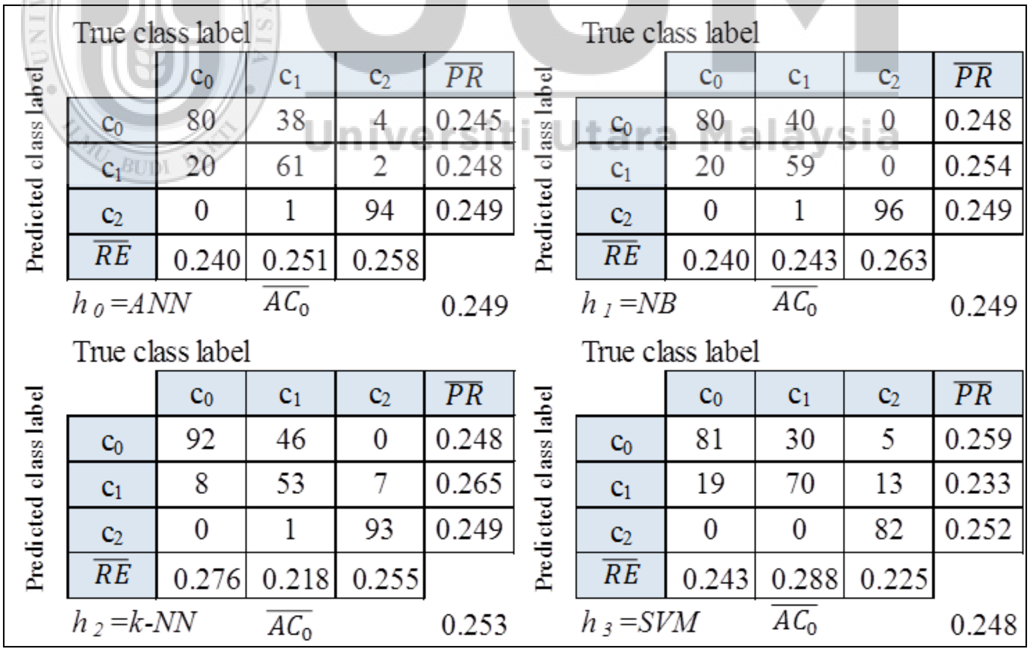
True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	97	19	7	0.250
c_1	3	81	12	0.251
c_2	0	0	81	0.256
$\overline{RE}$	0.264	0.269	0.221	

h_3 =SVM $\overline{AC}_0$ 0.250

Figure 4.7. Confusion matrix for noiseless data



True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	80	38	4	0.245
c_1	20	61	2	0.248
c_2	0	1	94	0.249
$\overline{RE}$	0.240	0.251	0.258	

h_0 =ANN $\overline{AC}_0$ 0.249

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	80	40	0	0.248
c_1	20	59	0	0.254
c_2	0	1	96	0.249
$\overline{RE}$	0.240	0.243	0.263	

h_1 =NB $\overline{AC}_0$ 0.249

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	92	46	0	0.248
c_1	8	53	7	0.265
c_2	0	1	93	0.249
$\overline{RE}$	0.276	0.218	0.255	

h_2 =k-NN $\overline{AC}_0$ 0.253

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	81	30	5	0.259
c_1	19	70	13	0.233
c_2	0	0	82	0.252
$\overline{RE}$	0.243	0.288	0.225	

h_3 =SVM $\overline{AC}_0$ 0.248

Figure 4.8. Confusion matrix for noisy data (SNR=1 dB)

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	84	31	2	0.262
c_1	16	68	3	0.279
c_2	0	1	95	0.249
$\overline{RE}$	0.240	0.264	0.262	

$h_0 = ANN$ $\overline{AC_0}$ 0.254

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	88	36	0	0.251
c_1	12	62	4	0.249
c_2	0	2	96	0.247
$\overline{RE}$	0.251	0.240	0.265	

$h_1 = NB$ $\overline{AC_0}$ 0.254

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	91	50	0	0.228
c_1	9	50	11	0.225
c_2	0	0	89	0.252
$\overline{RE}$	0.260	0.194	0.246	

$h_2 = k\text{-}NN$ $\overline{AC_0}$ 0.238

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	87	22	10	0.259
c_1	13	78	8	0.246
c_2	0	0	82	0.252
$\overline{RE}$	0.249	0.302	0.227	

$h_3 = SVM$ $\overline{AC_0}$ 0.254

Figure 4.9. Confusion matrix for noisy data (SNR=5 dB)

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	88	19	3	0.264
c_1	12	75	1	0.255
c_2	0	2	96	0.248
$\overline{RE}$	0.237	0.265	0.266	

$h_0 = ANN$ $\overline{AC_0}$ 0.254

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	90	28	0	0.251
c_1	10	69	4	0.249
c_2	0	3	96	0.246
$\overline{RE}$	0.242	0.244	0.266	

$h_1 = NB$ $\overline{AC_0}$ 0.251

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	98	41	0	0.232
c_1	2	59	13	0.239
c_2	0	0	87	0.253
$\overline{RE}$	0.263	0.208	0.241	

$h_2 = k\text{-}NN$ $\overline{AC_0}$ 0.240

True class label

Predicted class label

	c_0	c_1	c_2	$\overline{PR}$
c_0	96	20	9	0.253
c_1	4	80	9	0.258
c_2	0	0	82	0.253
$\overline{RE}$	0.258	0.283	0.227	

$h_3 = SVM$ $\overline{AC_0}$ 0.254

Figure 4.10. Confusion matrix for noisy data (SNR=10 dB)

At the end of each experiment, the algorithm calculates the performance for each classifier including the proposed ensemble classifier NSC based on the recorded test results. The NSC achieved the desired improved classification accuracy with noisy data: classification accuracies of 80% for SNR=1 dB, 84% for SNR=5 dB, and 88% for SNR=10 dB were obtained at CR =85.35%. Moreover, NSC may provide several demonstrable benefits, such as simplicity, and NSC improves the overall classification accuracy.

Compared with previous works, the classification accuracy by the proposed NSC of noiseless EEG data, is 90% achieved, which is 4.1% higher than the work done in [87] (85.9% accuracy), and 0.5% higher than the reported in [84] (89.5% accuracy), considering the same dataset.

In addition, to the best of our knowledge, no similar evaluation approaches for EEG-based epileptic seizure when considering AWGN with different SNR values.

4.5 Summary

In this chapter, in details the compression and classification framework for EEG-based epileptic seizure and methods perspectives have been explained. Classification accuracy and compression ratio have been employed to evaluate the performance of the utilized classifiers. A simulation program has been implemented for data compression, using an AWGN wireless model with different values of noise as part of this work. DWT is used to extract statistical features that represent values of all-important information about the original signal data. These features are calculated from each cross-correlation sequence to create feature vector sets. EEG classification methods have been utilized by using four classifiers. The four individual classifiers are ANN, NB, k-NN, and SVM. Another simulation program has been developed to demonstrate a proposed NSC ensemble technique to measure the noisy EEG-based epileptic seizure classification accuracy. The outputs of the above classifiers are used as input to the proposed combination method. The following parameters classifier prediction, class precision, class recall, confidence, and overall classifier accuracy have been selected. Finally, the confusion matrix has been used to represent each classifier vector performance in terms of the prediction and actual classification.

CHAPTER FIVE
RESULTS AND DISCUSSIONS

This chapter presents the experimental results of the two proposed approaches, namely, the unified sensor-based compression and classification (SCC) framework for data delivery and noise-aware signal combination (NSC) method. The EEG-based epileptic seizure dataset consists of three different classes that can distinguish the epilepsy and non-epilepsy subjects. In this study, all experimental results for the dataset are presented based on testing sets.

Three different experiments have been conducted to evaluate the proposed SCC EEG-based epileptic seizure framework. First, investigate the effect of compression of the EEG data at different values of signal-to-noise-ratio (SNR) at CR=85.35. Second, show the measure of classification accuracy for a noiseless as well as a noisy version of the compressed EEG data. Specifically, use SNR values equal to 1, 5 and 10 dB. We choose to simulate the performance of our proposed algorithm at the SNR levels presented earlier, which represents three levels of SNR, low: SNR = 1dB, moderate SNR =5 dB and acceptable to high SNR =10dB [75, 76], more informtion refer to Chapter 3. Low SNR level represents low data accurcacy, i.e., high noise effect. Conseqently, moderate SNR level represents med-level adat accuracy, and so forth for the case of acceptable of SNR. Third, measure the classification accuracy against the CR in all cases after adding three different noise levels on the transmitted compressed EEG data.

To do the second and third goals highlighted in the previous paragraph, four different classification methods were used. These methods are ANN, NB, *k*-NN and SVM. Initially, the classifiers were developed to work individually. A more detailed description about these four classifiers is presented in Chapter 2 and Chapter 4. As each classifier is working independently, an ensemble Noise-aware Signal Combination (NSC) method has been proposed to enhance the ensemble classification accuracy of EEG-based data.

The next section reports the obtained results and provides the illustrations and discussions relevant to each classifier. In addition to show, the classification accuracy of the proposed NSC compared to all other individual classifiers.

5.1 Classification Accuracy using Four Classifiers

Using a simulation platform, a framework that represents the transmitter and receiver of our proposed technique that combines several methods has been developed. On the transmitter side, the compressive sensing (CS) and DCT techniques were utilized. DCT uses a measurement matrix to compress the raw EEG-based epileptic seizure and send it over the wireless channel. On the other side, it received and reconstructed the compressed data back to the original size using iDCT. At the feature extraction stage, it extracted statistical features in excel sheets using DWT. These features are used to measure the data classification accuracy. The data mining tool is used to study the CS complexity-accuracy trade-off.

For each CR value, a series of ten experiments has been conducted, each of which was accomplished with an independent random measurement matrix Φ of size $M * N$.

Then, for each CR value, the classification accuracy has been computed as the average of the ten accuracies corresponding to the ten experiments performed for that CR value. Concerning execution time, for each classifier (ANN, NB, k-NN, SVM), all the CR values considered above were taken. Then, a series of classifications (ten experiments per CR value, all CR values per each classifier) were performed and the average running times for each classifier were measured.

The goal of the running time experiments is to decide which of the classifiers is more efficient in our context. On average, ANN took around 15 seconds; SVM took 1 second, while both of *k*-NN and Bayes took less than 1 second. These timings were obtained while the classifier parameters were configured to the default values. ANN training cycles were set to 500 and its learning rate set to 0.3. NB used Laplace correction to avoid high potential impact zero, so the default is set to true. The type of the *k-NN* is set to radial and parameter *k* was set to 10. Finally, SVM type was nu-support vector classification (nu-SVC) and the cache size was defined as 80 megabytes.

5.1.1 Cross-validation

Cross-validation is a nested operator and the most important part in the classification method. It is used to estimate a statistical performance of a learning operator. It has two sub-processes, namely training and testing. The training sub-process is used to provide a training model. Then, the resulting training model is applied in the testing sub-process to measure the model performance. A *k*-fold cross-validation is used to compute the classification accuracy, by first dividing the initial dataset into *k* subsets of approximately equal size, and performing the cross-validation on each of the subsets. In each execution, one of the subsets is used as a testing set and the remaining *k*-1 subsets are used as a training set. Using the number of cross-validations *k*, the average accuracy across all *k* iterations is computed.

The 10-fold cross-validation procedure has been used to evaluate the performance of the classifiers. Classification accuracies have been provided for four individual classifiers, namely ANN, NB, k-NN, and SVM for an EEG benchmark dataset. All experiments on the dataset are evaluated by 10-fold cross-validation process, which indicates the generalizability of the classifier and reliability of the obtained results. In contrast, in the proposed NSC method, only considered a 2-fold cross-validation procedure to reduce the computation time and the number of experiments. The reader shall note that the generalizability term refers to the ability of a model to perform well on unseen datasets.

5.1.2 Performance Evaluation Measures

Several types of methods for performance evaluation measures are presented. In this research study, the compression ratio and the overall classification accuracy are measured. The purpose is to evaluate the performance of the used individual classifiers. In the proposed technique, the parameters are represented by different statistical measures, such as classifier prediction, class precision or specificity, class recall or sensitivity, confidence, and overall classifier accuracy. The classification accuracy is evaluated based on the measurements of the highest probability of the hypothesis (Chapter 4, eq. 4.12). The sensitivity, specificity, class label and accuracy are defined as follows:

- *Specificity (PR_i)*: is the number of true negative decisions divided by the number of actual negative cases;
- *Sensitivity (RE_i)*: is the number of true positive decisions divided by the number of actual positive cases;
- *Classifier class label (PL_i)*: is the classifier prediction field of class labels, which contains the class labels of all classes during the historical classification.
- *Classification accuracy (AC_i)*: is the number of correct decisions divided by the total number of cases.

5.2 Experiments and Results

The following sections will focus on the experiments and results that were obtained in this research work.

5.2.1 Experiment 1: Effect of Compression on Noiseless Data

The performance of the proposed framework has been evaluated in two ways: (1) by dividing EEG epileptic seizure feature set into two groups, as a training and testing sets, and (2) by using the 10-fold cross-validation process. This indicates the classifier generalization and the reliability of the obtained results. The experiment results are shown in the following five graphs in Figures 5.1 - 5.5.

The coming Figure 5.1 illustrates the classification accuracy in function of the CR in the case of a noiseless wireless channel, using the four aforementioned classifiers. Figure 5.1 also shows the results for ANN, NB, *k*-NN, and SVM classifiers accuracy for the noiseless wireless channel. The results show that accuracy decreases logarithmically with the increase of CR. The results were divided into three main regions: at CR = 75% and CR=85%, respectively. Each point on the graphs, 10 experiments has been conducted and calculated the average of classification accuracy accordingly. While accuracy remains stable above 90% for all classifiers in the first region, ANN, and Bayes seem to have a better accuracy of about 5%.

The decay in the accuracy seems to be reasonable in the second region, showing that ANN outperforms all others in high compression values, but it decays exponentially in the third region for all classifiers. While ANN consistently outperforms the other three classifiers in all regions, the high classification time and complexity of implementation makes ANN prohibitive in real-time wirelessly tele-monitoring applications.

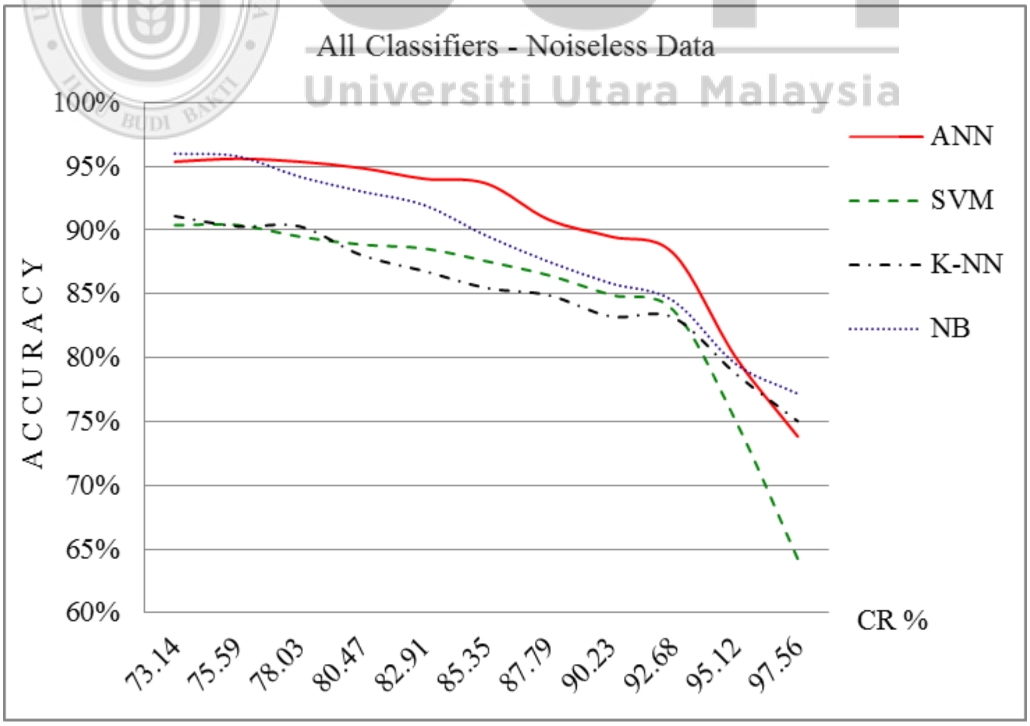


Figure 5.1. Classification accuracy in function of the CR, noiseless data

Hence, the following figures from Figures 5.2 – 5.5 show the results of classification accuracy in function of the CR for different SNR classifier values. These results include the effect of an additive white Gaussian noise (AWGN) channel model by different SNR values of 1 dB, 5 dB, and 10 dB on the transmitted compressed EEG data.

5.2.2 Experiment 2: Effect of Compression on Noisy data

The aim of this experiment was to show the accuracy of the effect of the compression on different EEG data classes. Each figure shows the effect on different data types, which are evaluated by each individual classifier.

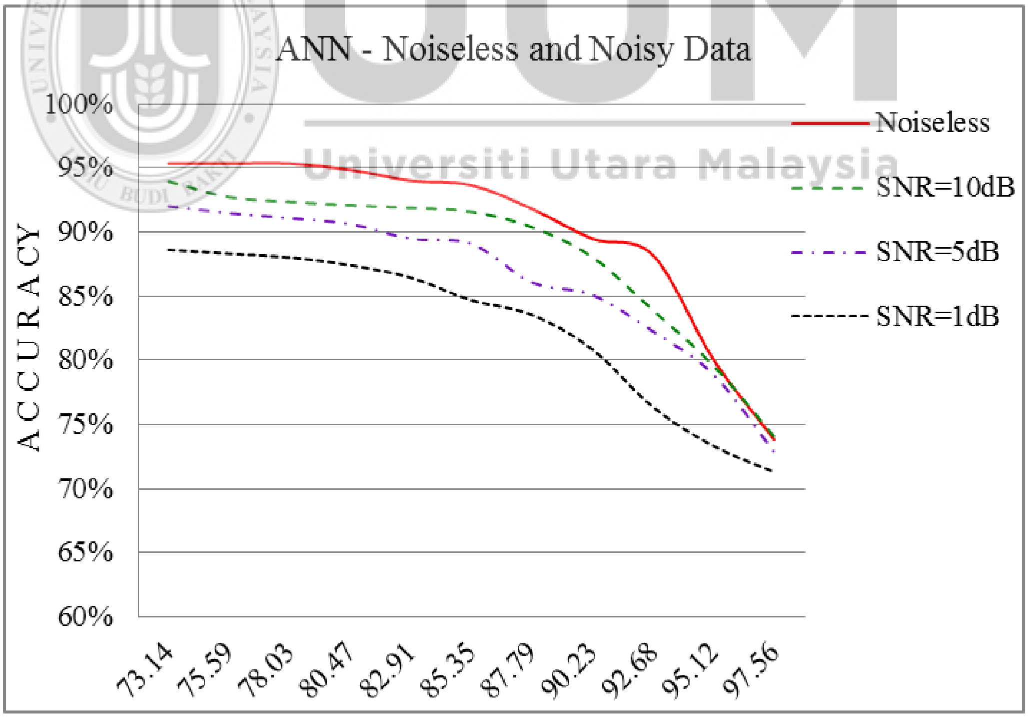


Figure 5.2. Classification accuracy using ANN classifier, noiseless and noisy data

Figure 5.2 corresponds to the use of the ANN classifier. The results show that at CR=85%, the accuracy starts to decrease regularly at the classification accuracy of 93.67% for noiseless data. Similarly, for noisy data, the classification accuracy was 91.60%, 89.10%, and 84.70%. The decay starts early with the increase of channel noise for SNR=10 dB, SNR=5 dB, and SNR=1 dB, respectively.

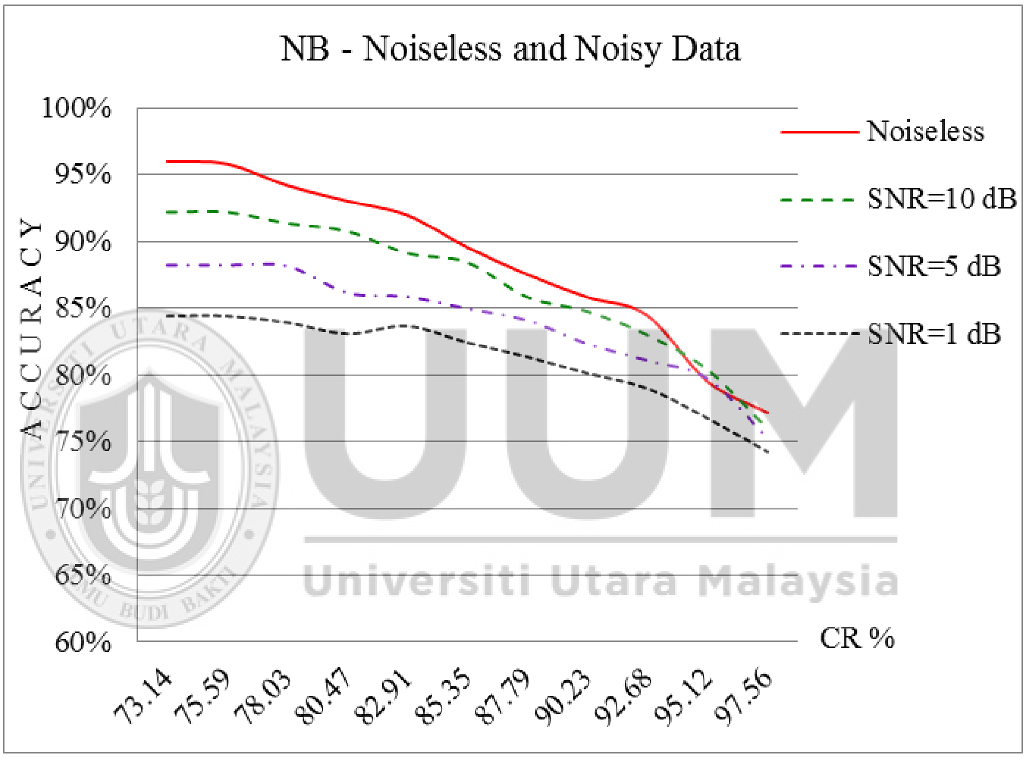


Figure 5.3. Classification accuracy using NB classifier for noisy data

In an analogous way, Figure 5.3 shows the usage of NB classifier. The results indicate that the classification accuracy decreases consistently, while the exponential decay starts early with the increase of channel noise. For example, the exponential decay starts at CR≈ 90% of the noiseless channel, while it starts at CR=85% when SNR=1dB.

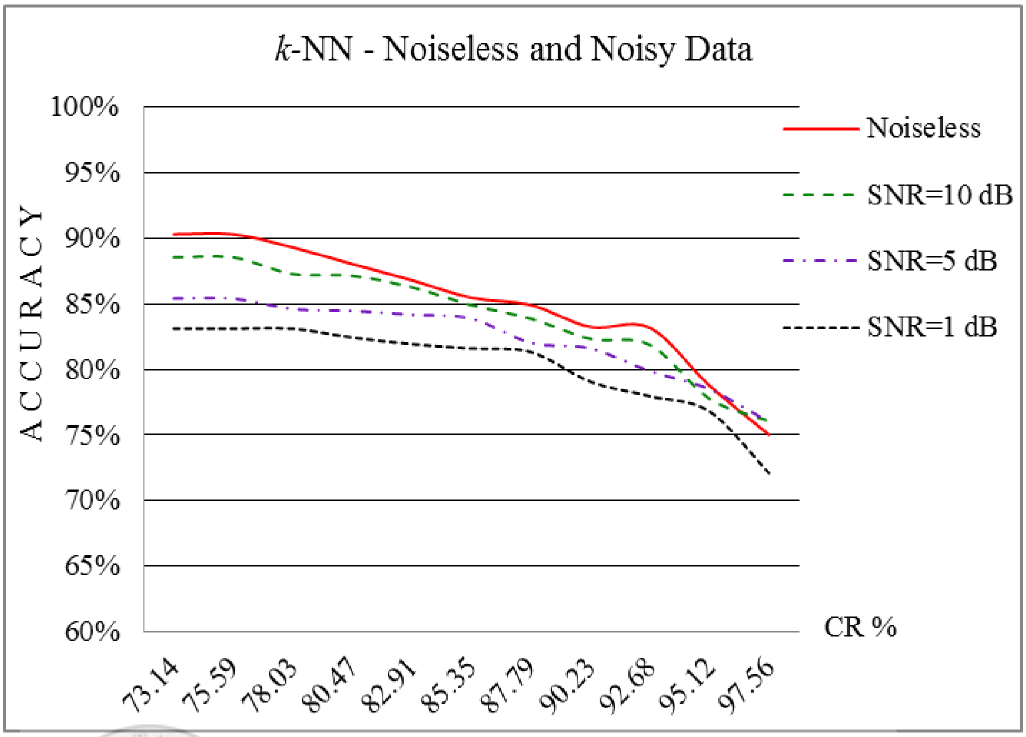


Figure 5.4. Classification accuracy using *k*-NN classifier for noisy data

Figure 5.4 shows *k*-NN classifier and the results indicate a slightly different behavior compared to Bayesian classifier. While the classification accuracy starts to decay linearly after CR=75%, the effect of noisy communication is more evident, causing the decrease of more than 10% when SNR=10dB. Eventually, Figure 5.5 shows steady decrease in function of both CR and SNR, nominating *k*-NN to be the best tolerable classifier to wireless channel noise and changes in CR.

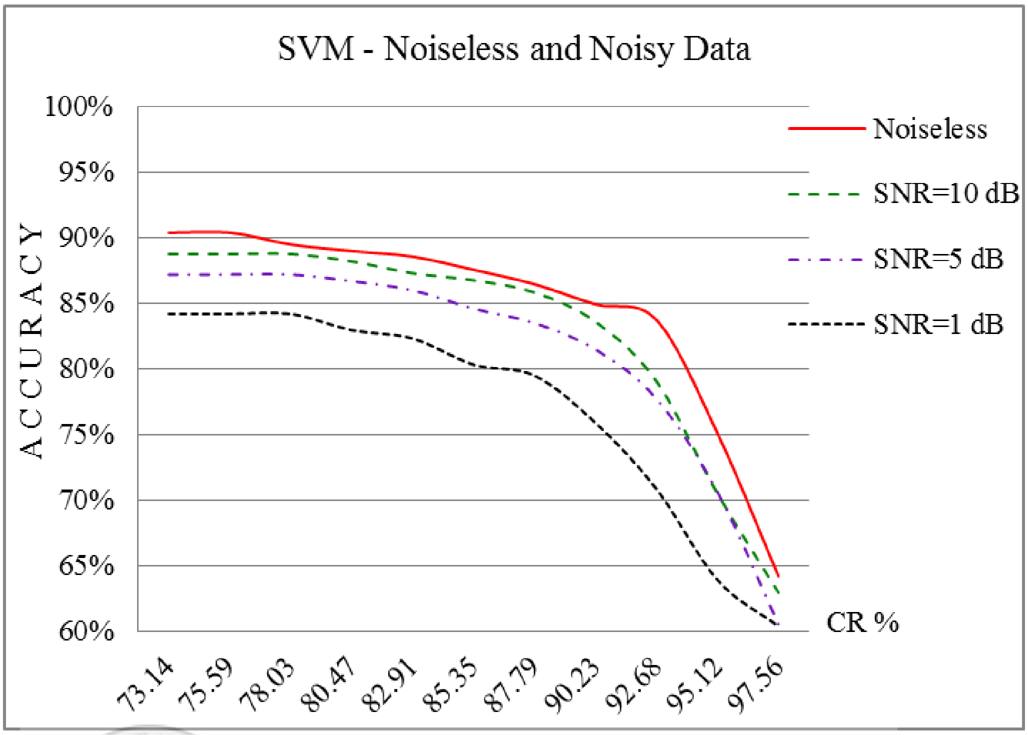


Figure 5.5. Classification accuracy using SVM classifier for noisy data

Figure 5.5 shows the SVM classifier and the results indicate that for all data types the accuracy decreases linearly with the increase of CR. In the first region, the accuracy remains stable above 84% for all data types with CR=83%, which seems to have less accuracy than the noiseless data of 5% approximately. The decay in the accuracy seems to be reasonable in this region, showing SVM with high compression values above 88% and classification accuracy starts to decrease exponentially in the third region for all data types.

Finally, as shown in Figures 5.2 to 5.5, the best compression ratio is 85.35% and gives the best classification accuracy around 85%. In addition, the ANN classifier is more accurate and gives a better accuracy of 95%. Bayes is less complex and uses only

Laplace correction. However, SVM and *k*-NN give less accuracy than ANN and Bayes. This is mainly because these classifier models use different classification strategies. For example, the Bayes classifier assumes that the features of the input pattern are independent. In the case of ANN, the dependency relationship can be learned from data. For the k-NN and SVM models, the default parameters were used.

From the above results, the best CR for both data types (noiseless and noisy) is 85.35%, which is equivalent to a compressed data of 600 samples. This CR percentage as 100 multiplied by the result of subtracting from 1 the value of *M* (size 600 s) divided by the original data *N* size (4096 s) as follows: $CR = \left(1 - \frac{M}{N}\right) * 100$.

5.3 Ensemble Classifiers Performance

In this work, all the classifiers by selecting *k*=2 for the *k*-fold cross-validation procedures have been evaluated individually. This 2-fold cross-validation was used in order to reduce the running time of the computation and the number of experiments. The advantage of this is that the data is divided into 50% training and 50% testing, so each data point is used for both training and validation on each fold. However, in 10-fold cross-validation which is commonly used, the data is randomly divided into *k* subsamples size. This *k* is booked as validation data for testing the model and the subsamples remaining are used as training. Then repeated *k* times (the folds) and the results from folds could be averaged to obtain a single estimation.

At this stage, a combined classifier called Noise-aware Signal Combination (NSC) technique has been proposed to measure the classification accuracy of certainty and uncertainty (noisy) data. In addition, based on the literature review, this study focused on getting more than 80% better accuracy than our target compression ratio of 85.35%. The stability of performance of the NSC classifier is assessed based on different statistical measurements, such as class prediction, sensitivity, specificity, and classification accuracy. Again, as the above mentioned that, for each point on the graphs, 10 experiments have been conducted and calculated the average accuracy accordingly.

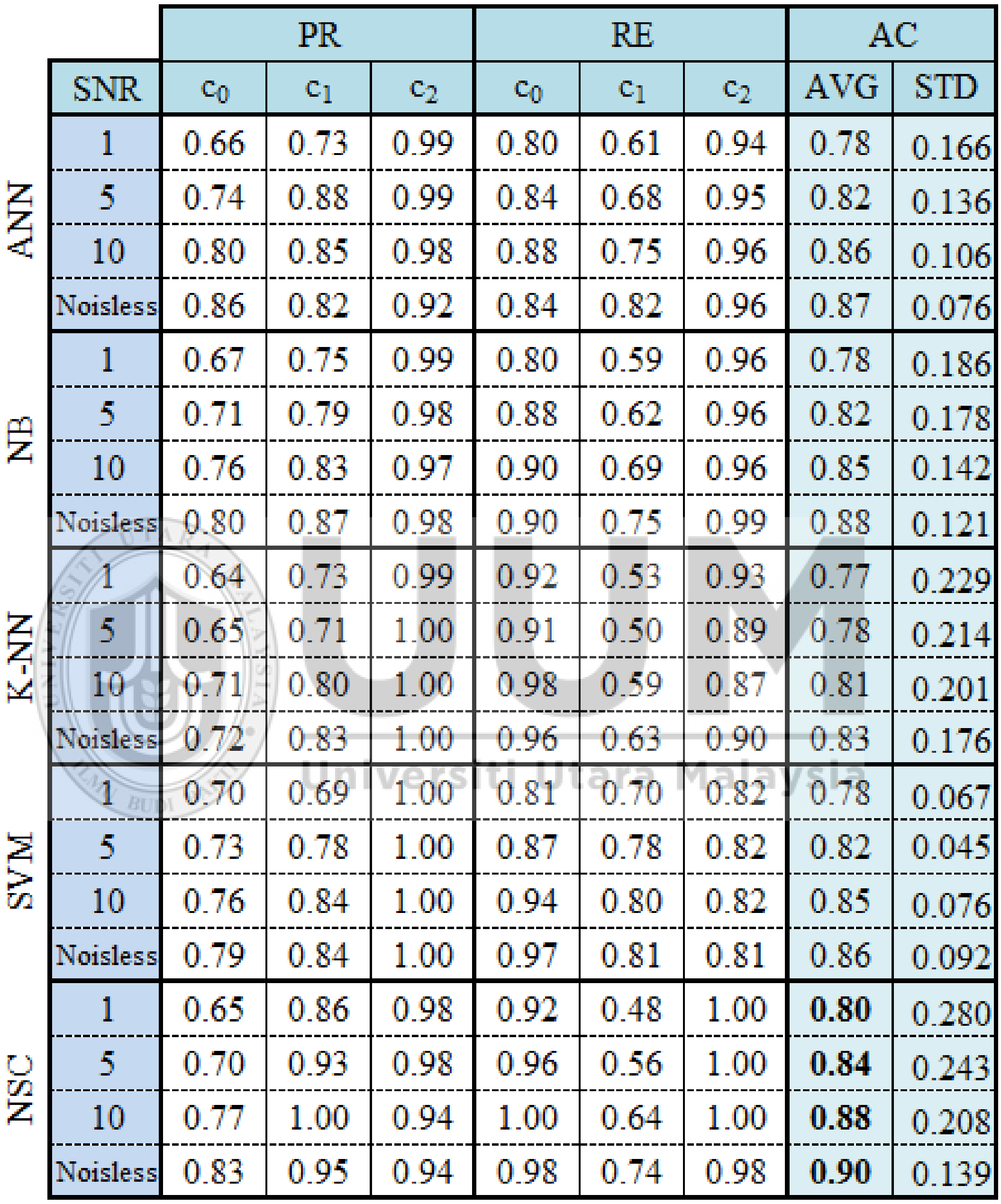
5.3.1 Confusion matrix for Performance Measures

In this research work, the EEG-based epileptic seizure data classified for different types of data, noiseless and different values of SNR to become noisy data. For each point on the graphs, 10 experiments have been conducted and calculated the average accuracy and its standard deviation accordingly. The standard deviation describes the distribution range, describing how much difference occurs between successful computations, which correspond to the data imperfection. In this case, the Standard Deviation (SD) is important to show the difference between successive measurements to make sure that the classifiers are not affected by data imperfection.

The calculated performance measures of the studied classifiers with EEG-epileptic seizure data compressed with CR= 85.35% for noiseless and add noise of SNR= 1, 5, and 10 dB are shown in table 5.1. This figure shows the class precision (PR), class recall (RE) and the classification average (AVG) accuracy (AC) and standard deviation (STD) for each classifier for different SNR and noiseless channel conditions. The figure represents the confusion matrix of the classifier outputs at the desired compression ratio. In addition, it includes most of the parameters that are used in this research, namely specificity, sensitivity, and overall accuracy of all classifiers. This table has seen that the NSC enhances the overall accuracy of the EEG data. The NSC result shows that for the noiseless data, this classifier achieves 90%, for SNR=1 dB, it achieves 80%, for SNR=5 dB, it achieves 84%, and for SNR=10 dB, it achieves 88%. The proposed NSC technique achieves a better accuracy than all existing classifiers.

Table 5.1:

Performance of the classifiers with CR=85.35% and SNR= 1, 5, and 10 dB as well as noiseless



		PR			RE			AC	
	SNR	c_0	c_1	c_2	c_0	c_1	c_2	AVG	STD
ANN	1	0.66	0.73	0.99	0.80	0.61	0.94	0.78	0.166
	5	0.74	0.88	0.99	0.84	0.68	0.95	0.82	0.136
	10	0.80	0.85	0.98	0.88	0.75	0.96	0.86	0.106
	Noisless	0.86	0.82	0.92	0.84	0.82	0.96	0.87	0.076
NB	1	0.67	0.75	0.99	0.80	0.59	0.96	0.78	0.186
	5	0.71	0.79	0.98	0.88	0.62	0.96	0.82	0.178
	10	0.76	0.83	0.97	0.90	0.69	0.96	0.85	0.142
	Noisless	0.80	0.87	0.98	0.90	0.75	0.99	0.88	0.121
K-NN	1	0.64	0.73	0.99	0.92	0.53	0.93	0.77	0.229
	5	0.65	0.71	1.00	0.91	0.50	0.89	0.78	0.214
	10	0.71	0.80	1.00	0.98	0.59	0.87	0.81	0.201
	Noisless	0.72	0.83	1.00	0.96	0.63	0.90	0.83	0.176
SVM	1	0.70	0.69	1.00	0.81	0.70	0.82	0.78	0.067
	5	0.73	0.78	1.00	0.87	0.78	0.82	0.82	0.045
	10	0.76	0.84	1.00	0.94	0.80	0.82	0.85	0.076
	Noisless	0.79	0.84	1.00	0.97	0.81	0.81	0.86	0.092
NSC	1	0.65	0.86	0.98	0.92	0.48	1.00	**0.80**	0.280
	5	0.70	0.93	0.98	0.96	0.56	1.00	**0.84**	0.243
	10	0.77	1.00	0.94	1.00	0.64	1.00	**0.88**	0.208
	Noisless	0.83	0.95	0.94	0.98	0.74	0.98	**0.90**	0.139

The above table shows the accuracy at CR=85.35% for all classifiers. It indicates that that accuracy at SNR=10dB is better than SNR=5dB and SNR=1dB which means that at 10dB the EEG characteristics are close to the noiseless data, while SNR=1db is representing high noise in EEG data. The corresponding results for the individual classifiers ANN, NB, k-NN (with *k*=10), and SVM in each SNR case together with that of NSC are plotted to illustrate the differences of their performances. Figures 5.6 - 5.10 illustrate that performance under SNR of 0, 1, 5, and 10 dB respectively. The corresponding readings of the accuracy in table 1 are emphasized in figures 5.6 - 5.10 with the line drawn at CR=85.35%.

The constraint on the desired accuracy in the case of noiseless data is to achieve 90%. The CR of 84.35% was the cutting edge of achieving this desired goal. Therefore, the performance of the classifiers is of its highest importance at this CR value. The overall accuracy results of all the experiments show that this constraint is met at CR=85.35%, while at the same time maintaining a high accuracy of 80% with very noisy data of SNR=1 dB.

5.3.2 Experiment 3: Effect of Compression on Noiseless data for NSC classifier

The aim of this experiment is to show the accuracies of the effect of the compression on different EEG data classes. Each figure shows the effect on noiseless data type, which is evaluated by the individual and the proposed NSC classifiers. The purpose from the following figures (5.6 - 5.10) is to show that even in a single test, the proposed technique still outperforms the others, however, based on average for each point could reduce these oscillation through averaging results and this shows the versatility of the results.

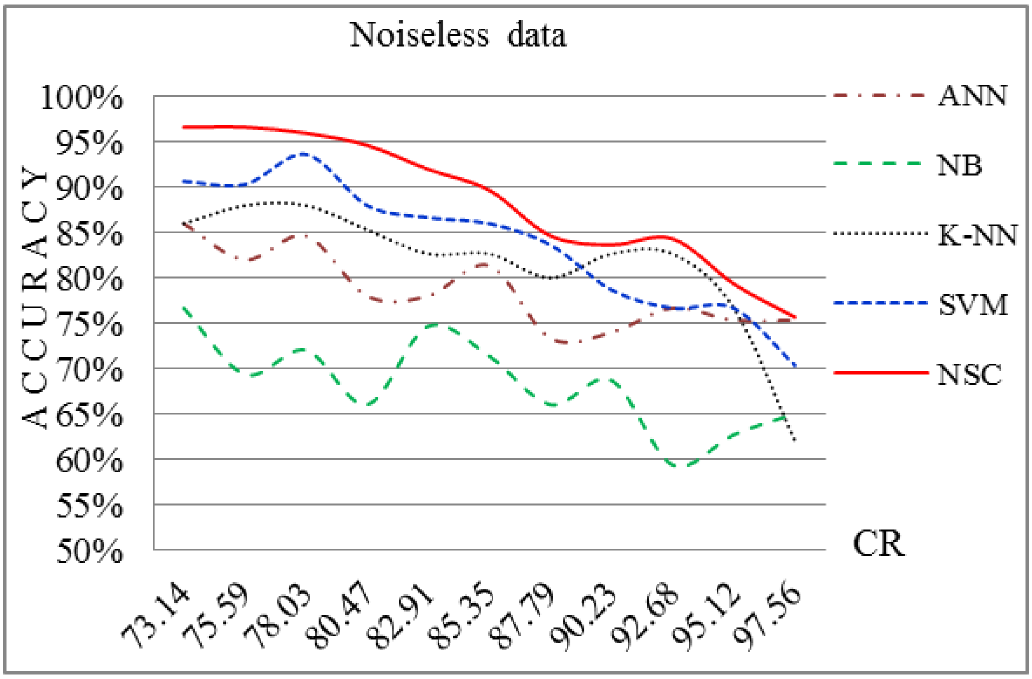


Figure 5.6. Classification accuracy against CR for noiseless

Figure 5.6 shows the trend of the classification accuracy of NSC increases almost linearly with the decrease of CR. The decay in the accuracy seems to be reasonable in all regions, showing that NSC outperforms the other classifiers in high accuracy values, and then it starts to decay exponentially like all of the other individual classifiers. Because the decay of *M* value gives high CR values, this means that some of the data will lose some of its components. Indeed, the model shows an almost horizontal linear curve and a moderate decrease in the second region, but in extremely CR values, the curve becomes steep.

5.3.3 Experiment 4: Effect of Compression on noisy data for NSC classifier

The experiment aims to show the accuracy of the effect of the compression on different EEG data classes. Each figure shows the effect on different data types, which are evaluated by each individual classifier and the proposed NSC classifier.

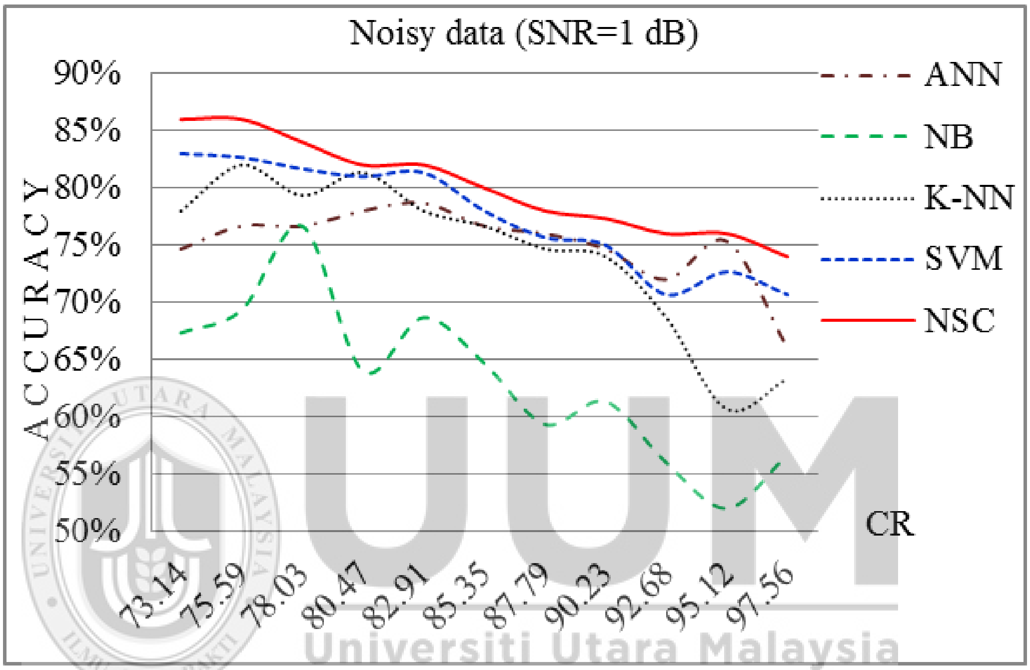


Figure 5.7. Classification accuracy in function of CR for SNR=1 dB

Figure 5.7 shows that the result is worse for all classifiers due to the quantity of added AWGN (SNR=1 dB), which is considered to be the highest noise added in all experiments. In this case, the NSC continued to perform consistently better than the rest of the classifiers. Bayes classifier still exhibits the poorest performance. The exact reported results at CR=85.35% show 80% as classification accuracy and then all classifier accuracies moderately decrease. Other results can be seen when for a moderate SNR of 5 dB in Figure 5.8.

As expected, Figure 5.8 shows that increasing the CR results in decreasing the overall accuracy for all classifiers and vise-versa. The results show that a model with high classification corresponds to a low compression ratio value. At the desired CR=85.35%, the NSC gives a classification accuracy of 84% of SNR=10 dB noised data. Also, at CR=90.23% the NSC accuracy is 83% and starts decreasing, and in all classifiers accuracy decreases exponentially.

Eventually, Figure 5.9 shows a slightly different behavior for all classifiers. While the classification accuracy of above 90% starts to extremely decay after CR=82.91%. The effect of added different SNR values using AWGN is that the noise in this case is much less where SNR=10 dB, which is close to the EEG data components. However, at those accuracy and CR values, the accuracy curve starts to decrease exponentially. This is due to the decrease in the value of *M*, that will give the high CR, so that the data goes a little bit to lose some of the information. Indeed, the model shows a decrease in the classification accuracy for all classifiers and an increase in the CR values. The curve is going to decay.

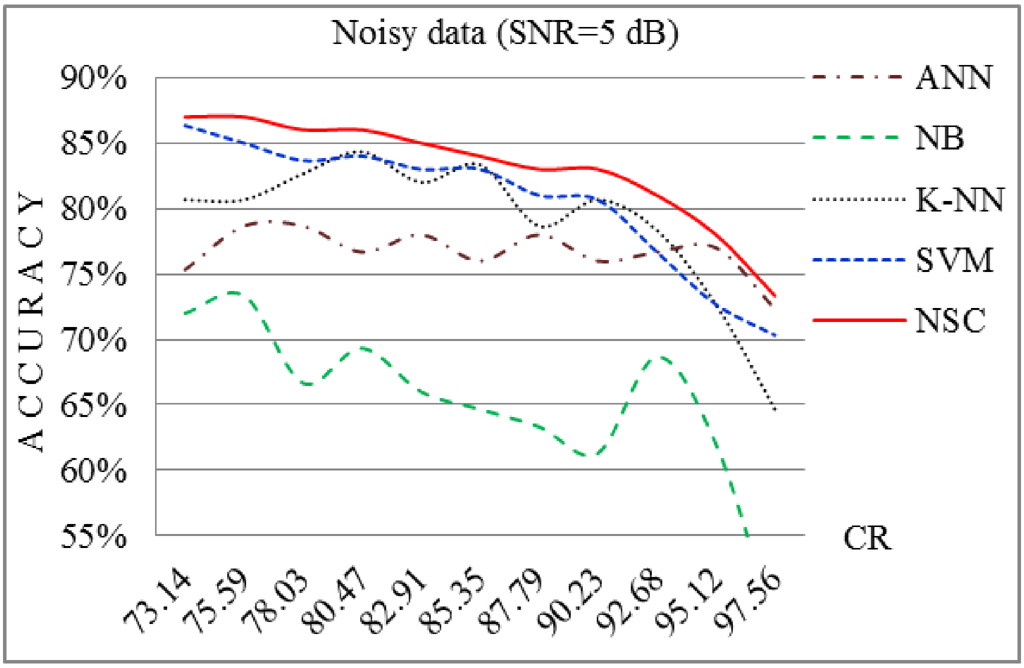


Figure 5.8. Classification accuracy in function of the CR, for SNR=5 dB

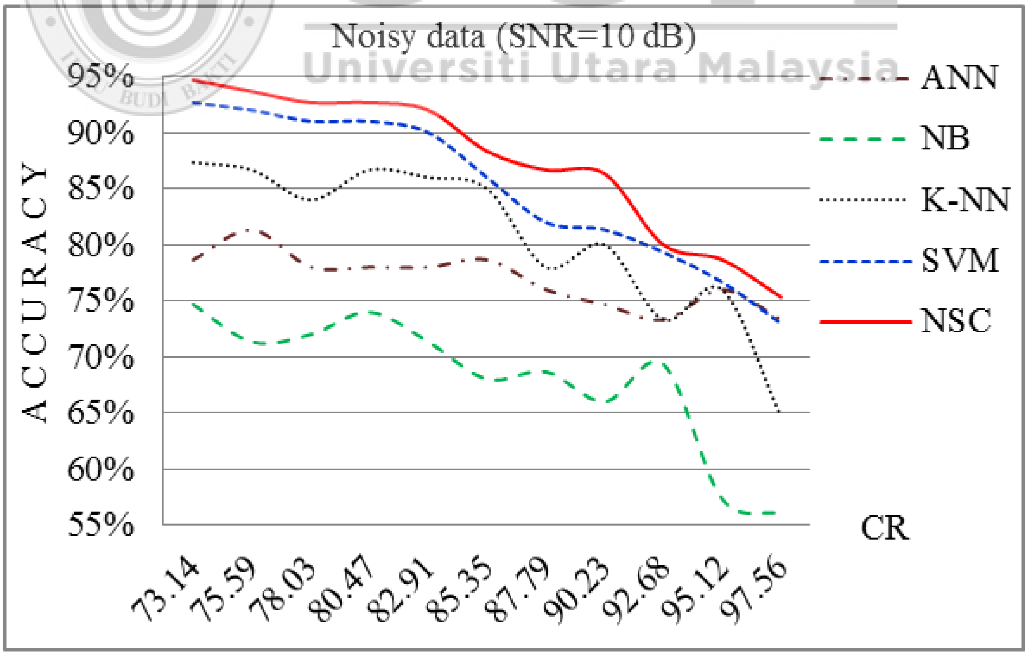


Figure 5.9. Classification accuracy in function of the CR, for SNR=10 dB

5.3.4 Average Accuracy for all CR with all SNR

Figure 5.10 shows the relationship between the average classification accuracy and the average of all data types (noiseless and noisy) for each classifier. This relationship is represented for all individual classifiers and the proposed NSC method. This shows the accuracy of the effect of the compression on different EEG data classes. In addition, this figure shows the effect on the average of accuracies for different data types which are evaluated by each the individual classifier and the proposed NSC classifier.

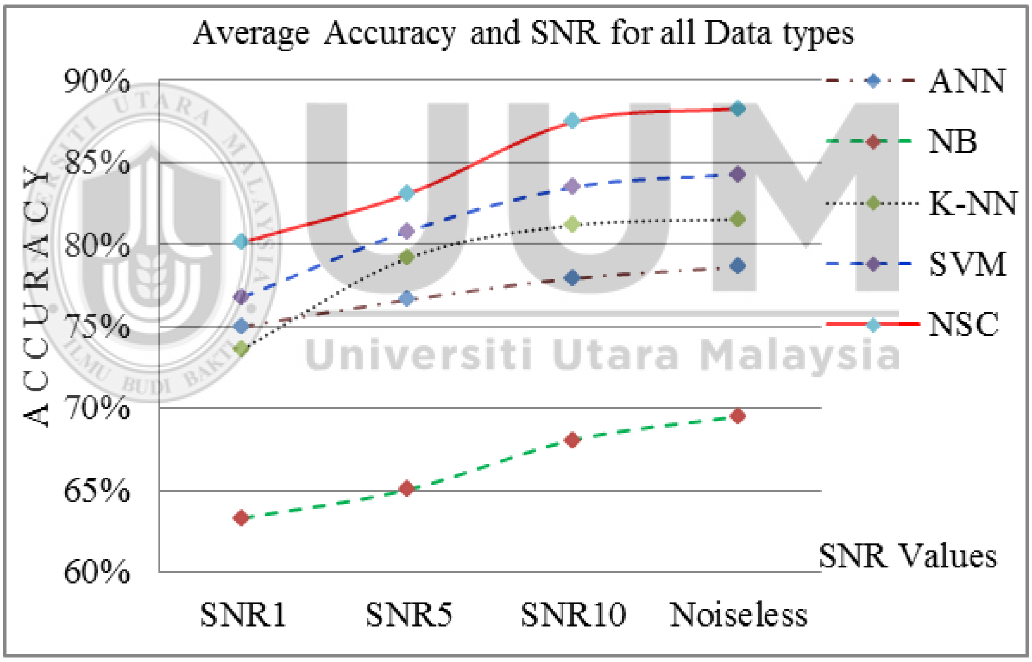


Figure 5.10. Average accuracy for all CR values with noiseless and SNR=1, 5, and 10 dB

Finally, overall and regardless of the compression ratio value, Figure 5.12 shows that the results for the average classification accuracy of NSC is constantly better than other classifiers. This statement is valid for EEG-epileptic seizure data in both noiseless and noisy. Also, this figure shows that the best accuracy average for the proposed NSC method is 88%, 80%, 83% and 87% for respectively (noiseless and noisy) data with SNR values of 1 dB, 5 dB, and 10 dB, respectively.

Table 5.2 shows the overall accuracy for all individual classifiers at training stage, in addition to proposed NSC at our intended compression ratio 85.35%.

Table 5.2:

Overall classification accuracy for all classifiers at CR=85.35%

Noise Values	**ANN**	**Bayes**	**k-NN**	**SVM**	**NSC**
Noiseless	87.00%	88.00%	83.00%	86.00%	**>90.00%**
SNR=1	78.00%	78.00%	79.00%	78.00%	**80.00%**
SNR=5	82.00%	82.00%	77.00%	82.00%	**84.00%**
SNR=10	86.00%	85.00%	86.00%	82.00%	**88.00%**

As shown in the above Table 5.2, at different values of SNR, the classification accuracy was obtained using four individual classifiers. At the SNR=1 dB the accuracy is lower since the noise quantity is high. At the SNR=5 dB the accuracy is moderately and better than the SNR=1 dB. At SNR=10 dB is almost close to or in some of classifiers it is better than no noise value, however, this is appeared based on the behavior of each classifier where each one is from different families.

Compared with previous works, the proposed NSC classification accuracy of noiseless EEG data has achieved 90%, which is 5% higher than the accuracy done in Sharma [104], 4.1% higher than the work done in Sadati's [87] (85.9% accuracy) especially for sets (A, D, E), 0.5% higher than the reported in Mohamed's [84] (89.5% accuracy). In addition, Liang achieved classification accuracy between 80% and 90% [92]. Tzallaz [81] achieved 89% only for noiseless dataset using one classifier. All of these approaches considered are using the same EEG dataset. In contrast to these approaches, the proposed approach achieved the desired improved classification accuracy with noisy data using different SNR values: 80% for SNR=1 dB, 84% for SNR=5 dB and 88% for SNR=10 dB. These results were obtained at CR =85.35%. Moreover, our approach provides several demonstrable benefits, such as simplicity, and improves the overall classification accuracy. In addition, to the best of our knowledge, no similar approaches have been evaluated for EEG-based epileptic seizure when considering AWGN channel with different SNR values.

The following table 5.3 shows the comparisons between the proposed NSC and others reported in the literature, most of literatures were worked only on noiseless data, which shows in table 5.3 by Not Available (N/A). In addition, the table shows that the proposed technique achieved more than 90% accuracy compared with others on the same dataset.

Table 5.3:

Comparison summaries of the previous work

Authors	Noiseless Datasets	Noisy Datasets	Classifiers	Accuracy
Proposed Method	A, C, E	A, C, E	ANN, NB, k-NN, SVM	>90%
Sharma 2014, [104]	Two different classes	N/A	LS-SVM	85%
Sadati's, 2006 [87]	A, D, E	N/A	SVM, FBNN, ANFIS, and proposed ANFN	85.9%
Mohamed 2013 [84]	A, B, C, D, E	N/A	NB, MLP, k-NN, LDA, AND SVM	89.5%
Tzallaz 2009 [81]	A, B, C, D, E	N/A	ANN	89%

5.4 Summary

In this chapter, the above experiments focused on results and obtaining a better classification accuracy of 85% for all reconstructed and noisy data types, noiseless and noisy with different values, according to the desired compression ratio of 85.35%. At this stage of research, all individual classifiers are evaluated by selecting *k*=10 as it is a common choice for the *k*-fold cross-validation method in order to achieve the classification accuracy. In this section, the results were obtained in two situations, for a 10-fold cross-validation and for a 2-fold cross-validation.

CHAPTER SIX
CONCLUSION AND FUTURE WORK

6.1 Summary and Conclusion

The electroencephalogram (EEG) signal is the most common measure of brain activity, making it used necessary in the diagnosis and treatment of brain neurodegenerative diseases and abnormalities. The identification of different types of EEG signals is challenging, as it requires the analysis of multiple sets of EEG data. The classification of EEG signals is essential in assessing the brain activity. Feature Extraction best classifies these signals.

In this thesis, EEG signal processing and classification techniques have been investigated and developed in order to identify different types of EEG signals, through the following three main objectives:

- Develop a unified framework for compression and delivery of EEG data across transmission channels. This is essential due to the large data size of the WBASN’s EEG data during data transmission.
- Develop methods for the classification of the reconstructed epileptic EEG noisy signals, in order to improve the classification rate with less execution time.
- Evaluate the efficiency of such technique by achieving better EEG classification accuracy and compression ratio.

In order to achieve the above objectives, this research work proposes two significant approaches, first: design a unified compression and classification framework. Second: propose a noise-aware signal combination (NSC) technique. The framework is used to down sample or compress EEG signal and combine both compressive sensing (CS) and discrete cosine transform (DCT) methods to reduce the size of data streaming over the channel. On the other hand, an inverse DCT is used to reconstruct the compressed EEG signal. For classification, a noise-aware signal combination (NSC) method is developed to enhance the accuracy. Ultimately, NSC will affect the sensor battery during the data transmission and contribute to the EEG-based epileptic seizure classification.

The proposed framework on a benchmark dataset consisting of three subjects, A, C, and E has been applied. Additionally the impact of normal and impaired wireless channels on the transmitted compressed EEG epileptic seizures signal has been investigated. The data was collected by a pre-surgical diagnosis; it was organized for the study to differentiate between the healthy subjects and epilepsy disease-suffering ones. More details on the benchmark EEG dataset are provided in Chapter 5.

The experiments of the framework are concentrated on the classification accuracy that gives better compression ratio (CR). The CR and accuracy were determined by calculated the CR with an original size of raw data N = 4096 samples and a different compressed value M. The compression process has been developed at different M values ranging from 100s/s up to 1100s/s. This variation of compressed values has been adopted in order to achieve better accuracy. In the SCC, four classifiers ANN, NB, k-NN, and SVM have been employed to categorize the reconstructed EEG epileptic signals. The performance of these classifiers through the 10-fold cross validation procedure was evaluated.

The framework has been developed into two stages. In the first stage, different classes of raw EEG data have been compressed. The purpose was to reduce the transmission cost at the transmitter. Hence, it combined different methods: CS, DCT, and random matrix, and then sent the compressed data. The aim of the compression is to evaluate the trade-off between the complexity of CS and accuracy in wireless tele monitoring. An additive white Gaussian noise (AWGN) was deployed as a wireless channel model between the transmitter and receiver. The second stage consists of reconstructing the compressed EEG data to the original size using CS and iDCT at the receiver side. The iDCT method was utilized for data reconstruction to achieve the low-complexity of compression paradigms. Similarly, the discrete wavelet transform (DWT) method was used to extract conventional statistical features. Finally, both conventional statistical features and the wavelet sub-band were combined to formulate a feature vector of 32 attributes. These attributes represent the distribution of EEG signals and are used as input vectors for a set of classifiers. These features are applied as input to the ANN, NB, k-NN, and SVM classifiers to classify three-class EEG signals in both noiseless and noisy data.

The results revealed that ANN has a better accuracy at 95%, which outperforms the other three classifiers, namely SVM, k-NN, and Bayes. However, the implementation complexity, and classification latency make ANN less preferable for real-time tele-monitoring applications. The experimental evaluation concluded that the classification accuracy is achieved 93.67% with a CR of 85.35% for EEG epileptic data. Therefore, the experimental results demonstrated that the SCC is promising for both compression and reconstruction characteristics of EEG signals as well as for the classification of EEG signals.

To reduce the computation time and the number of experiments as well as to improve the classification performance, four classifiers through the 2-fold cross-validation procedure have been evaluated. Noise is introduced to the data at different levels of signal-to-noise-ratio (SNR), 1, 5, and 10 dB. The compression paradigm with low complexity is achieved by utilizing the iDCT method for data reconstruction. Some features have been extracted from the reconstructed data using DWT. An EEG noise-aware signal combination (NSC) method for EEG-based epileptic detection has been developed (Chapter 4). The NSC method as well as the ANN, Bayes, k-NN, and SVM classifiers are tested with different categories of EEG-based epileptic seizure data. The proposed NSC combination method constantly performs better than the above four classifiers.

The experimental results show that the proposed NSC technique is efficient with noisy data, while satisfying the constraint of 90% accuracy for noiseless data compared with 85.9% and 89.5%. NSC achieved the desired improved classification accuracy in case of noisy data. NSC achieved a classification accuracy of 80% for SNR=1 dB, 84% for SNR=5 dB, and 88% for SNR=10 dB, at CR =85.35%. While improving the overall classification accuracy, NSC also provides several significant benefits, such as structural and computational simplicity.

Comparing the proposed NSC method results over legacy classifiers, the proposed method proved greatly efficient and obtained better accuracy for both noiseless and noisy data. Compared with previous works, the classification accuracy by the proposed NSC of noiseless EEG data achieved 90%, which is 4.1% higher that the work done in [87] (85.9% accuracy), 0.5% higher than that reported in [84] (89.5% accuracy), and finally, Liang achieved classification accuracy between 80% and 90% [92] considering the same dataset. In addition, there are no similar evaluation approaches for EEG-based epileptic seizure when considering AWGN with different SNR values. Moreover, new interesting results could be realized that the thermal noise using AWGN clearly affects the classification accuracy. The NSC method consumed less time in comparison to the classical legacy classifier techniques and some methods reported in the literature.

This research achieved the objective to reduce the size of EEG signals at the transmitter side, which eventually reduces the transmission time and saves the sensor's battery during data transmission. The research work has been made in the following main steps:

- Design and develop a sensor-based compression and classification framework for a trade-off between the compression ratio (CR) and classification accuracy of the transmitted compressed EEG data. The metrics were CR and accuracy, including specificity and sensitivity (Chapter 4). The results show good compression ratio that could be attained without affecting the classification accuracy for all classifiers.
- Design and develop a noise-aware signal combination method for data with some degree of uncertainty (imperfection) representing the EEG signal. The output of four legacy classifiers is fed in the proposed technique (Chapter 4).
- Devise a mathematical formulation of the prediction of the combined classifier to calculate the hypothesis with the highest probability (Chapter 4). The formula

calculates the goodness probability of a hypothesis *j* for predicting class *k* by dividing the summation of the weighted $0 < \sigma < 1$ terms, by the total of goodness for all hypotheses if they were to predict the same class *k*. The two terms of class precision and class recall are dependent and grouped in one term, leaving the overall accuracy as second term. These two terms have direct ratio to the goodness of selecting a classifier predicting class *k*. The result value represents the goodness of hypothesis *j* reporting class *k* with respect to other hypotheses reporting the class *k*.

- Evaluate the research work on the benchmark EEG dataset of three different classes. After comparing our results with previous work [84, 87, 92] using three data sets, the results show that our proposed method outperformed other classifiers (Chapter 5).

6.2 Research Contributions

Thus, the main contributions of this research are,

1. A unified framework for EEG compression using CS and Additive White Gaussian Noise (AWGN) channel transmission has been developed,
2. A new noise-aware signal combination (NSC) method that supports both types of biomedical EEG data for both noiseless and noisy is proposed,
3. A series of comprehensive experiments were conducted to examine and evaluate the effectiveness and robustness of the NSC method for classifying contaminated EEG data.

6.3 Future Work

Using the proposed framework, it has been verified that regardless of the classifier type, classification accuracy would not be affected by signal impairments for compression ratios up to 80%, which opens the door for enhancing an energy-efficient delivery of medical data over wireless channels. By emphasizing compression, it has been verified that the power consumed for data communication can be reduced drastically without affecting the application accuracy, which is a significant advantage.

In addition, seizure detection approaches were studied intensively over the last years, and automatic seizure detection is playing a key role in a long-term. This research work can be extended to develop a popularized mHealth (m for mobile) system for reliable and efficient monitoring of EEG medical data, and leveraging sensors and mobile technologies to facilitate the remote connecting of patients with medical systems. The system raised a solution for accessible network architecture, active signal processing, reliable wireless channel, looking for the energy-efficient communication, and analysis for perfect EEG medical diagnosis.

In order to achieve such extension, the following directions can be addressed:

- Develop a signal processing technique to leverage a compressive sensing method for precise reconstruction of medical data.
- Leverage communication layers by developing a communication technique to deliver energy-efficient transmission of vital signs.
- Develop a model to demonstrate the classification accuracy for efficient remote monitoring of medical conditions, such as seizure detection and detect the mental task for real-time brain computer system interface.

REFERENCES

[1] M. Hanson, "Wireless Body Area Sensor Network Technology for Motion-Based Health Assessment", A Dissertation presented to University of Virginia, for the Degree Doctor of Philosophy, USA, August 2009.

[2] I. Stojmenovi, "Handbook of Sensor Networks: Algorithms and Architectures", University of Ottawa, John Wiley & Sons, Inc., 2005.

[3] J. Lach, J. Aylor, N. Merris, M. Hanson, C. Rehorn, "Wearable Gait Data Collection for Longitudinal Fall Analysis", International Conference on Aging, Disability and Independence, Virginia, USA, 2003.

[4] D. Culler, D. Estrin, and M. Srivastava, "Overview of Sensor Networks", Journal of IEEE Computer Society, Vol. 37, pp. 41-49, August 2004.

[5] S. Chen, H. Lee, C. Chen, H. Huang, and C. Luo "Wireless Body Sensor Network with Adaptive Low-Power Design for Biometrics and Healthcare Applications", Journal of IEEE Systems, Vol. 3, December 2009.

[6] A. Darwish, and A. Hassanien, "Wearable and Implantable Wireless Sensor Network Solutions for Healthcare Monitoring," Journal of Sensors, ISSN 1424-8220, 2011.

[7] K. Sohraby, D. Minoli, T. Znati, "Wireless Sensor Networks Technology, Protocols, and Applications", by John Wiley & Sons, ISBN: 9780471743002, 2007.

[8] Wikipedia Electrocardiogram, http://en.wikipedia.org/wiki/Electrocardiogram; accessed on Nov.2nd, 2012.

[9] Wikipedia Accelerometer, http://en.wikipedia.org/wiki/Accelerometer; accessed on November 2nd, 2012.

[10] L. Huang, M. Ashouei, F. Yaziciogl, J. Penders, R. Vullers, G. Dolmans, P. Merken, J. Huisken, H. de Groot, C. Van Hoof, and B. Gyselinckx, "Ultra-Low Power Sensor Design for Wireless Body Area Networks: Challenges, Potential Solutions, and Applications", Journal of Digital Content Technology and its Applications Vol: 3, September 2009.

[11] K. Abualsaud, M. Mahmuddin, A. Mohamed, "WBASN Signal Processing and Communication Framework: Survey on Sensing, Communication Technologies, Delivery and Feedback", Journal of Computer Science (JCS), pp:121-132, New York, Vol. 8, 2012.

[12] The Kaiser Family Foundation, http://kff.org/health-costs/issue-brief/

[13] OnWorld "Wireless Sensor Networks for Healthcare", industrial market report, http://onworld.com/healthcare/index.html, August 2010.

[14] OnWorld Inc. provides global business intelligence on smart technology markets, California, USA, "Health & Wellness Wireless Sensor Networks", industrial market report, http://ww.onworld.com/healthcare/index.html, 2013.

[15] C. Chong, S. Kumar, ''Sensor Networks: Evolution, Opportunities, and Challenges", Proceedings of the IEEE, Vol. 91, August 2003.

[16] C. García-Hernández, P. H. Ibargüengoytia-González, J. García-Hernández, and J. A. Pérez-Díaz, "Wireless Sensor Networks and Applications: A Survey", IJCSNS International Journal of Computer Science and Network Security, Vol.7, March 2007.

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

[18] W. Xue-qing, 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, Chengdu, China, May 2009.

[19] 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.

[20] B. Dolan, Home monitoring: "Berg Insight, Home health monitoring market", published on Tuesday - December 21st, 2010.

[21] C. Raghavendra, K. Sivalingam, T. Eds., “Wireless Sensor Networks: Architectures and Protocols”, ISBN: 0849318238, Kluwer Academic, New York, 2004.

[22] J. Li and J. Li, “Data sampling control, compression and query in sensor networks”, International Journal of Sensor Networks, Vol 2, 2007.

[23] People living with chronic disease http://www.welldocinc.com accessed 07/02/2012.

[24] Intel Corporation, Reconceiving disease management Steps toward achieving the promise. http://www.intel.com/healthcare/pdf/7INCO001_WhitePaper_LO6.pdf, November 7th, 2014

[25] World health organization strategy “Country Cooperation Strategy for WHO and Qatar” http://www.who.int/countryfocus/cooperation_strategy/ccs_qat_en.pdf accessed on November 1st, 2014.

[26] Chronic disease report from WHO “The impact of chronic disease in Qatar” http://www.who.int/chp/chronic_disease_report/media/impact/en/index.html accessed on April 1st, 2014.

[27] J. Y., Khan, and M. R., Yuce,“Wireless Body Area Network (WBAN) for Medical Applications” In: Campolo, D. (Ed.), New Developments in Biomedical Engineering. INTECH, ISBN: 978-953-7619-57-2, 2010.

[28] K. Abualsaud, A. Mohamed and M. Mahmuddin, "Survey on Wireless Body Area Sensor Networks for Healthcare Applications: Signal Processing, Data Analysis and Feedback", Proceedings of the 3rd International Conference on Computing and Informatics, ICOCI2011, 8-9 June, Bandung, Indonesia 2011.

[29] H. Karl and A. Willig, “Protocols and Architectures for Wireless Sensor Networks”, Handbook, John Wiley & Sons, Ltd. ISBN: 0-470-09510-5, 2005.

[30] C. Shen, C. Chan, F. Lin, and M. Chiu, “Epileptic Seizure Detection for Multichannel EEG Signals with Support Vector Machines”, in the proceeding of the 11th IEEE International Conference on Bioinformatics and Bioengineering, 2011.

[31] I. Güler, E. Beyli, “Adaptive Neuro-Fuzzy Inference System for Classification of EEG Signals using Wavelet Coefficients”, Journal of Neuroscience Methods, Elsevier, Vol. 148, 2005.

[32] A. Subasi, “EEG Signal Classification using Wavelet Feature Extraction and a Mixture of Expert Model”, Journal of Expert Systems with Applications, Elsevier 32, 2007.

[33] L. Iasemidis, D. Shiau, W. Chaovalitwongse, J. Sackellares, P. Pardalos, *et al.* “Adaptive Epileptic Seizure Prediction System”, Journal of IEEE Transactions on Biomedical Engineering, 50(5), 616–627, 2003.

[34] J. Wang, P. Liu, M. She, S. Nahavandi, A. Kouzani, “Bag-of-words Representation for Biomedical Time Series Classification”, Journal of Biomedical Signal Processing and Control Journal, Vol. 8, Elsevier 2013.

[35] Sensing processing: http://en.wikipedia.org/wiki/Sensing, accessed November 30, 2014.

[36] R. Latuske, ARS Software GmbH, “Bluetooth Health Device Profile (HDP)”, ARS Software GmbH, Starnberger Str. 22, D-82131 GAUTING/Munich. 2009

[37] M. Adil, K. Abbasi, and M. R. Ashraf, “Wireless Body Area Networks for Medical Applications”, published in CIIT Student Transactions on Radio Communications, Vol. 2, June 2013.

[38] Polysomnography: http://en.wikipedia.org/wiki/Polysomnography

[39] EEG-Electrodes: http://www.wisegeek.com/topics/eeg-electrodes.htm#

[40] A. Tzallas, M. Tsipouras, and D. Fotiadis, “Automatic Seizure Detection Based on Time-Frequency Analysis and Artificial Neural Networks,” Hindawi Publishing Corporation, Computational Intelligence and Neuroscience, Vol. 2007, Article ID 80510, 13 pages, 2007.

[41] H. Cao, V. C.M. Leung, C. Chow and H. Chan, “Enabling Technologies for Wireless Body Area Networks: A Survey and Outlook” IEEE Communication Magazine, December 2009.

[42] A. P. and N. Bourbakis, “A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis”, Journal of IEEE Transactions on Systems, MAN, and CYBERNETICS-PART C: Applications and Reviews, Vol. 40, January 2010.

[43] UC Berkeley, “TinyOS Community Forum, an Open-Source OS for the networked sensor regime” http://www.tinyos.net/ accessed on January, 2013.

[44] Preprocessing: http://searchsqlserver.techtarget.com/definition/data-preprocessing October 13, 2012

[45] S. Kotsiantis, D. Kanellopoulos and P. Pintelas, “Data Preprocessing for Supervised Leaning*”* International Journal of Computer Science, Vol. 1, 2006.

[46] M. Xu; “Research and Design of Data Preprocessing of Wireless Sensor Networks based on Multi-Agents” in proceeding of IEEE International Conference on Network Infrastructure and Digital Content, November 6-8, Beijing, China, 2009.

[47] N. Ahmad, D. Hoang, and M. Phung, “Robust Preprocessing for Health Care Monitoring Framework”, the proceeding in 11th IEEE International Conference on e-Health Networking, Applications and Services, 2009, Australia – Dec. 16-18, 2009.

[48] H., Adeli, Z. Zhou, and N., Dadmehr,“Analysis of EEG Records in an Epileptic Patient using Wavelet Transform”, Journal of Neuroscience Methods, Vol. 123, pp. 69-87, 2003.

[49] Epilepsy: http://www.who.int/mediacentre/factsheets/fs999/en/ accessed May 26, 2014

[50] D. Rivero, V. Aguiar-Pulido, E. Fernández-Blanco, and M. Gestal, “Using Genetic Algorithms for Automatic Rrecurrent ANN Development: An Application to EEG Signals Classification,” IJDMMM 5(2): 182-191, 2013.

[51] D. Wang, D. Miao, and C. Xie, "Best Basis-based Wavelet Packet Entropy Feature Extraction and Hierarchical EEG Classification for Epileptic Detection," Expert Systems with Applications, Elsevier, pp. 14314-14320, Vol. 38, 2011.

[52] V. Bajaj and R. B. Pachori; “Classification of Seizure and Non-seizure EEG Signals Using Empirical Mode Decomposition”; Journal of IEEE Transactions on Information Technology in Biomedicine, Vol. 16, November 2012.

[53] M. Valderrama, C. Alvarado, S. Nikolopoulos, J. Maritnerie, C. Adam, V. Navarro, and M. Le Van Quyen, “Identifying an Increased Risk of Epileptic Seizures using a Multi-feature EEG-ECG Classification”, Journal of Biomedical Signal Processing and Control, Elsevier, pp. 237-244, Vol. 7, May 2012.

[54] B. Olshausen, “Aliasing, Nyquist sampling theorem", PSC 129 - Sensory Processes, October 10, 2000.

[55] A. Jain and E. Y. Chang, “Adaptive Sampling for Sensor Networks*”,* proceeding in the 1st international workshop on Data management for sensor networks, ACM New York, USA, 2004.

[56] B. Gedik, L. Liu, and P. Yu, “ASAP: An Adaptive Sampling Approach to Data Collection in Sensor Networks*”*, Journal of IEEE Transactions on Parallel and Distributed Systems, Vol. 18, December, 2007.

[57] M. Umer, E. Tanin, L. Kulik, “Opportunistic Sampling in Wireless Sensor Networks*”*, ACM-GIS, Seattle, USA, ACM 978-1-60558-649-6, November 4-6, 2009.

[58] R. Huang, G. -Huixu, “The Design of Energy-Saving Filtering Mechanism for Sensor Networks”, Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, China, July 11-14, 2010.

[59] I. Mohomed, A. Misra, M. Ebling, and W. Jerome, “HARMONI: Context-aware Filtering of Sensor Data for Continuous Remote Health Monitoring”, Sixth Annual IEEE International Conference on Pervasive Computing and Communications, Hong Kong, March 17-21, 2008.

[60] K. Hua, H. Wang, W. Wang, and S. Wu, "Adaptive Data Compression in Wireless Body Sensor Networks", 13th IEEE International Conference on Computational Science and Engineering, 2010

[61] A. Abdulghani, A. Casson, and E. Rodriguez-Villegas, “Quantifying the Feasibility of Compressive Sensing in Portable Electroencephalography Systems” in Proc. Int. Conf. Foundations Augmented Cognition, Neuroergonomics Oper. Neurosci. FAC, pp.319–328, Springer, 2009.

[62] R. Vidhyapriya and P. Vanathi, “Energy Efficient Data Compression in Wireless Sensor Networks”, International Arab Journal of Information Technology (IAIJT), Vol. 6, July 2009.

[63] H. Kim, R. Yazicioglu, P. Merken, C. Hoof, and H. Yoo, “ECG Signal Compression and Classification Algorithm with Quad Level Vector for ECG Holter System”, Journal of IEEE Transactions on Information Technology in Biomedicine, Vol. 14, January 2010.

[64] J. Kolo, S. Shanmugam, D. Lim, L. Ang, and K. Seng, “An Adaptive Lossless Data Compression Scheme for Wireless Sensor Networks", Hindawi Publishing, Journal of Sensors, Vol. 2012.

[65] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, “Compressed Sensing for Real-Time Energy-Efficient ECG Compression on Wireless Body Sensor Nodes”, Journal of IEEE Transactions on Biomedical Engineering, Vol. 58, September 2011.

[66] S. Aviyente, “Compressed sensing framework for EEG compression,” in Proc. IEEE Workshop Statistical Signal Processing (SSP), pp. 181–184, Wisconsin, August 2007.

[67] S. Senay, L. Chaparro, M. Sun, and R. Sclabassi, “Compressive Sensing and Random Filtering of EEG Signals using Slepian Basis,” In Proc. EURASIP Eusipco, Switzerland, August 2008.

[68] W. Chen and I. Wassell, *“*Energy Efficient Signal Acquisition via Compressive Sensing in Wireless Sensor Networks*”,* IEEE 6th International Symposium on Wireless and Pervasive Computing (ISWPC), Hong Kong, 23-25 February 2011.

[69] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, “Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals”, Journal of IEEE Transactions on Information Theory, Vol. 56, January 2010.

[70] R. Baraniuk, “Compressive sensing”, Lecture Note, Journal of IEEE Signal Processing Magazine, Vol. 24, July 2007.

[71] E. Candès, “Compressive sampling”, in Proceedings of the International Congress of Mathematicians, Madrid, Spain, August 2006.

[72] E. Candes and T. Tao, “Near Optimal Signal Recovery from Random Projections: Universal Encoding Strategies”, Journal of IEEE Transactions on Information Theory, Vol. 52, December 2006

[73] D. Donoho, “Compressed Sensing”, Journal of IEEE Transactions on Information Theory, Vol. 52, April 2006.

[74] S. Fauvel and R. Ward, “An Energy Efficient Compressed Sensing Framework for the Compression of Electroencephalogram Signals”, Journal of Sensors, 2014.

[75] A. Abdulghani, A. Casson and E. Rodriguez-Villegas, “Quantifying the performance of compressive sensing on scalp EEG signals”, 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL), Rome, Italy, November 7-10, 2010.

[76] A. Abdulghani, A. Casson, and E. Rodriguez-Villegas, “Compressive Sensing Scalp EEG Signals: Implementations and Practical Performance”, Journal of Medical and Biological Engineering and Computing, 50 (11) P:1137-1145, Springer, 2012.

[77] J. Proakis, and D. Manolakis, “Digital Signal Processing: Principles, Algorithms and Applications”, Pearson Prentice Hall, 4th edition, 2007.

[78] D. Tkach, H. Huang, and T. Kuiken, “Study of Stability of Time-domain Features for Electromyographic Pattern Recognition”, Journal of NeuroEngineering and Rehabilitation, Springer, Vol. 7, 2010.

[79] S. Sun, and C. Zhang, “Adaptive Feature Extraction for EEG Signal Classification”, Journal of Medical & Biological Engineering and Computing, Springer, Vol. 44, 2006

[80] P. Durka, H. Klekowicz, K. J., W. Blinowska, and S. Niemcewicz. “A Simple System for Detection of EEG Artifacts in Polysomnographic Recordings”, Journal of IEEE Transactions on Biomedical Engineering, Vol. 40, 2003.

[81] A. Tzallas, M. Tsipouras, and D. Fotiadis, “Epileptic Seizure Detection in Electroencephalograms using Time-Frequency Analysis,” IEEE Transactions on Information Technology in Biomedicine, Vol. 13, September 2009.

[82] W. Zhou, Y. Liu, Q. Yuan, and X. Li, “Epileptic Seizure Detection Using Lacunarity and Bayesian Linear Discriminant Analysis in Intracranial EEG”, Journal of IEEE Transactions on Biomedical Engineering, Vol. 60, December 2013.

[83] C. Vidaurre, N. Kramer, B. Blankertz, and A. Schlogl, “Time Domain Parameters as a Feature for EEG-based Brain Computer Interfaces,” Journal of Neural Networks, Elsevier, Vol. 22, November 2009.

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

[85] P. Mirowski, D. Madhavan, Y. LeCun, R. Kuzniecky, “Classification of Patterns of EEG Synchronization for Seizure Prediction”, Electroencephalography and Clinical Neurophysiology, Vol. 120, Elsevier 2009.

[86] M. Fakhr, M. Moshrefi-Torbati, M. Hill, C. M. Hill, P. R. White, “Signal Processing Techniques Applied to Human Sleep EEG Signals—A review Shayan”, Journal of Biomedical Signal Processing and Control 10, Elsevier, 2014.

[87] N. Sadati, H. Mohseni, and A. Maghsoudi, “ Epileptic Seizure Detection using Neural Fuzzy Networks,” IEEE International Conference on Fuzzy Systems, Vancouver, BC, Canada, July 16-21, 2006.

[88] M. Pregenzer and G. Pfurtscheller, “Frequency Component Selection for an EEG-Based Brain to Computer Interface,” Journal of IEEE Transactions on Rehabilitation Engineering, Vol. 7, December 1999.

[89] Y. Song, P. Liò, “A New Approach for Epileptic Seizure Detection: Sample Entropy Based Feature Extraction and Extreme Learning Machine”, Journal of Biomedical Science and Engineering, Vol. 3, Elsevier, 2010.

[90] L. Guo, D. Rivero, J. Dorado, J. R. Rabunal, A. Pazos, “Automatic Epileptic Seizure Detection in EEGs Based on Line Length Feature and Artificial Neural Networks”, Journal of Neuroscience Methods, Elsevier, 2010.

[91] M. Mercy, “Performance Analysis of Epileptic Seizure Detection Using DWT & ICA with Neural Networks”, International Journal of Computational Engineering Research, Vol. 2, 2012.

[92] S. Liang, H. Wang, and W. Chang, “Combination of EEG Complexity and Spectral Analysis for Epilepsy Diagnosis and Seizure Detection,” EURASIP Journal Advanced Signal Processing, Springer, Vol. 2010, 2010.

[93] W. Weng, and K. Khorasani, “An Adaptive Structure Neural Network with Application to EEG Automatic Seizure Detection”, Journal of Neural Network, Elsevier, Vol. 9, 1996.

[94] N. Pradhan, P. Sadasivan, and G. Arunodaya, “Detection of Seizure Activity in EEG by an Artificial Neural Network: A Preliminary Study,” Journal of Computer And Biomedical Research, Vol. 29, 1996.

[95] V. Nigam, and D. Graupe, “A Neural-network-based Detection of Epilepsy”, Journal of Neurological Research, Vol. 26, pp. 55–60, 2004.

[96] A. Ahangi, M. Karamnejad, N. Mohammadi, R. Ebrahimpour, N. Bagheri, “Multiple Classifier System for EEG Signal Classification with Application to Brain–Computer Interfaces”, Journal of Neural Computer & Applications, Springer, 2012.

[97] A. Subasi, E. Ercelebi, “Classification of EEG Signals using Neural Network and Logistic Regression”, Journal of Computer Methods and Programs in Biomedicine, Elsevier, 78, 2005.

[98] U. Orhan, M. Hekim, M. Ozer, “EEG Signal Classification using the K-means Clustering and a Multilayer Perceptron Neural Network Model”, Journal of Expert Systems with Applications, Elsevier, 38, 2011.

[99] S. Shafiul Alam and M. Bhuiyan, “Detection of Seizure and Epilepsy Using Higher Order Statistics in the EMD Domain”, Journal of IEEE Journal of Biomedical and Health Informatics, Vol. 17, March 2013.

[100] Y. Liu, W. Zhou, Q. Yuan, and S. Chen, “Automatic Seizure Detection Using Wavelet Transform and SVM in Long-Term Intracranial EEG,” Journal of IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 20, November 2012.

[101] D. Rivero, E. Fernandez-Blanco, J. Dorado, A. Pazos (2011a), “A New Signal Classification Technique by Means of Genetic Algorithms and k-NN,” IEEE Congress on Evolutionary Computation (CEC), pp: 581–586, 2011.

[102] N. Ba-Karait, S. Shamsuddin, and R. Sudirman, “EEG Signals Classification Using a Hybrid Method Based on Negative Selection and Particle Swarm Optimization”, Hournal of Machine Learning and Data Mining in Pattern Recognition, Springer, 2012.

[103] M. Arab, A. Suratgar, A. Ashtiani, “Electroencephalogram Signals Processing for Topographic Brain Mapping and Epilepsies Classification”, Journl of Computers in Biology and Medicine 40, Elsevier, 2010.

[104] R. Sharma, R. Pachori, and S. Gautam, “Empirical Mode Decomposition Based Classification of Focal and Non-Focal EEG Signals”, in the proc. IEEE International Conference on Medical Biometrics, Shenzhen, China, 30 May-01 June, 2014.

[105] R. Pachori and S. Patidar, “Epileptic Seizure Classification in EEG Signals using Second-Order Difference Plot of Intrinsic Mode Functions”, Computer Methods and Programs in Biomedicine, Vol. 113, Elsevier, February 2014.

[106] R. Pachori and V. Bajaj, “Analysis of Normal and Epileptic Seizure EEG Signals using Empirical Mode Decomposition”, Computer Methods and Programs in Biomedicine, Vol. 104, Elsevier, December 2011.

[107] R. Pachori, “Discrimination between Ictal and Seizure-free EEG Signals using Empirical Mode Decomposition”, Research Letters in Signal Processing, Vol. 2008, Article ID 293056, 5 pages, Elsevier, December 2008.

[108] V. Joshi, R. Pachori, and A. Vijesh, “Classification of Ictal and Seizure-free EEG Signals using Fractional Linear Prediction”, Biomedical Signal Processing and Control, Vol. 9, pp. 1-5, Elsevier, January 2014.

[109] M. Yu, S. Naqvi, A. Rhuma and L. Chambers, “Fall Detection in A Smart Room by Using A Fuzzy One Class Support Vector Machine and Imperfect Training Data,” IEEE International Conference on Acoustics, Speech and Signal Processing, 2011.

[110] H. He, and Y. Cao, “SSC: A Classifier Combination Method Based on Signal Strength”, Journal of IEEE Transactions on Neural Networks and Learning Systems, Vol. 23, July 2012.

[111] Y. Yang, X. Liu, “A Re-examination of Text Categorization Methods”, Annual ACM Conference on Research and Development in Information Retrieval, USA, 1999.

[112] J. Han, M. Kamber, “Data Mining: Concepts and Techniques”, Elsevier, Second Edition, 2006.

[113] D. Rivero L. Guo J.A. Seoane J. Dorado, “Using Genetic Algorithms and K-Nearest Neighbor for Automatic Frequency Band Selection for Signal Classification”, IET Signal Processing, 2012.

[114] R. Hussein, A. Mohamed, and M.Alghoniemy, “Adaptive Compression and Optimization for Real-Time Energy-Efficient Wireless EEG Monitoring Systems”, published in the Biomedical Engineering International Conference (BMEiCON-2013), 2013.

[115] J. Peña, “Finding Consensus Bayesian Network Structures”, Journal of Artificial Intelligence Research, Vol. 42, 2011.

[116] S. Sun, C. Zhang, D. Zhang, “An Experimental Evaluation of Ensemble Methods for EEG Signal Classification”, Journal of Pattern Recognition Letters, 28, Elsevier 2007.

[117] H. Bostrom, R. Johansson, A. Karlsson, “On Evidential Combination Rules for Ensemble Classifiers”, 11th International Conference on Information Fusion, June 30 -July 3 2008.

[118] S. Dehuri, A. Kumar Jagadev, and S. Cho, “Epileptic Seizure Identification from Electroencephalography Signal using DE-RBFNs Ensemble”, Journal of Procedia Computer Science 23, Elsevier, 2013.

[119] H. Lu, H. Eng, C. Guan, K. Plataniotis, and A. Venetsanopoulos, “Regularized Common Spatial Patterns with Aggregation for EEG Classification in Small-Sample Setting”, Journal of IEEE Transactions on Biomedical Engineering, Vol. 57, 2010.

[120] C. Chen, “A Survey on Sub-Nyquist Sampling”, Journal on Bioinformatics and Systems Biology, 2009.

[121] J. Haupt, W. Bajwa, M. Rabbat, and R. Nowak, “Compressed Sensing for Networked Data”, Journal of IEEE Signal Processing Magazine, - Special Issue on Compressing Sensing, Vol.25, pp. 92-101, March 2008.

[122] S. Mallat, “A Wavelet Tour of Signal Processing”, Handbook, Second Edition (Wavelet Analysis & Its Applications), Elsevier, 2009.

[123] N. Timmons and W. Scanlon, “An Adaptive Energy Efficient MAC Protocol for the Medical Body Area Network”, 1st International Conference on Wireless Communication, Wireless VITAE, Denmark, May 17-20, 2009.

[124] M. Unser, “Sampling—50 Years after Shannon”, Proceedings of the IEEE, Vol. 88, 2000.

[125] Inverse Discrete Cosine Transform (iDCT): http://www.mathworks.com/help/signal/ref/idct.html

[126] R. 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", Physical Review E Journal, American Physical Society , Vol. 64, 2001.

[127] S. Kumar, N. Sriraam, P. Benakop, B. Jinaga, "Entropies Based Detection of Epileptic Seizures with Artificial Neural Network Classifiers," Expert System Application, 37, 3284–3291, Elsevier, 2010.

[128] K. Abualsaud, M. Mahmuddin, M. Saleh, A. Mohamed, "Ensemble Classifier for Epileptic Seizure Detection for imperfect EEG Signals", in The Scientific World Journal, Hindawi Publishing Corporation, Vol. 2015, Article ID: 945689, 15 pages, 2015.

[129] H. Ocbagabir, K. Aboalayon, and M. Faezipour, "Efficient EEG Analysis for Seizure Monitoring in Epileptic Patients", IEEE Long Island Systems, Applications and Technology Conference (LISAT), Farmingdale, NY, 3 May 2013.

[130] P. Khan, M. Hussain, and K. Kwak, "Medical Applications of Wireless Body Area Networks", International Journal of Digital Content Technology and its Applications, Vol. 3, September 2009.

[131] A. Awad, A. Mohamed, A. A. El-Sherif, O. A. Nasr, "Interference-aware Energy-efficient Cross-layer Design for Healthcare Monitoring Applications", Elsevier, Computer Networks 74 (2014) 64–77, 2014.

[132] P. Durka and K. Blinowska, "A Unified Time Frequency Parametrization of EEGs," Journal of IEEE Engineering in Medicine and Biology, Vol. 20, pp. 47–53, September/October 2001.

[133] S. Aviyente, E. Bernat, S. Malone, and W. Iacono, "Analysis of Event Related Potentials using PCA and Matching Pursuit on the Time-Frequency Plane," in IEEE International Conference of the Engineering in Medicine and Biology Society, Vol. 3, pp. 424–427, 2006.

[134] Data Mining, Machine Learning Software, http://rapid-i.com/

[135] S. Koco, C. Capponi, "On Multi-Class Classification through the Minimization of the Confusion Matrix Norm", JMLR: Workshop and Conference Proceedings 29, ACML 2013.

Appendix A

Khalid Abualsaud, Massudi Mahmuddin, Mohammad Saleh, Amr Mohamed, "Ensemble Classifier for Epileptic Seizure Detection for imperfect EEG Signals", in *The Scientific World Journal, Hindawi Publishing Corporation*, Vol. 2015, Article ID: 945689, 15 pages, 2015.

Appendix B

Khalid Abualsaud, Massudi Mahmuddin, Mohammad Saleh, Amr Mohamed "Performance Comparison of Classification Algorithms for EEG-based Remote Epileptic Seizure detection in Wireless Sensor Networks" in *The ACS/IEEE International Conference on Computer Systems and Applications* (*AICCSA*), 2014.

Appendix C

Khalid Abualsaud, Massudi Mahmuddin, Ramy Hussein, Amr Mohamed, "Performance Evaluation of Compression-Accuracy trade-off Using Compressive Sensing for EEG-based Epileptic Seizure Detection in Wireless Tele-monitoring", in the proceeding of *The 9th IEEE International Wireless Communications and Mobile Computing Conference (IWCMC 2013)*, Sardinia – Italy, July 1-5, 2013.

Appendix D

Khalid Abu Al-Saud, Massudi Mahmuddin, Amr Mohamed, "WBASN signal processing and communication framework: Survey on Sensing, Communication Technologies, Delivery and feedback", in *Journal of Computer Science (JCS),* New York, USA, vol. 8, Issue 1, 2012**.**

Appendix E

Khalid Abu Al-Saud, Amr Mohamed and Massudi Mahmuddin, "Survey on Wireless Body Area Sensor Networks for healthcare applications: Signal Processing, data analysis and feedback", in the Proceedings of *The 3rd International Conference on Computing and Informatics, ICOCI2011*, 8-9 June, 2011 Bandung, Indonesia.