Non-invasive Detection of Hypoglycemia in Patients with Type 1 Diabetes using Electroencephalography Signals

---

by

Bich Lien Nguyen

Submitted to the Faculty of Engineering and Information Technology

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

at the University of Technology, Sydney

Sydney, May 2014

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree, except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Candidate

Bich Lien Nguyen

Sydney, May 2014

Acknowledgements

First of all, I would like to express my sincere gratitude and acknowledgement to my supervisor, Professor Hung Tan Nguyen, for his continuous and thoughtful guidance and support during my PhD studies. His expertise, knowledge and spirit are always a great inspiration to me, not only for my doctoral research, but also for my future career.

I am very grateful to my co-supervisor, Dr. Steve Ling for his valuable suggestions and supports in improving my research.

Many thanks to Ms. Lesley Nguyen for her great assistance in editing this thesis.

Special appreciation to my father, my mother, my aunt, my older sister and my younger brother who always encouraged me during my stay in Sydney.

Deepest thanks to my dearest husband, Mr. Van Anh Nguyen, for the incessant support and unconditional love he gives me in the past four years as well as the rest of our life.

Finally, this thesis is dedicated to my beloved little daughter, Gia Han Nguyen. Her presence in the last year of my studies provided me great strength to go to the end of my PhD journey.

Bich Lien Nguyen

Sydney, May 2014

Contents

List of Figures

List of Tables

Abstract

For patients with type 1 diabetes mellitus (T1DM), hypoglycemia, or the state of abnormally low blood glucose level (BGL), is the most common but acute complication which limits their intellectual as well as physical activities. Mild hypoglycemic episodes cause sweating, nervousness, heart plumping, confusion, anxiety, etc. which can be fixed by eating or drinking glucose-rich food. However, if left untreated, severe episodes of hypoglycemia may lead to unconsciousness, coma, or even death. Nocturnal episodes of hypoglycemia are especially dangerous because sleep reduces and obscures early symptoms, so that an initially mild episode may become severe. Because of its severity, it is essential for T1DM patients to be monitored and alarmed whenever a hypoglycemic episode occurs, especially during the night.

For the purpose of hypoglycemia detection, using parameters extracted from the electroencephalogram (EEG) is one of the most promising methods. Because it depends on a continuous supply of glucose and is vulnerable to any glucose deprivation, the human brain is one of the first affected organs under the occurrence of hypoglycemia. Since the EEG is directly related to the metabolism of brain cells, a failure of cerebral glucose supply can cause early changes in EEG signals that can be utilized in hypoglycemia detecting devices.

The main aim of this thesis is to develop a computational methodology of non-invasively detecting the onset of nocturnal hypoglycemia for patients with T1DM from their EEG signals. There are two core tasks to be implemented: feature extraction and classification.

Feature extraction analyses a variety of EEG parameters to find features that significantly respond to the onset of hypoglycemia. Important features will be used as inputs of the classification in order to classify and detect hypoglycemic episodes.

Using raw EEG signals collected at four EEG channels (C3, C4, O1, O2) from five T1DM patients who participated in an overnight hypoglycemia-induced study, four EEG parameters (power level, centroid frequency, spectral variance, spectral entropy) within three frequency bands (theta, alpha, beta) are extracted by spectral analysis. Statistical analysis is applied to find parameters that significantly correlate to the transitions of patients' states during the study, from normal to hypoglycemic and then to recovery state. The statistical results show that under hypoglycemic conditions, there are early changes in the theta and alpha bands of EEG signals. The decrease in centroid alpha frequency is the most significant feature which is consistently observed in all patients at all EEG channels ($p<0.0001$). Besides, by analysing the data from the BGL range of 3.3-3.9 mmol/l, it is established that the EEG responses to hypoglycemia only significantly occur when patients' BGLs fall to the threshold of 3.3 mmol/l. This threshold is used to distinguish between hypoglycemic state and non-hypoglycemic state for the classification purpose in this thesis. As a result of the feature extraction, two EEG features of centroid theta frequency and centroid alpha frequency are derived at two channels C3 and O2 to be used as inputs of the classification.

In terms of classification algorithm, in this thesis, the standard multi-layer feed-forward neural network is utilised as the classification unit. Three different training techniques are applied to train the developed neural network, including the LM algorithm, the LM+GA algorithm, and the LM+GA+Adaptive algorithm. The LM algorithm is based on the popular Levenberg-Marquardt algorithm and the cross-validation technique in order to direct the training process to one of local optimal solutions. The LM+GA algorithm includes two consecutive steps of global search and local search. The global search is based on a genetic algorithm which helps the training process direct to the area of the global optimal solution. The local search is based on the LM algorithm which acts as a fine tuner to get the training process closer to the final optimised solution. Lastly, the LM+GA+Adaptive algorithm is introduced, as the final neural network training procedure

proposed by this thesis. This algorithm consists of two consecutive stages including an adaptive training stage implemented after the GA+LM algorithm. The stage of adaptive training helps to adapt the optimised network yielded by the GA+LM algorithm to the new EEG patterns of unseen subjects, thus limiting effects of the EEG variability on classification results and enhancing the generalisation ability of the developed neural network.

The final classification results produced by this thesis (80% sensitivity and 60% specificity on the training set and 78% sensitivity and 62% specificity on the testing set) indicate that by applying the proposed computational methodology, nocturnal episodes of hypoglycemia can be successfully detected in patients with T1DM from their non-invasive EEG signals. With the final performance achieved by this thesis, future works will be carried out to pursue the final aim of the current research which is developing a non-invasive EEG-based system for detecting hypoglycemia that can be applied into the real clinical environment.

Chapter 1

Introduction

1.1 Problem statement

Diabetes mellitus is recognized as one of the most challenging global health problems. Type 1 diabetes mellitus (T1DM) is a form of diabetes mellitus which is caused by the loss of insulin-producing beta cells in the pancreas leading to insulin deficiency. T1DM is a chronic condition and can cause a variety of serious complications for patients. The Diabetes Control and Complication Trial (DCCT) Research Group in 1993 emphasized the significant benefits of the intensive insulin therapy for T1DM patients. Results of the report showed that the therapy for a mean of six years (as opposed to conventional therapy) efficiently delayed the appearance as well as reduced the risk of retinopathy by 47%, nephropathy by 54% and neuropathy by 60%. However, it was also highlighted that patients who participated in the DCCT experienced a threefold-increase incidence of severe hypoglycemia episodes over those receiving conventional therapy. This is considered as the most common and acute complication for T1DM patients and a barrier which limits the glycemic control therapy for diabetes patients.

Hypoglycemia is the medical term for the state produced by a lower than normal level of blood glucose. A hypoglycemic episode can be defined as one in which a patient has a

blood glucose level (or glycemic level) lower than 60 mg/dl (3.3 mmol/l). Episodes of hypoglycemia are considered as a fact of life for most patients with T1DM who owe their lives to insulin to maintain a normal range of blood glucose level. Hypoglycemia can produce a variety of symptoms, from mild to severe episodes (Clarke et al. 2009; Klonoff 2001). Mild hypoglycemia causes symptoms like sweating, nervousness, heart plumping, confusion, anxiety, etc. It can be treated quickly and easily by eating or drinking glucose-rich food to restore the blood glucose level to normal. If left untreated, a mild episode of hypoglycemia can become a severe episode in which the patient needs assistance to treat the event. More seriously, it can lead to seizures, coma, and even death. Hypoglycemia also reduces the quality of life of patients as well as caregivers by causing chronic anxiety about the future potential hypoglycemic episodes (Warren & Frier 2005).

One of the most dangerous effects of hypoglycemia is hypoglycemic unawareness which is caused by the frequent exposure to hypoglycemia. In this situation, patients' bodies do not release counter-regulatory hormones which are the origin of early warning symptoms for patients like shaking, sweating, hunger, anxiety, etc. Because of the lack of warning, patients cannot realize the occurrence of hypoglycemia until it becomes severe and could lead to fatal damage.

Nocturnal hypoglycemia is also especially fearful for T1DM patients because sleep can make the symptoms unclear. Nocturnal hypoglycemia is common in patients with T1DM and usually asymptomatic. It was reported previously that almost 50% of all episodes of severe hypoglycemia occur at night during sleep (Group 1991). Such episodes can cause convulsions and coma, and have been implicated as a precipitating factor in cardiac arrhythmias resulting in sudden death--the "dead-in-bed syndrome" (Sovik & Thordarson 1999). Recurrent exposure to nocturnal hypoglycemia may gradually impair patients' cognitive function, as well as lead to other substantial long-term morbidities including the development of acquired hypoglycemic syndromes, such as impaired awareness of hypoglycemia, through the putative effect of unsuspected recurrent episodes of nocturnal hypoglycemia.

Because of its prevalence and severity, a variety of studies have been carried out, using

different techniques to produce systems that can monitor hypoglycemic conditions in T1DM patients. Some of them require intermittently taking patients' blood samples to monitor the blood glucose levels during the day. This method gives relatively exact information about hypoglycemic status. However, taking blood is uncomfortable for patients, and very inconvenient to monitor continuously, especially during the night. A non-invasive technique is obviously a better solution for these disadvantages. Continuous glucose monitoring systems (CGMSs) use different techniques which allow monitoring the blood glucose levels continuously and providing better information about glycemic shifting throughout the day. However, the low accuracy of CGMSs is a prominent disadvantage of the technique. Also, these systems still involve the minimally invasive procedure of inserting and changing sensors over time. Currently, on the market, there are some devices which monitor hypoglycemia non-invasively by using patients' physiological parameters such as heart rate, skin impedance and electrocardiography (ECG) outputs. It is no doubt that exploring and developing new techniques to monitor and detect the onset of hypoglycemia for T1DM patients is still an open research area which has drawn the attention of researchers around the world.

Electroencephalogram (EEG) is the recording of electrical activity along the scalp of the human brain. Recently, EEG signals have been shown to be a powerful tool for diagnosing and detecting various diseases and conditions. Under the occurrence of hypoglycemia, the human brain is one of the first affected organs. Because it cannot synthesize as well as store this primary metabolic fuel, the brain depends on a continuous supply of glucose and is vulnerable to any glucose deprivation. In previous works, the EEG, which is directly related to the metabolism of brain cells, was shown to have early responses to the onset of hypoglycemia (Howorka et al. 1996; Pramming et al. 1988; Tallroth et al. 1990; Tribl et al. 1996). This research direction leads to the essential core of this thesis which is exploring early changes in EEG parameters induced by hypoglycemic episodes which can be detected by an advanced computational algorithm.

Inspired by the human central nervous system, neural networks have attracted many researchers in a wide range of research areas. Neural networks have been successfully applied to various classification and pattern recognition problems in industrial as well as

real-life sectors. In the health and biomedical sectors, neural networks have also been widely used as an effective diagnosis and detection technique. In order to avoid inherent shortcomings of standard neural networks, different strategies of network training will be explored to determine the most suitable network for the application of hypoglycemia detection based on EEG signals.

1.2 Objectives of the thesis

The main objective of this thesis is to develop a computational methodology of detecting hypoglycemic episodes non-invasively for patients with T1DM using their EEG signals. In order to achieve this purpose, two main tasks will be focused upon.

- First, EEG signals from T1DM patients will be processed to extract important features which significantly respond to the occurrence of hypoglycemic episodes. This involves a signal processing step to get rid of unwanted artifacts from the raw signals and then transform EEG signals from time domain to frequency domain to get the power spectra of the signal. Based on the power spectra, a feature extraction step will be implemented to derive various EEG parameters. The extracted parameters will be analysed by statistical techniques to find important features which can significantly contribute to the performance of the classification.

- After finding potential EEG features, an advanced classification algorithm will be developed for the purpose of detecting hypoglycemia, using the extracted features as inputs. At first phase, a standard multi-layer feed-forward neural network will be developed and utilised as the classification unit to verify the potentiality of detecting hypoglycemia using extracted features derived from EEG signals. Then, different advanced strategies and algorithms for training the developed neural network will be explored in order to enhance the general performance of the developed neural network.

1.3 Thesis contributions

This thesis presents a comprehensive computational methodology for non-invasive detection of nocturnal hypoglycemic episodes from EEG signals for patients with type 1 diabetes. The five main contributions of the thesis are presented as follows:

- Firstly, the thesis presents a thorough analysis of EEG signals from type 1 diabetes adolescents during an insulin-induced study to investigate responses of EEG signals to nocturnal hypoglycemia. By exploring various EEG parameters of patients over three phases of the study including normal, hypoglycemia and recovery, it is shown that EEG signals are highly correlated with patients' conditions during the study. A comprehensive literature review demonstrates that until now, most previous studies exploring the relationship between the EEG and hypoglycemia only stopped at using peak values at each frequency bin of the EEG power spectra without any trend to extract meaningful features. The point of exploring various EEG features and proving that they significantly respond to hypoglycemic conditions by statistical analysis makes an important contribution of this thesis.

- Secondly, the thesis provides a comparison between two blood glucose levels of 3.3 mmol/l and 3.9 mmol/l in order to investigate a blood glucose threshold at which the EEG signals of T1DM patients starts to respond to hypoglycemic conditions. This threshold will be used later to distinguish between non-hypoglycemic and hypoglycemic states for the purpose of classification.

- Thirdly, the introduction of two EEG features of centroid alpha frequency and centroid theta frequency in the application of non-invasively detecting hypoglycemia from EEG signals is another significant contribution of this thesis. As shown in the literature review, previous works which also aimed to develop systems for detecting hypoglycemia from EEG signals only stopped at using peak values at each frequency bin of the EEG power spectra. This method can lead to very noisy inputs for the purpose of classification. In this thesis, it will be demonstrated that the two EEG parameters of centroid theta frequency and centroid alpha frequency are

the most important extracted features which can significantly contribute to the performance of the hypoglycemia detection algorithm.

- Fourthly, the thesis proposes a computational algorithm based on neural network for the purpose of classification and detection of hypoglycemia. Using extracted EEG features as inputs, feed-forward multi-layer neural networks are developed as classification units to identify episodes of nocturnal hypoglycemia. The data from five T1DM patients will be used to develop and validate the performance of the proposed classification algorithms.

- Fifthly, with the aim of enhancing the classification performance, a combination of genetic algorithm (GA) and Levenberg-Marquardt (LM) algorithm, named the GA+LM algorithm, will be proposed for training the developed neural networks. In this way, the advantages of each algorithm, including the global search ability of GA algorithm and the local search ability of LM algorithm, can be utilized in order to direct the network training process to the global optimal without trapping into one of the local solutions. The thesis demonstrates that by applying a properly combined strategy to train neural network, the performance of hypoglycemia detection using only two EEG channels can be improved markedly.

- Lastly, in order to overcome one of well-known limitations of using EEG signals in healthcare application which is the signal variability from person to person to person, an adaptive strategy of training neural network will be introduced. The proposed adaptive training strategy will be proved to allow the neural network to adapt itself to a new individual user and help to enhance the generalisation ability of the hypoglycemia detection from EEG signals. Implementing the proposed adaptive strategy in conjunction with the GA+LM algorithm, the GA+LM+Adaptive algorithm will be introduced as the final procedure proposed by this thesis for training neural network in the application of detecting hypoglycemia using EEG signals from only two channels.

1.4 The structure of the thesis

This thesis consists of six chapters, a bibliography and appendices. The remaining chapters of the thesis are organized as follows:

- ***Chapter 2*** presents a comprehensive literature review of various techniques for detecting the onset of hypoglycemic episodes in patients with type 1 diabetes. First, the prevalence and severity of type 1 diabetes and its most dangerous complication, hypoglycemia, will be mentioned to show the inspiration of this thesis. After that, common techniques and devices of glucose monitoring and hypoglycemia detecting which are currently available on the market, as well as the positives and drawbacks of each technique will be presented. Finally, this chapter provides a review of using EEG signals and computational intelligence, generally in biomedical systems and specifically in the application of detecting hypoglycemia.

- ***Chapter 3*** proposes a computational methodology for EEG-based hypoglycemia detection. This chapter consists of two main parts: feature extraction and classification. In the feature extraction part, using data of five T1DM patients from an overnight insulin-induced study (also called glucose clamp study), EEG signals from four EEG channels including C3, C4, O1 and O2 are processed and analysed by spectral analysis to extract various EEG parameters. The main aim of this part is to explore the response of extracted parameters during the clamp study in order to determine the most important features that significantly change under hypoglycemic conditions. In order to do this, the correlation of each EEG parameter with the BGL transition during the insulin-induced study including three phases of Normal, Hypoglycemia and Recovery will be explored. Also, data at the BGL range of 3.3-3.9 mmol/l (named as the Early Onset phase of the study) will be analysed to find a blood glucose threshold to distinguish between non-hypoglycemic and hypoglycemic states.

  Using extracted EEG parameters from all four channels as inputs, a standard neural network algorithm is introduced for the purpose of classification. The developed

neural network has a feed-forward multi-layer structure which is trained by using Levenberg-Marquardt (LM) algorithm and cross-validation technique. Classification performance will be presented to evaluate the capability of the detection of hypoglycemic episodes from EEG signals.

- ***Chapter 4*** aims to propose an advanced algorithm for training neural network in order to enhance the performance of the developed classification algorithm. To do this, a combination of genetic algorithm (GA) and the LM algorithm will be explored to utilize advantages as well as avoid limitations of each algorithm in training neural network. The GA algorithm is used to locate the region of the global optimal consistently. The LM algorithm acts as a fine tuner to help the training process quickly converges toward the global solution. Our main objective is to demonstrate that by applying a properly combined strategy to train neural network, the performance of hypoglycemia detection using only two EEG channels can be improved markedly.

- ***Chapter 5*** introduces an adaptive strategy for training neural network in order to improve the generalisation ability of the developed neural network. It has been noted in previous works that EEG patterns significantly vary from person to person, which leads to a difficulty in generalising an EEG-based system to a new user. To overcome this, a strategy of training neural network adaptively, which allows the neural network to adapt itself to each new individual user, will be implemented. The results will be presented to demonstrate that by applying partly individual training, the hypoglycemia detecting system which uses only two EEG channels can perform efficiently.

Based on the performance of algorithms and strategies for training neural networks which have been explored so far in this and previous chapters, the final network training procedure for hypoglycemia detection using EEG signals proposed by this thesis will be presented. This training procedure, named the GA+LM+Adaptive algorithm, consists of two sequential stages including the GA+LM algorithm implemented in conjunction with the adaptive strategy. Classification results of this

training procedure provided at the end of this chapter are considered as the final performance of the computational methodology of hypoglycemia detection using only 2 EEG channels developed by this thesis.

- ***Chapter 6*** presents the overall conclusions for this research. Advantages as well as limitations of the proposed computational methodology for non-invasively detecting hypoglycemia in patients with type 1 diabetes using EEG signals will be mentioned in this chapter. Finally, the potentiality and different directions of pursuing the research in future works will be discussed.

1.5 Publications related to the thesis

There are four fully refereed international conference papers of the Institute of Electrical and Electronics Engineers (IEEE) related to the thesis. Details about these papers are provided as follows:

- **L.B. Nguyen**, S.S.H. Ling, T.W. Jones & H.T. Nguyen 2011, 'Identification of hypoglycemic states for patients with T1DM using various parameters derived from EEG signals', *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, Massachusetts, USA, pp. 2760-3.

- **L.B. Nguyen**, A.V. Nguyen, S.H. Ling & H.T. Nguyen 2012, 'An adaptive strategy of classification for detecting hypoglycemia using only two EEG channels', *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, USA, pp. 3515-8.

- **L.B. Nguyen**, A.V. Nguyen, L. Sai Ho & H.T. Nguyen 2013, 'Analysing EEG signals under insulin-induced hypoglycemia in type 1 diabetes patients', *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* Osaka, Japan, pp. 1980-3.

- **L.B. Nguyen**, A.V. Nguyen, L. Sai Ho & H.T. Nguyen 2013, 'Combining genetic

algorithm and Levenberg-Marquardt algorithm in training neural network for hypoglycemia detection using EEG signals', *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Osaka, Japan, pp. 5386-9.

Chapter 2

Literature Review

2.1 Introduction

Diabetes mellitus is recognized as a challenging global health problem. Untreated diabetes can cause many long-term severe complications including damages to the heart, blood vessels, eyes, kidneys, and nerves. It can also lead to acute complications such as diabetic ketoacidosis, hyperosmolar hyperglycemic state and hypoglycemia which have the potential to lead to cognitive impairments, coma, or even progress to death.

Type 1 diabetes mellitus (T1DM) is a form of diabetes that results from an autoimmune destruction of insulin-making cells in the pancreas. T1DM typically occurs in people under 30 years old and is the major form of diabetes in those under 10 years old. The cause of T1DM is not known and it is not preventable with current knowledge. Patients with T1DM depend on external insulin for their survival and the treatment must be continued indefinitely in all cases.

The Diabetes Control and Complication Trial Research Group highlighted significant benefits of the intensive insulin therapy, include efficiently delaying the onset as well as reducing the risk of acute diabetic complications (Group 1993). However, the main

adverse effect associated with intensive therapy is a two-to-threefold increase of severe hypoglycemia which is the medical term for the state of an abnormally low blood glucose level. Hypoglycemia is considered as the most acute but common complication of T1DM which can lead to unconsciousness, coma, or even death if left untreated. Obviously this is a significant barrier to the achievement of desired glycemic control in the insulin-treated diabetic patients.

Under the occurrence of hypoglycemia, the human brain is one of the first affected organs. Because it cannot synthesize as well as store this primary metabolic fuel, the brain depends on a continuous supply of glucose and is vulnerable to any glucose deprivation. The electroencephalogram (EEG), which is directly related to the metabolism of brain cells, is shown to have early responses to the onset of hypoglycemia. Efficiently detecting early changes in the EEG under hypoglycemia conditions are important towards developing a hypoglycemia monitoring system which can alarm patients and caregivers at onset.

2.2 Type 1 diabetes mellitus and the intensive insulin therapy

Diabetes mellitus is a chronic condition in which a patient's body cannot maintain the normal blood glucose levels. The human body requires a hormone called insulin, which is produced by the pancreas, for stimulating the body's cells to convert glucose (sugar) from food into energy. In people with diabetes mellitus, insulin is no longer produced in sufficient amounts, or the body cells do not respond adequately to the hormone. So in these cases, when diabetes patients eat glucose, instead of being turned into energy, the glucose stays in the blood, causing higher blood glucose levels than normal people.

Type 1 diabetes mellitus (T1DM) is one of the main forms of diabetes that occurs when the pancreas no longer produces significant amounts of the hormone insulin, owing to the destruction of the insulin-producing beta cells of the pancreas. Accounting for about 5-10% of all diabetes cases (Daneman 2006), or 17.35-34.7 million people worldwide (Organization 2012), T1DM is a serious chronic disorder which may lead to dangerous short-term as well as long-term complications.

T1DM symptoms can include excessive thirst, constant hunger, frequent urination, skin infections or blotches, drowsiness, sudden weight loss. Although the disease onset can occur at any age, T1DM is often referred to as juvenile-onset diabetes with over half of the cases being diagnosed in childhood (younger than 16-18 years old) (Daneman 2006). The incidence of T1DM in this group of ages also increases rapidly with the rate of 3% per year (Henk-Jan Aanstoot 2007). It is also the major diabetes form in those under 10 years old and is one of the leading chronic diseases of childhood among developed countries (Yach 2004). In the period of 2000-2009, it was reported that in Australia, the average incidence rate of type 1 diabetes was 11.5 cases per 100,000 population per annum, in which there were 9,308 new cases of type 1 diabetes among children aged 0–14 years and 13,756 new cases of type 1 diabetes among those aged 15+ years (Welfare 2011).

There are several risk factors that are considered to cause T1DM even though it is still not fully understood. It has been shown in the literature that the presence and combination of some certain genes indicates an increased risk of developing T1DM (Bluestone, Herold & Eisenbarth 2010; Daneman 2006; Poretsky & Ali 2010). Besides, non-genetic factors also play an important role in the development of the disease, including older age of the mother (Flood, Brink & Gleason 1982), first pregnancy (Patterson et al. 1994), breast feeding (Borch-Johnsen et al. 1984), low maternal educational level, low family income (Blom et al. 1989), exposure to some types of virus, chemicals and drugs (Poretsky & Ali 2010), etc. Ongoing research has been devoted to the influences of individual as well as combined factors on the disease onset.

T1DM patients with prolonged, poor glycemic control have increased likelihood of developing vascular complications which later result in the chance that blood vessels in different body organs are completely weakened or blocked. The long-term damage to blood vessels may lead to cardiovascular diseases and heart attacks. Excess glucose can injure the walls of the tiny blood vessels that nourish patients' nerves and cause tingling, burning or pains at the tops of the toes, fingers and gradually spread upwards even until patients lose all sense of feeling in the affected limbs. It also often leads to kidney damage (nephropathy) which, in many cases, results in irreversible end-stage kidney disease. In this situation, patients require dialysis or a kidney transplant for their survival. Retinopathy, in which

blocked or leaky blood vessels in the retina (a light-sensitive layer of tissue at the back of eyes) gradually prevent the light from fully passing through and potentially leading to blindness, is also another serious complication for patients with T1DM.

Unlike the other type of diabetes mellitus which can be efficiently managed by the combinations of oral medications, diet and exercise, the treatment of T1DM involves an indefinite procedure which is required for the whole of the patient's life. To stay alive, people with T1DM must have a constant supply of insulin through injections or an insulin pump, along with attention to dietary management and careful monitoring of blood glucose levels. Untreated T1DM often leads to diabetic ketoacidosis which causes cerebral edema (accumulation of liquid in the brain). This complication is very life-threatening and considered as the most common cause of death in pediatric diabetes (Rosenbloom & Hanas 1996).

In 1993, the Diabetes Control and Complication Trial (DCCT) Research Group highlighted the significant benefits of intensive insulin therapy in glycemic control (the medical term for achieving target level of glycated hemoglobin HbA1c or blood glucose for T1DM patients). Results of DCCT showed that the intensive therapy for a mean of six years (maintaining glycemic levels to a target HbA1c level of 7%) as opposed to conventional therapy (with resultant mean HbA1c level of 9%) significantly lowered the risk for retinopathy by 47%, nephropathy by 54% and for neuropathy by 60% (Group 1993). The results indicated that the intensive therapy efficiently delays the onset and slows the progression of the aforementioned serious complications. Besides, compared to the conventional therapy, this therapy allows greater flexibility of meal times, carbohydrate quantities, and physical activities, which is important in balancing the normal life of patients.

On the other hand, it was shown in the report that the intensive therapy increases three times the incidence of hypoglycemia among patients with T1DM over conventional therapy. For diabetes patients, hypoglycemia is the medical term of the state produced by an abnormally low level of blood glucose. This is considered as the most common but

highly severe complication for patients with T1DM and a serious barrier of the intensive therapy in achieving improved diabetes control.

2.3 Hypoglycemia

Hypoglycemia, or abnormally low blood glucose concentration, is considered as a fact of life for most patients with T1DM who owe their lives to insulin to maintain a normal range of blood glucose level. Being firmly established since the discovery of the insulin therapy for diabetic patients, diabetic hypoglycemia is the most feared complication and a limiting factor of glycemic control for T1DM patients. Hypoglycemia definition varies from study to study due to differences in study purpose and circumstance; participants' age and their period of taking glycemic control; measurement method to determine the concentration of blood glucose; etc. The definition of hypoglycemia occurrence from some prominent research groups are presented in Table 2.1.

Table 2.1: Definition of hypoglycemia

Research Group	Definition	Associated blood glucose concentration	Associated response
DCCT Research Group (1997)	Severe hypoglycemia	Blood glucose level < 50 mg/dl	Requirement of assistance to recover
American Diabetes Association (2005)	Hypoglycemia	Plasma glucose level = 70 mg/dl	
Christopher et al. (2006)	Mild hypoglycemia	Blood glucose level = 55 - 70 mg/dl	No loss of consciousness
	Moderate hypoglycemia	Blood glucose level < 55 mg/dl	No loss of consciousness
	Severe hypoglycemia	Blood glucose level < 70 mg/dl	Loss of consciousness
Cryer (2003, 2007)	Hypoglycemia	Plasma glucose level = 65 - 70 mg/dl	Increased glucagon and epinephrine secretion
		Plasma glucose level = 50 - 55 mg/dl	Neurogenic and neuroglycopenic symptoms occur

Plasma glucose level = blood glucose level * 1.15
A blood glucose value of mg/dl = a blood glucose value of mmol/l * 18

Hypoglycemia develops when the rates of glucose entry into the body's systemic circulation are reduced relative to glucose uptake by tissues. For patients with diabetes mellitus, hypoglycemia is caused by a mismatch between the insulin therapy and the body's physiological demand. An inaccurate combination of insulin dosages, meals, and physical activities can lead to a decrease in patient's blood glucose concentration or a hypoglycemia episode. Cryer et al. (2003) classified the causes of hypoglycemia into six groups which are listed in Table 2.2.

Table 2.2: Causes of hypoglycemia in type 1 diabetes mellitus
(Cryer, Davis & Shamoon 2003; Strachan 2007)

Causes	Examples
Inappropriate insulin injection	- Excessive dose - Inappropriate timing - Inappropriate insulin formulation
Inadequate exogenous carbohydrate	- Misses meal - Overnight fast
Increased carbohydrate utilization	- Exercise
Decreased endogenous glucose production	- Excessive alcohol intake
Increased insulin sensitivity	- Night time - Exercise - Weight loss
Decreased insulin clearance	- Kidney damage

A summary of responses to hypoglycemia as consequences of lowering blood glucose concentration is provided in Figure 2.1 (Wolpert 2007). When the glucose concentration falls to a certain level, hypoglycemic symptoms which reflect effects of hypoglycemia on the patient' body start to appear. These symptoms are normally categorized into two main groups: autonomic group and neuroglycopenic group. Autonomic symptoms arise from the activation of the autonomous central nervous system, including sweating, shaking, heart pounding, hunger, and nervousness. Neuroglycopenic symptoms result from the reduced consumption of glucose by the brain, especially in the cerebral cortex. The neuroglycopenic group includes symptoms such as confusion, tiredness, difficulty thinking, drowsiness, difficulty speaking. Lower blood glucose levels may lead to severe problems of unconsciousness, convulsions, coma, flattening EEG signals and even death.

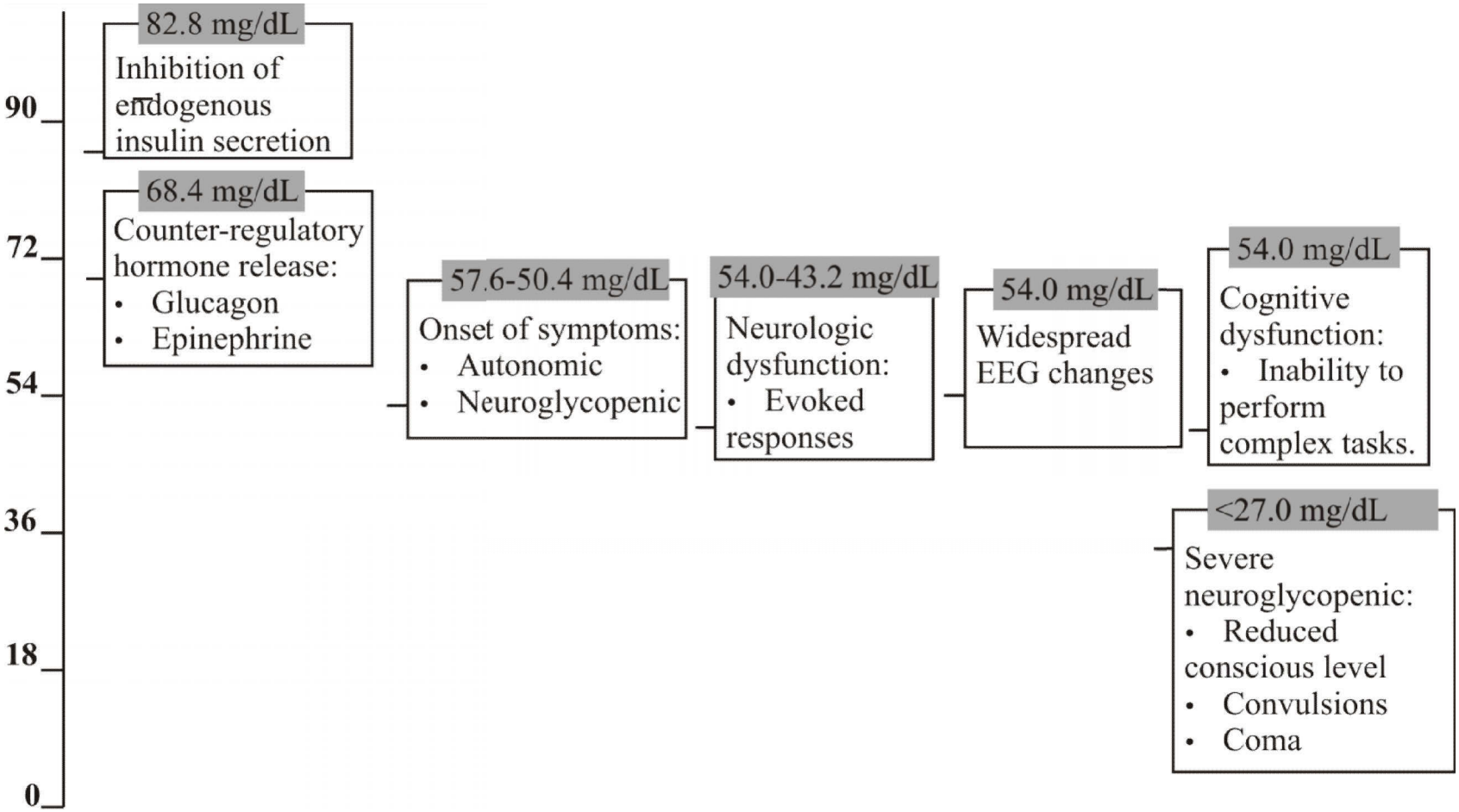


Figure 2.1: Hierarchy of responses to hypoglycemia

[Adopted from (Wolpert 2007)].

In healthy people without diabetes, hypoglycemia is usually corrected naturally by the combination of a number of the human body's defence mechanisms. Initially, a decrease in insulin secretion in response to declining blood glucose levels occurs. As glucose levels continue to fall, a number of redundant glucose counter-regulatory factors (glucagon, epinephrine, growth hormone) are sequentially activated at specific thresholds to ensure sufficient glucose uptake to the brain and other central nervous system tissue metabolism. Besides, a low blood glucose level can activate some early warning symptoms which help people acknowledge the situation and fix it by taking glucose rich food or drink.

Hypoglycemia in diabetes is a special case regarding the ability of the patient's body to amend the situation. In T1DM patients undergoing intensive insulin therapy, falling glucose concentrations often do not elicit counter-regulatory responses at normal glycemic thresholds, allowing glucose levels to drop to dangerously low values. Within the first few years of diabetes, the glucagon secretion in response to hypoglycemia, which stimulates the liver to convert stored glycogen into glucose released into the bloodstream, becomes lost. After the next few years, the epinephrine production in response to hypoglycemia may also be defective and gradually lost. Once the epinephrine response which is the origin of early warning symptoms for the patient is lost, patients normally cannot realize the occurrence of hypoglycemia until it becomes severe and could lead to fatal complications such as coma and convulsions. Studies in T1DM patients have demonstrated that as few as two episodes of antecedent hypoglycemia can blunt responses to subsequent hypoglycemia (Davis & Alonso 2004). When counter-regulatory responses are lost, the risk of a severe hypoglycemic episode is reported to increase at least ten times (Klonoff 2001). This phenomenon is known as hypoglycemia unawareness which is one of the most feared complications for T1DM patients undergoing glycemic control.

Nocturnal hypoglycemia is particularly dangerous for T1DM patients because sleep obscures early warning symptoms, so that an initially mild episode may become severe. The DCCT Research Group in 1991 reported that almost 50% of all episodes of severe hypoglycemia occur at night during sleep. Such episodes can cause convulsions and coma and have been referred to as a factor resulting in the dead-in-bed syndrome in young diabetic patients (Sovik & Thordarson 1999). Even mild episodes of nocturnal

hypoglycemia can lead to a variety of fearful consequences. Nocturnal hypoglycemia seems to have no immediate detrimental effect on cognitive function; however, on the following day, mood and well-being may be adversely affected. Recurrent exposure to hypoglycemic episodes during the night also induces changes in counter-regulatory responses to hypoglycemia, which later leads to impaired awareness of hypoglycemia. Because of the potentially life threatening nature of severe hypoglycemic episodes, hypoglycemia has been shown to reduce the quality of life for patients by causing chronic anxiety about future potential episodes of hypoglycemia.

2.4 Hypoglycemia detection techniques

The danger of hypoglycemic episodes and adverse effects on patients' lives have led to a huge demand for devices that can detect the onset of hypoglycemia and give alarm to provide enough time for patients and their caregivers to take action. A variety of techniques have been developed which can be categorized into two main groups of blood glucose monitoring and physiological glucose monitoring. Patients with T1DM depend on monitoring the concentration of glucose in blood during the day to assess the effectiveness of their insulin therapy, to adjust their diet and exercise plans. Most importantly, it allows the detection of hypoglycemic episodes in T1DM patients, which is especially crucial for those with hypoglycemic unawareness.

2.4.1 Blood glucose monitoring

Blood glucose monitoring is a technique of testing the concentration of glucose in the blood which is particularly important in the care of diabetes mellitus. It includes two different ways of monitoring: intermittent and continuous. Different manufacturers use different technologies which are suitable for the demand and preference of patients based on their different pathological as well as financial conditions. Generally, they are classified into intermittent and continuous monitors.

Intermittent monitoring is commonly performed by piercing the skin (finger) to draw blood, then applying the blood sample to a chemically active disposable test-strip to determine the glucose level in the blood (Figure 2.2). The greatest advantage of this method is that it provides fairly accurate results of blood glucose concentration. Being marketed at very low prices compared to other techniques is another advantage of this type of blood glucose monitor. However, it can only provide the discrete blood glucose values at times of testing, without information of trends in glucose levels which can help warn coming unwanted episodes of hypoglycemia. Moreover, piercing the finger from time to time during day and night brings a lot of inconvenience for patients and their families.

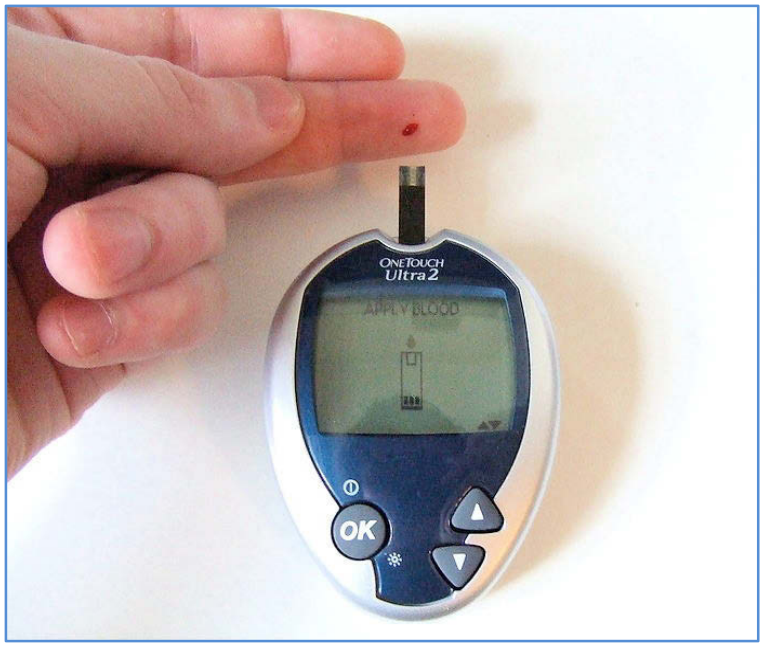


Figure 2.2: An intermittent blood glucose monitoring device

Continuous blood glucose monitors, on the other hand, provide better information about shifting blood glucose levels throughout the day. Because of its continuous characteristic, the technology applied to this kind of monitor is totally different to intermittent monitors. Instead of taking blood, to determine the glucose concentration in blood, continuous monitors require sensors which measure blood glucose with minimal invasiveness through continuous measurement of interstitial fluid or with the non-invasive method of applying electromagnetic radiation through the skin to blood vessels in the body. Table 2.3 provides a comparison of three continuous monitors that have been approved by the U.S. Food and Drug Administration (FDA). Compared to the type of intermittent monitoring, continuous monitors are marked at much higher prices. Low accuracy is one of the most prominent disadvantages of continuous monitors. Moreover, the requirement of frequently inserting sensors and continuous calibration also cause a lot of trouble for users, especially for young patients.

Table 2.3: Continuous glucose monitoring systems (CGMS)

	GUARDIAN REAL-TIME CGMS ®	Dexcom SEVEN Plus	Abbott FreeStyle Navigator®
FDA approval	February 2007	March, 2006	March 2008
			
Price	• $1339 for monitor, transmitter, charger, and 4 sensors • $35 per sensor	• $1248 for receiver, case, charger, transmitter • $399 per 1 month 4 sensors	No longer marketed since April 2010
Sensor life	3 days	7 days	5 days
Length of sensor probe	13 mm	13 mm	6 mm
Calibration	First calibration is 2 hours after insertion. Second calibration within next 6 hours after first, then every 12 hours. Will alarm if calibration value not entered.	Every 12 hours with another device named One Touch Ultra	Calibrate at 10, 12, 24 and 72 hours after insertion with no further calibration for the final 2 days of the 5 day wear

2.4.2 Physiological glucose monitoring

Physiological glucose monitoring is another technique of determining the blood glucose concentration based on physiological parameters of the patient's body. By this way, the monitoring process is uninvolved with the patient's blood flow and hypoglycemic episodes can be detected non-invasively. This technique is desirable for patients as well as their caregivers because even minimally invasive devices bring the inconvenience of sensor inserting and changing procedures, as well as the risk of infection. On the other hand, the main limitation of this technique is a high false positive rate (alarm for reasons other than hypoglycemia). This is because physiological parameters are sensitive to not only hypoglycemia but also a huge range of reactions of the body.

Figure 2.3: The Diabetes Sentry Monitor

Designed over 20 years ago by Teledyne Avionics, The Diabetes Sentry® is a monitor worn on the wrist which monitors for two symptoms of hypoglycemia: perspiration and a drop in skin temperature (Figure 2.3). The presence of either sweating or a two degree Fahrenheit drop in body temperature triggers an audible alarm that will awaken most people. As other physiological glucose monitors, this device has the advantages of providing protection to patients without any procedure of inserting and continuously

changing sensors. However, it was reported that the accuracy is considerably low with true positive rate of 50% at quite low plasma glucose level of 37 ± 11 mg/dl (Nguyen et al. 2013b). Moreover, this device is only recommended for patients without hypoglycemia unawareness who exhibit the symptoms of perspiration and/or a drop in skin temperature under hypoglycemic conditions. The producer also comments that there are false positives which occur whenever perspiration is present or a drop in temperature occurs due to various reasons such as when patients are dreaming, sick, on medication or the room is too hot.

Figure 2.4: The Glucowatch G2 Biographer

The Glucowatch G2 Biographer (from Cygnus Inc., California, USA) is another non-invasive wristwatch-like device that was approved by the U.S. Food and Drug Administration (FDA) in 2001 for detecting trends and tracking patterns in glucose levels in adults (age 18 and older) and children/adolescents (age 7 to 17) with diabetes. This device provides real-time measurements of interstitial glucose concentrations at 10-min intervals by conducting a constant low-level electric current through the skin between two electrodes (as shown in Figure 2.4). It was reported by clinical studies (Group 2004) that the device performs better at higher glucose levels and is not reliable with low accuracy in detecting hypoglycemia (23% sensitivity and 51% specificity when setting the alarm level at 60 mg/dl). With this level of accuracy, coming along with various limitations such as the requirement of calibration using another blood glucose meter, replacement of sensors every

23 hours, etc., this device is not recommended to be used as an independent tool for detecting hypoglycemia non-invasively.

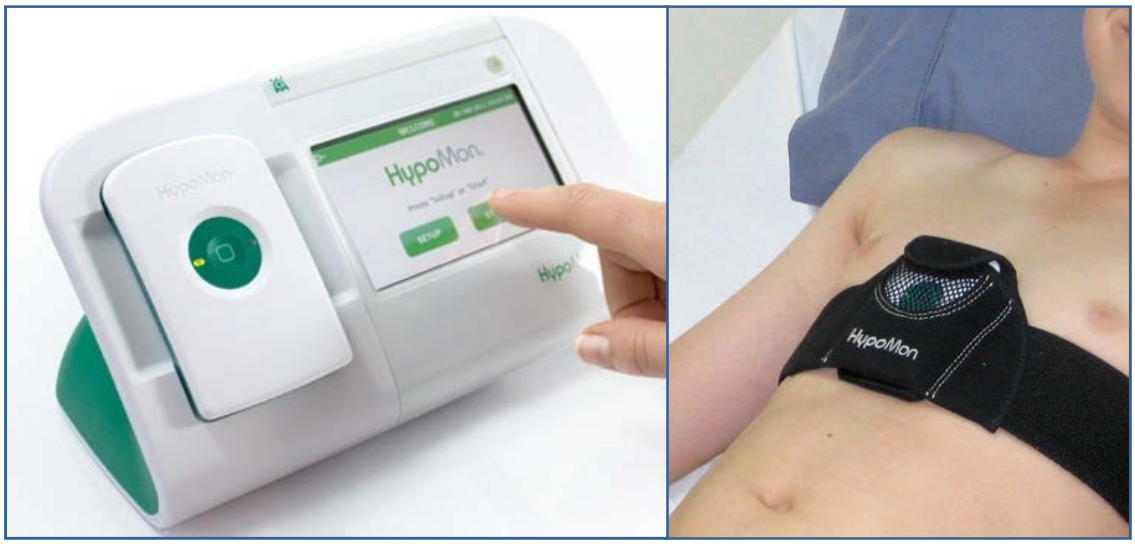


Figure 2.5: The HypoMon Monitor

Designed by Australian researchers at the University of Technology Sydney and later developed by the company AIMEDICS, HypoMon is a non-invasive alarm system that identifies night-time hypoglycemia in children and young adults aged between 10 and 25 years with type 1 diabetes. This device is intended for use during sleep-time at home, enabling users to increase their ability to detect nocturnal hypoglycemic episodes. The HypoMon includes a Monitor and a Belt (as shown in Figure 2.5). The HypoMon Belt, worn comfortably around the upper chest, contains a matchbox-sized, needle-free transmitter. The sensors communicate wirelessly with the HypoMon Monitor, logging the body's physiological signals, including heart rate and cardiac outputs during sleep. Using an advanced computational algorithm to identify the onset of hypoglycemic episodes, this device has been reported with desired accuracy and has been approved by The Therapeutics Goods Administration (TGA) as an Australian developed, potentially life-saving, non-invasive alarm that can protect young people with type 1 diabetes against nocturnal hypoglycemia (Nguyen, Ghevondian & Jones 2009; Skladnev et al. 2010).

2.5 Hypoglycemia detection from EEG signals

The human brain constitutes only about 2% of body weight, but consumes up to 20% of the energy used by the body, more than any other organ. Glucose is an obligate fuel for the metabolism of the brain. Because it cannot synthesize as well as store this primary metabolic fuel, the brain depends on a continuous supply of glucose from blood circulation and is vulnerable to any glucose deprivation.

Although hypoglycemia can produce a large number of symptoms, like sweating or increased cardiac output, the principal problems arise from an inadequate supply of glucose, which is the primary metabolic substrate, to the brain. Since the electroencephalography (EEG) signal is directly related to the metabolism of brain cells, it has been believed that hypoglycemia can cause early changes in EEG that can be non-invasively detected. Previous studies have attempted to find out EEG changes caused by hypoglycemia.

2.5.1 Electroencephalogram and its applications in biomedical systems

By definition, electroencephalogram (EEG) is the recording of electrical activities along the scalp of human brain. The brain's electrical charge is maintained by billions of neurons. Neurons are electrically charged (or "polarized") by transport proteins that pump ions across their membranes. When a neuron receives a signal from its neighbour via an action potential, it responds by releasing ions into the space outside the cell. Ions of like charge repel each other, and when many ions are pushed out of many neurons at the same time, they can push their neighbours, who push their neighbours, and so on, in a wave. When the wave of ions reaches the electrodes on the scalp, they can push or pull electrons on the metal on the electrodes. Since metal conducts the push and pull of electrons easily, the difference in push, or voltage, between any two electrodes can be measured by a voltmeter. Recording these voltages over time gives us the EEG signal.

The electric potentials generated by single neurons are far too small to be acquired by EEG electrodes. EEG activity therefore always reflects the summation of the synchronous

activity of thousands or millions of neurons that have similar spatial orientation. If the cells do not have similar spatial orientation, their ions do not line up and create waves to be detected and recorded.

In conventional applications, the EEG recording is obtained by placing electrodes on the scalp with a conductive gel or paste to reduce the skin impedance. When measuring from the scalp, typical adult human EEG signals are about 10μV to 100 μV in amplitude. As a result, the signals after being recorded need to be amplified and then digitized to display and store in computers.

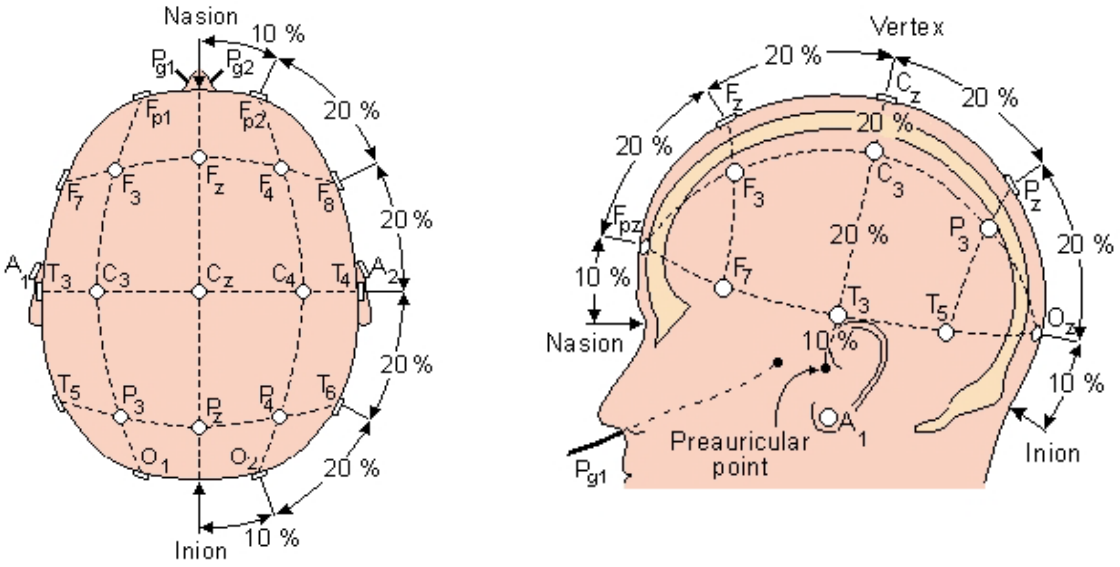


Figure 2.6: The international 10-20 system seen from (A) above and (B) left the head [Adopted from (BCI2000)]

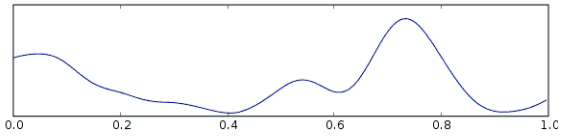
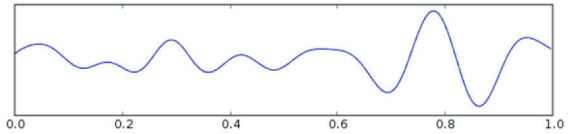
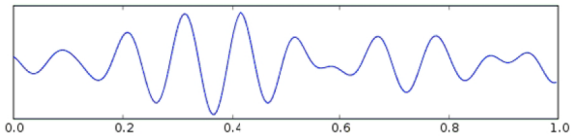
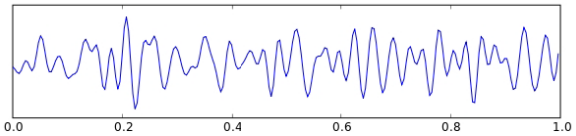
In order to achieve uniformity in the interpretation of the EEG activity recorded globally, the spatial locations and names for electrode placement are standardized according the International 10–20 system as in Figure 2.6 (Towle et al. 1993). The numbers "10" and "20" refer to the fact that the actual distances between adjacent electrodes are either 10% or 20% of the total front-back or right-left distance of the skull. Each site of the brain has a letter to identify the lobe and a number to identify the hemisphere location. The letters F, T, C, P and O stand for Frontal, Temporal, Central, Parietal, and Occipital, respectively. Even numbers (2, 4, 6 and 8) refer to electrode positions on the right hemisphere, whereas odd numbers (1, 3, 5 and 7) refer to electrode positions on the left hemisphere of the brain.

The EEG signal is popularly described in terms of rhythmic activities or brain wave patterns. By means of Fourier Transform, the raw EEG signals in time domain can be transformed into frequency domain to derive brain waves. These waves are often divided into four basic frequency bands: delta band (0.5-4Hz), theta band (4-8 Hz), alpha band (8-13Hz) and beta band (13-30 Hz). These divisions which are based on the characteristics of each band are somewhat different between research groups. A summary of features of each frequency band is presented in Table 2.4

The EEG has been widely used in a variety of health and medical applications. One of the best known applications is the detection of epileptic seizures and localization of the seizure origin (Smith 2005). EEG helps determine seizure type and epilepsy syndrome in patients with epilepsy, and thereby choice of antiepileptic medication and prediction of prognosis. EEG findings contribute to the multi-axial diagnosis of epilepsy, in terms of whether the seizure disorder is focal or generalised, idiopathic or symptomatic, or part of a specific epilepsy syndrome. Another well-known EEG-based application is to monitor drowsiness and detect the onset of fatigue in drivers (King, Nguyen & Lal 2006; Subasi & Kiymik 2010). Most recently, researchers have shown the potentiality of EEG signals in detecting the freezing of gait in patients with Parkinson's disease (Cole, Roy & Nawab 2011; Handojoseno et al. 2012). The EEG is also recognized as a prominent non-invasive means applied in the area of brain-computer interface. The EEG based brain-computer interface has been pursued extensively by a number of research labs to develop assistive devices designed for use by disabled people, such as intelligent wheelchairs (Craig & Nguyen 2007; Luzheng, Xin-An & Yili).

Table 2.4: Basic EEG rhythms

(Rowan & Tolunsky 2003; Sadasivan & Narayana Dutt 1994)

Wave patterns	Example of 1-second activity	Features
Delta (0.5-4Hz)		- the highest in amplitude (250 – 300 μV) and the slowest waves - prone to eye-movement artifact contamination - prominent in infants and young children - present in adult sleep, not present in normal adult waking state
Theta (4-8 Hz)		- normal amplitude in the range of 100 – 150 μV - prominent in young children - present in drowsiness in older children and adults
Alpha (8-13 Hz)		- normal amplitude of 40-50 μV - maximal amplitude in the occipital regions - most prominent when measured in the relaxed, waking state with eyes closed - increases during drowsiness even when eyes opened
Beta (13-30 Hz)		- relatively small amplitude compared to other bands - usually present on both brain sides in symmetrical distribution and most evident frontally - low amplitude with multiple and varying frequencies is often associated with active, busy or anxious thinking and active concentration

2.5.2 Correlation of EEG signals and hypoglycemia

The correlation of EEG signals and hypoglycemia has been reported by many research groups round the world. In general, these studies can be categorized into two different types. The first type focuses on exploring changes in EEG signals under the occurrence of hypoglycemic episodes which can be natural or insulin-induced episodes. In terms of hypoglycemia detection, an extensive literature review on the first type helps to identify important EEG changes under hypoglycemic conditions which can be used as input parameters for the detecting algorithm. Meanwhile, the second type aims to analyse the correlation of abnormalities in EEG signals which are induced by the frequent exposure to hypoglycemia in T1DM patients.

In this section, various studies exploring the association between EEG signals and hypoglycemia will be reviewed. Besides, results of some recent studies carried out to develop the alarm system to hypoglycemia using EEG signals will also be mentioned as references for this thesis.

2.5.2.1 Studies were carried out to discover the EEG changes during hypoglycemia

One of the earliest studies was conducted in 1988 by Pramming et al. to examine the effect of induced hypoglycemia on the electroencephalogram of 13 patients with type 1 diabetes (Pramming et al. 1988). No changes were seen in the EEG when the BGL was above 3 mmo/l (Figure 2.7). The spectral analysis of patients' EEG signals revealed that at BGL of 2 mmol/l, there is an abrupt drop in alpha activity, accompanied by an increase in theta activity. The changes were found in all of the patients with the exception of one, and in all of the EEG channels but most prominently in the frontotemporal region.

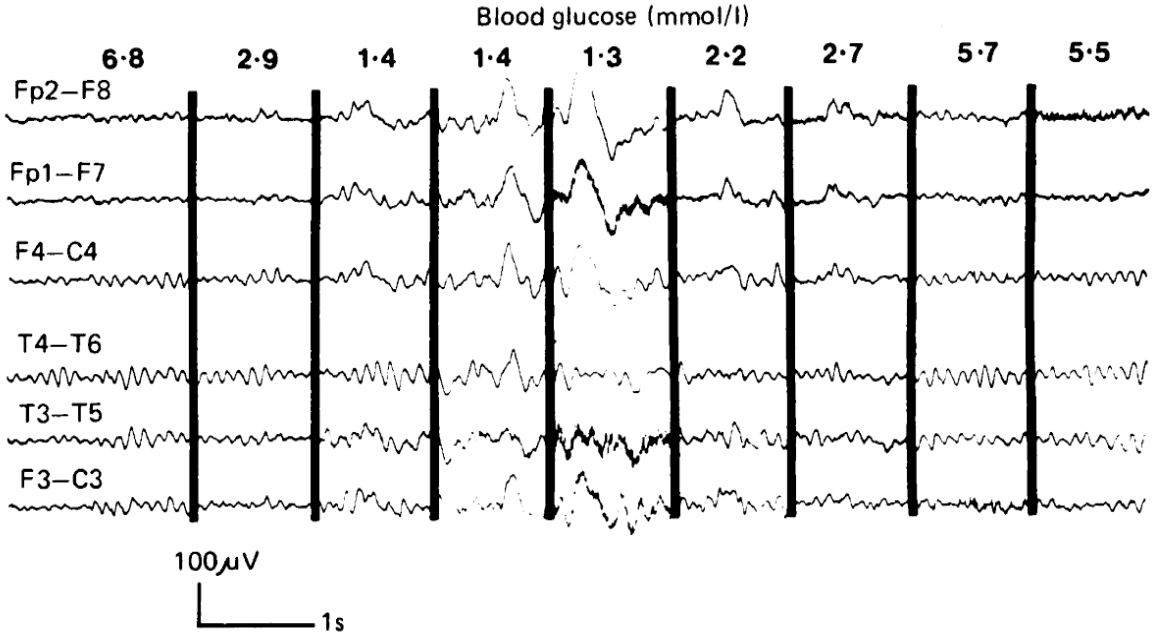


Figure 2.7: EEG signals at different blood glucose concentrations from one patient [Adopted from (Pramming et al. 1988)].

In 1990, Tallroth et al. implemented a study to measure and evaluate the influence of hypoglycemia on the cerebral function before, during and right after insulin-induced hypoglycemia by exploring quantified EEG, P300 and somatosensory evoked (Tallroth et al. 1990). The EEG-related results of this study showed that under hypoglycemic conditions, there are widespread increases of delta and theta activity (0-8 Hz) and an increase of low frequency alpha activity (8.12-9.62 Hz) anteriorly over the brain, along with a decrease of high frequency alpha activity (9.75-12.87 Hz) and beta activity (13.0-19.37 Hz) over the posterior regions. Comparing between a diabetic group of 8 T1DM patients and a control group of 12 age-matched healthy subjects, it was reported that the increase of theta activity and the anterior release of alpha activity during hypoglycemia was more marked in the diabetic group. This study concluded that most of the significant EEG changes in the present study were found within the anterior areas.

In 1991, Bendtson et al. reported a nocturnal hypoglycemia – associated study with 8 T1DM patients which was implemented in two consecutive and one subsequent night with

continuous monitoring of EEG (Bendtson et al. 1991). This study aimed to evaluate the influence of both spontaneous and insulin-induced hypoglycemia on nocturnal electroencephalogram sleep-patterns. This results of this study showed an increase in delta and theta activity in EEG signals of only 3 patients at blood glucose levels below 2.0 mml/l. The EEG changes were found equally in all regions of the brain. These changes were not identical in each patient, however they were reproducible.

Howorka et al. in 1996 reported an important study, investigating the relationship between EEG parameters of vigilance in response to insulin-induced hypoglycemia in T1DM patients with different levels of hypoglycemic unawareness (Howorka et al. 1996). From the point of hypoglycemia detection, this study led to crucial findings which are an early and immediate reduction in vigilance as well as changes in vigilance-associated EEG parameters under hypoglycemic conditions. The study established that if these results are confirmed in further investigations, EEG characteristics of vigilance might be important tools for predicting the onset of hypoglycemia in diabetic patients.

Another study was conducted by Tribl et al. (Tribl et al. 1996) with the aim of determining the EEG power spectra in diabetic patients with insulin dependent diabetes mellitus at different levels of hypoglycemia, and documenting the topographical distribution of the EEG changes (as shown in Figure 2.8). The results confirmed earlier studies on insulin-induced hypoglycemia concerning the deceleration of electrical activity with an increase in delta and theta activity, and a decrease in alpha activity. The changes were reported to be most pronounced in the theta band and the most sensitive parameter was the alpha/theta ratio. At the glucose level of 50-60 mg/dl, the increase in delta and theta activity was most pronounced in later frontal regions. At the lower glucose level, the increase in slow frequencies appeared in the posterior part of the brain.

In 1998, Bjørgaas et al. reported a study exploring quantitative spectral analysis of EEG signals of children (19 diabetic and 17 non-diabetic) during a gradual decline in plasma glucose. This study aims to elucidate how EEG signals respond to hypoglycemic symptoms; to determine the plasma glucose level at which EEG changes first appear; and to investigate whether the anterior regions of the brain are more sensitive to hypoglycemia

than other areas (Bjørgaas et al. 1998). The results showed that with children, EEG changes occur already at around 4 mmol/l. At around 3 mmol/l, the glucose level often experienced by diabetic children, the EEG deterioration is substantial, and appears over the entire cerebral cortex (as shown in Figure 2.9).

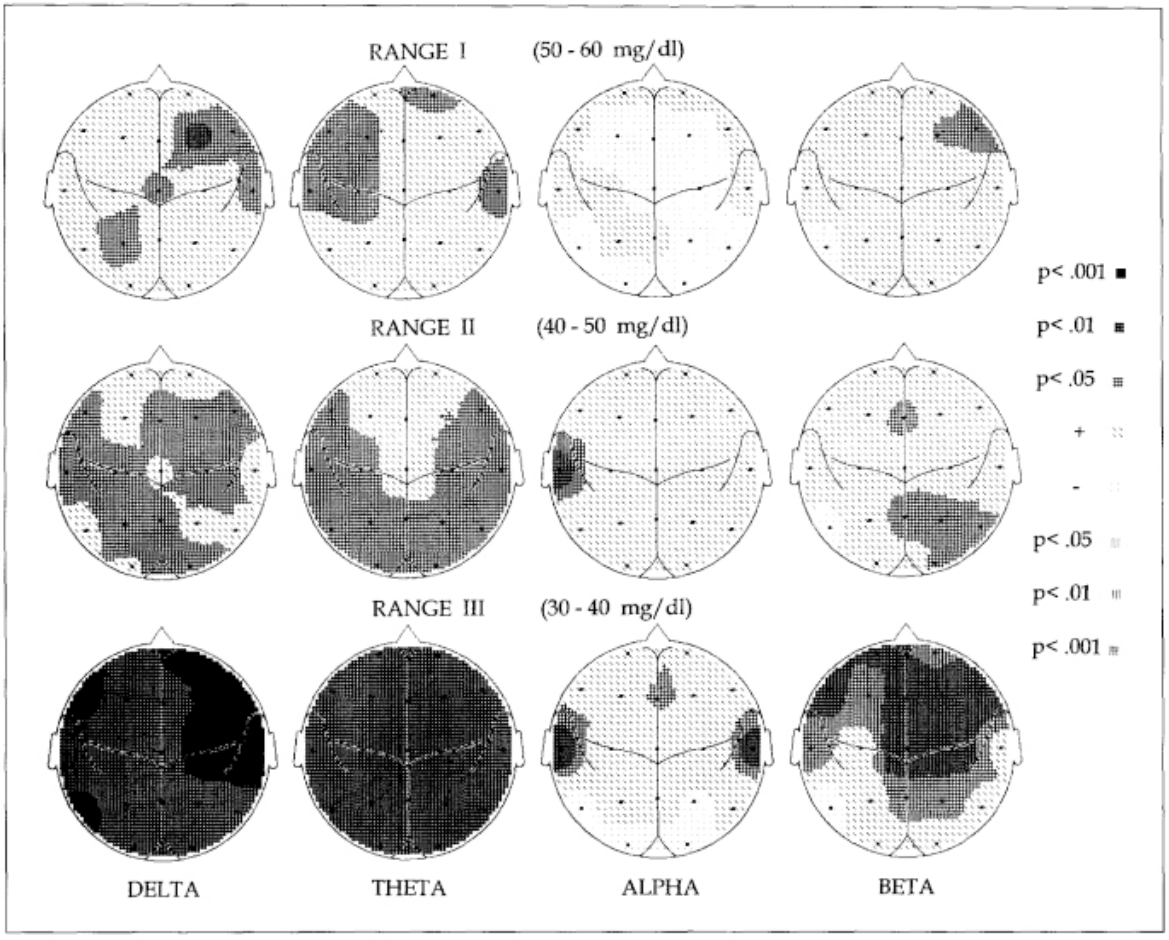


Figure 2.8: Differences between hypoglycemia and nomoglycemia in type 1 diabetic patients for absolute EEG power depicted in significance probability maps [Adopted from (Tribl et al. 1996)].

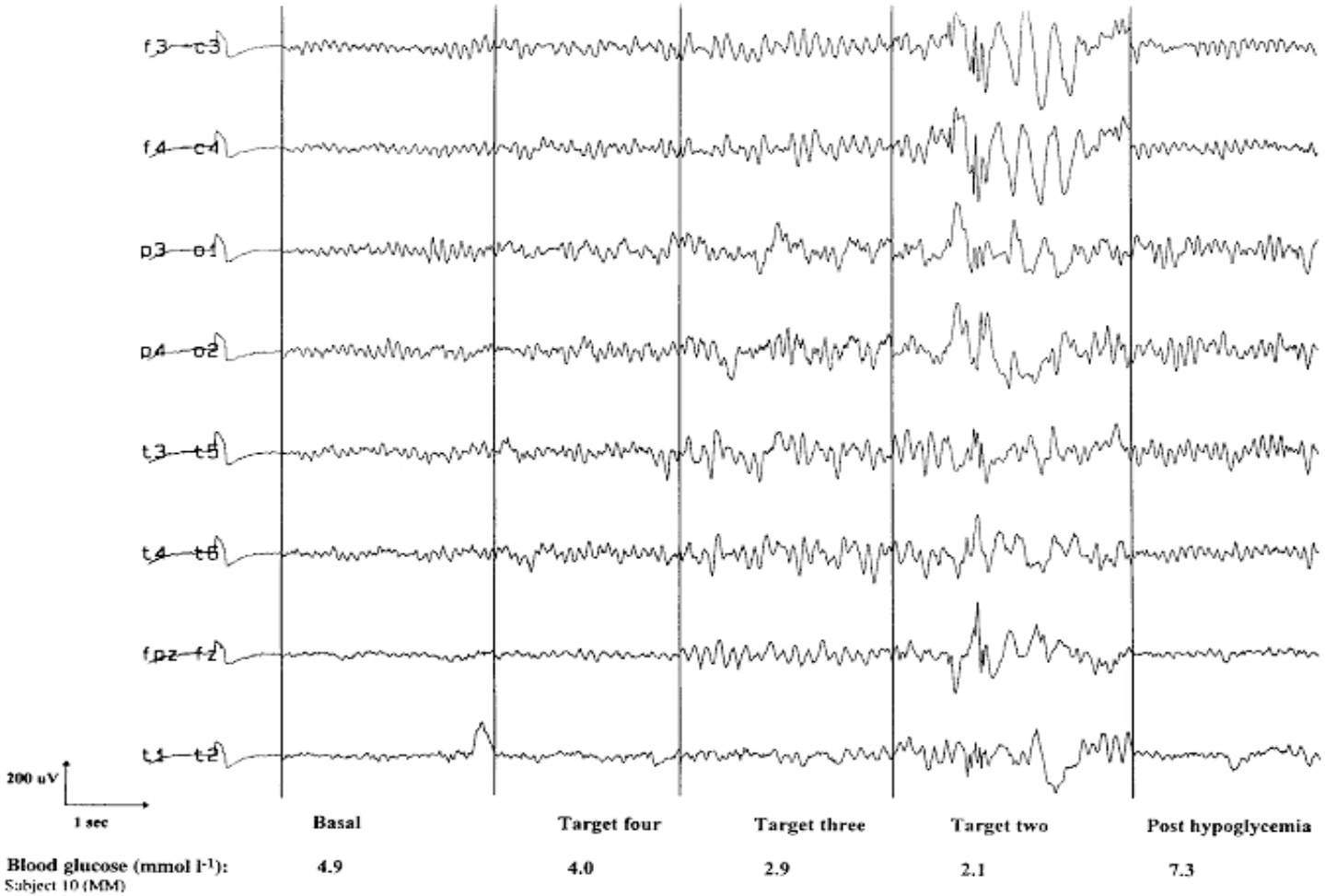


Figure 2.9: Electroencephalograms from one 13-year-old diabetic patient, showing increasing amounts of slow activity at plasma glucose 4.0, 2.9, and 2.1 mmol/l. [Adopted from (Bjørgaas et al. 1998)].

2.5.2.2 Studies were carried out to discover the EEG abnormalities during non-hypoglycemia state in subjects with type 1 diabetes

In 1996, Bjørgaas et al. (Bjørgaas, Sand & Gimse 1996) conducted a study to investigate if EEG signals are affected in diabetic children with and without episodes of severe hypoglycemia, as well as to test the assumption that the frontal lobe is more sensitive to hypoglycemia than other cortical areas. In this study, there were 28 children with T1DM (15 had experienced episodes of severe hypoglycemia – SH patients and 13 had not experienced such episodes – non SH patients) and 28 age and sex-matched control children. The study results showed an increased theta activity in the SH group bilaterally in the frontocentral region and a slight trend toward more delta activity in diabetic children than control children bilaterally in the occipital electrodes. The relative alpha amplitude was decreased in the SH group at several locations. The study supported the hypothesis that the frontal lobe is especially vulnerable to hypoglycemia.

Another study conducted in 2002 by Brismar et.al aimed to investigate signs of brain dysfunction in T1DM patients by exploring quantitative EEG parameters in 49 T1DM adults (with good glycemic control and without history of recurrent hypoglycemia) and 51 control subjects. The study established that its most pronounced finding was a loss of fast oscillations (alpha, beta, gamma) in both posterior temporal regions in diabetic patients (most significant for beta activity $p<0.001$). There also was a decrease in beta activity bilaterally in the anterior temporal and occipital regions as well as a slight slowness of theta alpha peak frequency.

Hyllienmark et al. in 2005 reported a study to identify whether adolescents with T1DM have EEG abnormalities (Hyllienmark et al. 2005). The study population included 35 T1DM patients with disease duration of 7.6±4.6 years and 45 healthy control subjects. Compared with control subjects, the EEG signals from diabetic patients showed an increase in slow activities (delta and theta) and a reduction in alpha peak frequency, both of which was most pronounced in the frontal regions. The results of this study also showed a decrease in fast activity (alpha, beta and gamma), which was most pronounced bilaterally in the posterior temporal regions.

2.5.2.3 Recent studies carried out to develop the alarm system to hypoglycemia using EEG signals

An extensive literature review on the correlation between EEG signals and hypoglycemia has shown a potential possibility of developing a hypoglycemia-monitoring system for T1DM patients from their non-invasive EEG signals. This section provides a brief review of several recent studies that have been carried out with the aim of developing real-time systems that can detect hypoglycemia using EEG signals.

In 2005, Laione and Marques proposed a methodology of hypoglycemia detection based on EEG signals (Laione & Marques 2005). This study developed a system to acquire, process and analyse EEG signals from subjects and then used EEG parameters to classify the state of patients into hypoglycemia or non-hypoglycemia. The classifying technique used in this study was neural network. This study led to the result of 49.2% accuracy rate, 76% sensitivity and 32.5% specificity when the neural network was trained and validated with different subject groups. These results showed that hypoglycemia can be detected using EEG signals. However, the proposed methodology needs to be developed further to enhance the performance. To do this, this study proposed that this system would need to be calibrated by acquiring trial EEG signals during some spontaneous episodes of hypoglycemia together with the corresponding blood glucose levels. Obviously, this is not a good solution because it is not an easy procedure for patients, especially since the data needs to be under real hypoglycemic conditions. With the reported results and limitations, this system apparently needs to be improved to get a more advanced method. However, since being published in 2005, this study has not been further developed and elaborated to apply to the clinical environment.

Most recently in 2010, Juhl et al. reported a study to test the hypothesis that specific changes in the electroencephalogram during hypoglycemia can be recorded by subcutaneous electrodes and processed by a general mathematical algorithm (Juhl et al. 2010). Although currently, this study has been continuously developed to produce a real-time system that can detect hypoglycemia using EEG signals from the brain, it encounters a great disadvantage of using implanted electrodes to record EEG signals. It is obvious that

this technique, with a troublesome surgical procedure of implanting and regular replacing electrodes, along with a high risk of infection, could not be a good solution for patients, especially for those who are young.

2.5.2 Computational Intelligence for EEG-based hypoglycemia detection

In recent years, with the rapid growth of computer technologies, Computational Intelligence (CI) has played an important role in the development of fundamental research as well as real-life applications. Basically, CI is a set of nature-inspired computational methodologies and approaches which aim to deal with complex problems to which traditional methods are ineffective or unfeasible. Primarily it covers the subjects of artificial neural networks, fuzzy logic, evolutionary computation and the combination of them in various real-world applications. In the field of biomedical engineering and technology, CI has made promising and essential contributions in developing biomedical systems that aim to detect human health problems as well as assist people with disabilities.

Since the first form was introduced in 1943 (McCulloch & Pitts 1943), neural networks (NNs) have grown progressively to be employed popularly in the biomedical area as a powerful tool of classification and pattern recognition. Different neural network – based algorithms have been developed to be used as a classification platform in various biomedical applications:

- Monitoring drowsiness and detecting the onset of fatigue in drivers (King, Nguyen & Lal 2006; Subasi & Kiymik 2010).
- Finding free spaces, performing obstacle avoidance and controlling travel directions of power wheelchair to assist patients with disabilities (Nguyen et al. 2012; Nguyen 2008).
- Detecting the freezing of gait in patients with Parkinson's disease (Cole, Roy & Nawab 2011; Handojoseno et al. 2012).
- Early diagnosis of skin cancer (Ercal et al. 1994; Sigurdsson et al. 2004); breast cancer (Karahaliou et al. 2008; Woten, Lusth & El-Shenawee 2007), etc.

Recently, studies implemented at the Centre for Health Technologies, University of Technology Sydney, Australia have shown the effectiveness of employing various CI algorithms to detect the onset of nocturnal hypoglycemia for patients with type 1 diabetes mellitus (T1DM) using physiological parameters including skin impedance, heart rate and a variety of cardiac outputs from electrocardiographic (ECG) signals. In 1999, Ghevondian and Nguyen first introduced the use of a neural network algorithm in modelling of patients' blood glucose profiles (Nguyen, Su & Nguyen 2011). This study has been leading to the successful production of the HypoMon® System which is a non-invasive alarm for night time hypoglycemic episodes. This HypoMon® System has been reported to employ an effective Bayesian neural network as the classification unit to detect the onset of hypoglycemic episodes in T1DM children (Nguyen, Ghevondian & Jones 2009). Most recently, the combinations of neural network with evolutionary algorithms were also reported to be effective tools in detecting hypoglycemic episodes from physiological parameters (Nguyen et al. 2012; Nguyen et al. 2013a).

Using electroencephalography (EEG) signals to detect hypoglycemic episodes is a new approach which has been shown to be a feasible and effective solution for patients with T1DM. To do this, the methodology consists of two main computational tasks which are EEG feature extraction and classification. Using extracted parameters from EEG signals as inputs, the classification algorithm determine that the patient's current state is hypoglycemia or non-hypoglycemia. Recent studies have shown that neural network, as a powerful classification unit, can potentially identify hypoglycemic episodes in T1DM patients (Juhl et al. 2010; Laione & Marques 2005). With reported classification results and limitations, these two studies need to be improved further to achieve better performance in order to be applied into the real clinical environment.

2.6 Discussion

Type 1 diabetes mellitus (T1DM) is a chronic illness characterized by the body's inability to produce insulin due to the autoimmune destruction of the beta cells in the pancreas. This type of diabetes is most common in juveniles, but it can also develop in adults in their late 30s and early 40s. T1DM patients depend indefinitely on the external insulin treatment to maintain their targeted blood glucose levels throughout their whole lives. However, this treatment brings extreme burdens for most patients and their families, due to the fear of the most common but dangerous complication which is hypoglycemia (or low blood glucose level). Hypoglycemia has been recognized to impact all patients with T1DM, impairing their quality of life and limiting attempts to achieve desired targets for glycemic control. Episodes of nocturnal hypoglycemia, in which symptoms can be obscured by sleep, may lead to significant anxiety and morbidity, especially for T1DM children and their families.

Due to the prevalence and severity of hypoglycemia in patients with T1DM, a variety of studies have been carried out, using different methods to produce systems that can monitor patients' blood glucose profiles, detect the onset of hypoglycemic episodes and give alarm to patients as well as caregivers. Some of them require gradually taking patients' blood samples to determine the blood glucose level (BGL). This method gives relatively exact information about hypoglycemic status. However, taking blood is uncomfortable for patients, and very inconvenient to monitor continuously, especially during the night. A non-invasive technique is obviously a better solution for these disadvantages. Currently, on the market, there are some devices which monitor hypoglycemia non-invasively using physiological parameters such as heart rate, skin impedance and electrocardiogram (ECG) parameters.

Under the occurrence of hypoglycemia, the human brain is one of the first affected organs. Because it cannot synthesize as well as store this primary metabolic fuel, the brain depends on a continuous supply of glucose and is vulnerable to any glucose deprivation (Cryer, Davis & Shamoon 2003). Since the electroencephalogram (EEG) is directly related to the metabolism of brain cells, a failure of cerebral glucose supply can cause early changes in EEG signals. A number of studies have reported important traces in EEG signals induced

by hypoglycemic episodes in T1DM patients (Bjørgaas, Sand & Gimse 1996; Howorka et al. 1996; Pramming et al. 1988). Recent studies also lead to acceptable results which show the potential ability of detecting hypoglycemia from EEG signals (Juhl et al. 2010; Laione & Marques 2005). Nevertheless, to our knowledge until now, none of these studies has been elaborated towards the aim of producing a monitoring and alarm system which can detect hypoglycemia from non-invasive EEG signals.

This research is the first stage of the project currently implemented at the Centre for Health Technologies, University of Technology Sydney which aims to develop a system that can detect the onset of hypoglycemic episodes from EEG signals, and then give an alarm to provide enough time for patients and caregivers to take action. The main objective of this thesis is to design a computational methodology to detect nocturnal hypoglycemic episodes for T1DM patients from their EEG signals. There are two core parts contained in the developed computation methodology, including the signal analysing and feature extraction part, and the classification part. The developed methodology in this thesis is based on two important criteria for developing real-life applications which are high performance and saving computational cost.

Chapter 3

Identification of Hypoglycemic States for Patients with T1DM using EEG Signals

---

3.1 Introduction

Hypoglycemia, or abnormally low blood glucose level (BGL), is the most common but dangerous complication of the intensive insulin therapy for patients with type 1 diabetes mellitus (T1DM). Recurrent exposure to hypoglycemic episodes impacts the life quality of all T1DM patients, limits their intellectual as well as physical activities, and potentially causes irreversible and severe effects, such as cognitive impairments, seizures, coma, and even death.

Under the occurrence of hypoglycemia, the human brain is one of the first affected organs. Because it cannot synthesize as well as store this primary metabolic fuel, the brain depends on a continuous supply of glucose and is vulnerable to any glucose deprivation (Cryer, Davis & Shamoon 2003). Since the electroencephalogram (EEG) is directly related to the metabolism of brain cells, a failure of cerebral glucose supply can cause early changes in EEG signals which can be non-invasively detected in order to identify hypoglycemic

episodes in T1DM patients. A number of studies have reported important traces in EEG signals induced by hypoglycemic episodes in T1DM patients (Howorka et al. 1996; Hyllienmark et al. 2005). Recent studies also lead to acceptable results which show the potential ability of detecting hypoglycemia from EEG signals (Juhl et al. 2010; Laione & Marques 2005). Nevertheless, all of these results need to be improved further in order to be applied into the real clinical environment.

The core objective of this chapter is to build up a computational framework for detecting hypoglycemia from non-invasive EEG signals for patients with T1DM. The developed framework consists of two main parts which are EEG feature extraction and classification. The raw data collected from five T1DM patients will be processed to get rid of unwanted noises before being analysed to extract important EEG parameters or features. The extracted EEG features will be used as inputs of a classification unit in order to detect the onset of hypoglycemic episodes.

Specifically, the feature extraction part aims to explore physiological responses of EEG signals during an insulin-induced nocturnal hypoglycemic study of five T1DM patients. The association between different EEG spectral parameters and the transition of patients' states during the study will be analysed. More explicitly, the correlation between four spectral EEG parameters from four different EEG channels and patients' conditions during the whole study including three different phases of Normal, Hypoglycemia and Recovery will be investigated. The data within the BGL range of 3.3-3.9 mmol/l (60-70 mg/dl whole blood glucose) will also be analysed to figure out the EEG responses to the potentially early onset of hypoglycemia. Based on those analyses, EEG parameters which are most sensitive to hypoglycemia will be established.

Using extracted EEG features as inputs, a classification algorithm for detecting the occurrence of hypoglycemic episodes will be developed. In this chapter, a standard neural network is created for the classification purpose. The Levenberg-Marquardt algorithm is applied to train the neural network. Classification performance will be determined in order to evaluate the potentiality of the detection of hypoglycemic episodes from EEG signals for patients with T1DM.

3.2 Methodology

3.2.1 Study protocols

All analyses of this thesis are based on the data from an overnight glucose clamp study which was implemented at the Princess Margaret Hospital for Children in Perth, Australia. There were five T1DM adolescents (between the ages of 12 and 18 years old) who participated in the study. The HbA1c levels of these patients were within the range of 6.5% and 8.9%. All experienced occasional mild hypoglycemia, as is usual during the course of treatment in adolescents with T1DM.

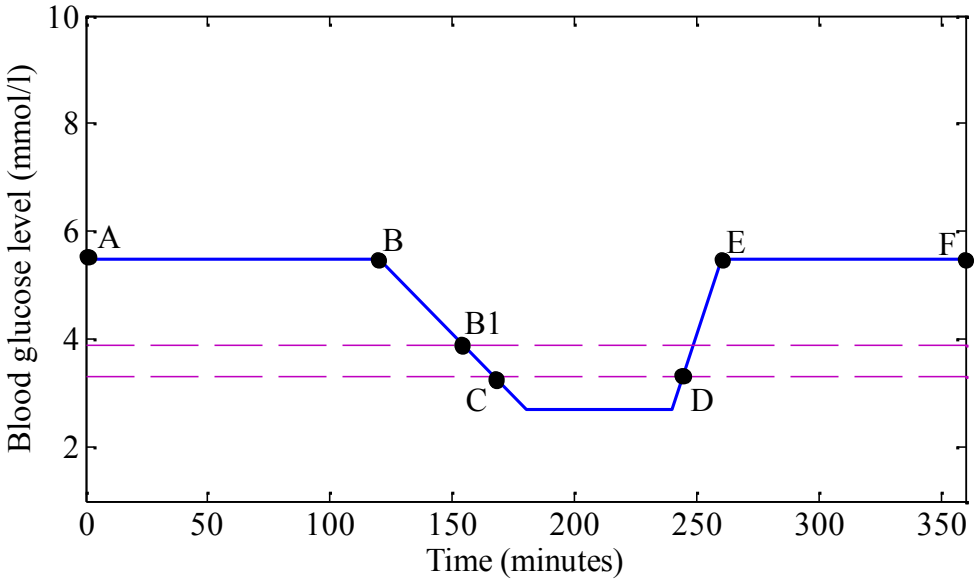


Figure 3.1: Target BGL profile of the induced hypoglycemia study

The target BGL profile of the study is plotted in Figure 3.1. Each patient underwent an overnight study which consisted of five phases approximately: baseline (around 30 minutes before insulin infusion, which is used for reference only), euglycemia (AB), ramp (BC), hypoglycemia (CD), and recovery (DEF). In this study, hypoglycemia is defined as blood glucose levels lower than 3.3mmol/l (equivalent to 60 mg/dl whole blood glucose).

During the study, EEG signals were continuously recorded and stored by using a Compumedics Siesta System with the sampling rate of 128 Hz. EEG electrodes were positioned at 4 channels O1, O2, C3 and C4 according to the International 10/20 system, referenced to A1 and A2, respectively (as shown in Figure 3.2). There were also 2 electrodes placed at patients' chins to acquire the electro-myogram (EMG) signals and 2 electrodes placed near patients' eyes to measure the electro-oculogram (EOG) signals. The actual BGLs were routinely collected to be used as reference using Yellow Spring Instruments with the general sampling period of 5 minutes.

The patients' sleep state was monitored during the study. It was determined that with dim lighting and experienced nursing and medical staff, arousal or activation was not a complicating factor. Also, it was concluded that the sleep quality of all patients was not impaired by the performance of these clamp study protocols.

All data used in this thesis were collected with the approval of the Women's and Children's Health Service, Department of Health, Government of Western Australia, and with informed consent.

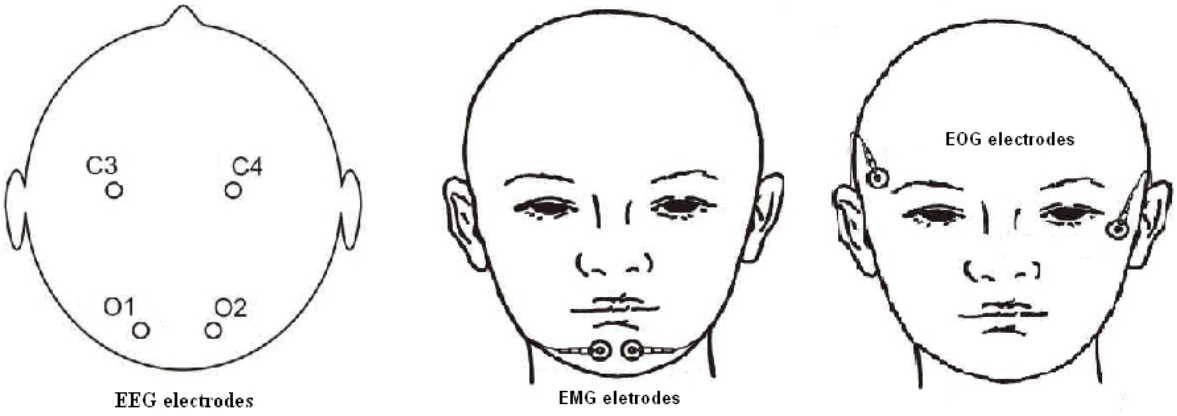


Figure 3.2: Electrode positions in the study

3.2.2 EEG feature extraction for identifying hypoglycemic states in T1DM patients

After collecting data from the clinical hypoglycemia-associated study, raw EEG signals from patients need to be processed to extract important parameters that can be used to identify the state of hypoglycemia. Thanks to this step of processing, extracting and analysing EEG parameters, significant features can be acquired to be used as inputs of the classification algorithm for detecting hypoglycemia. A diagram of implementing feature extraction in this thesis is provided in Figure 3.3. Details about each task are presented in sections 3.2.2.1- 3.2.2.3.

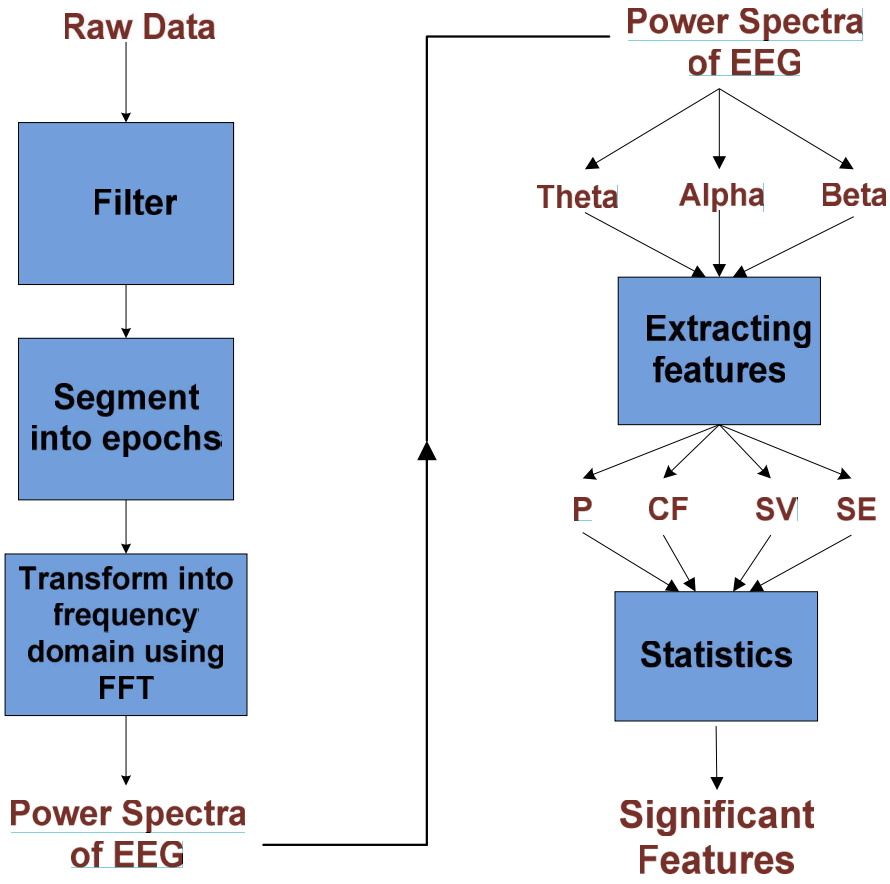


Figure 3.3: EEG feature extraction procedure

(P : Power ; CF : Centroid frequency ; SV : Spectral Variance ; SE : Spectral Entropy)

3.2.2.1 Signal processing

EEG rhythms are often mixed with other biological signals, for instance EMG and EOG signals. The presence of these signals and other noise makes analysing EEG signals difficult. Thus, to enhance the system's performance, noise and unexpected artifacts need to be eliminated. To do this, in this study, raw EEG data are filtered by using an IIR highpass filter with a cut-off frequency of 2 Hz to get rid of low frequency artifacts and a notch filter at 50Hz to remove power noise. A visual artifact rejection method is applied to exclude epochs contaminated with artifacts. Segments containing significant artifacts are discarded, based on EMG and EOG signals.

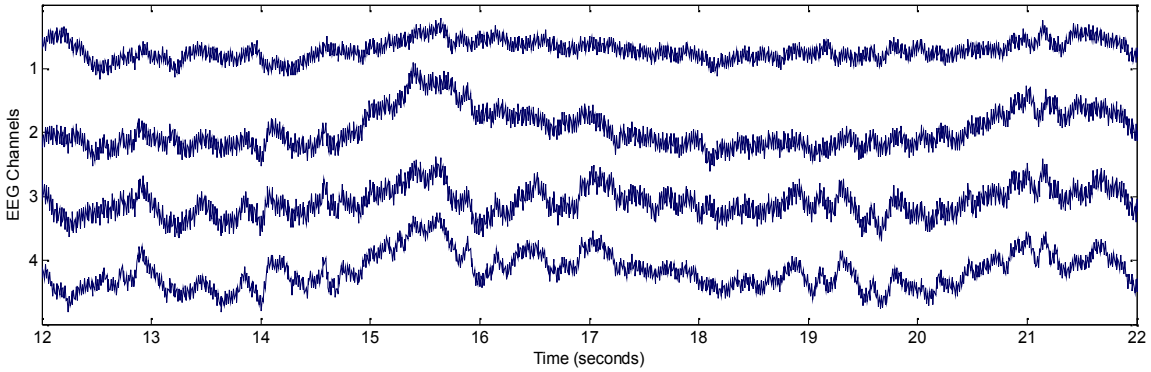


Figure 3.4: 10-second segment of raw EEG signals

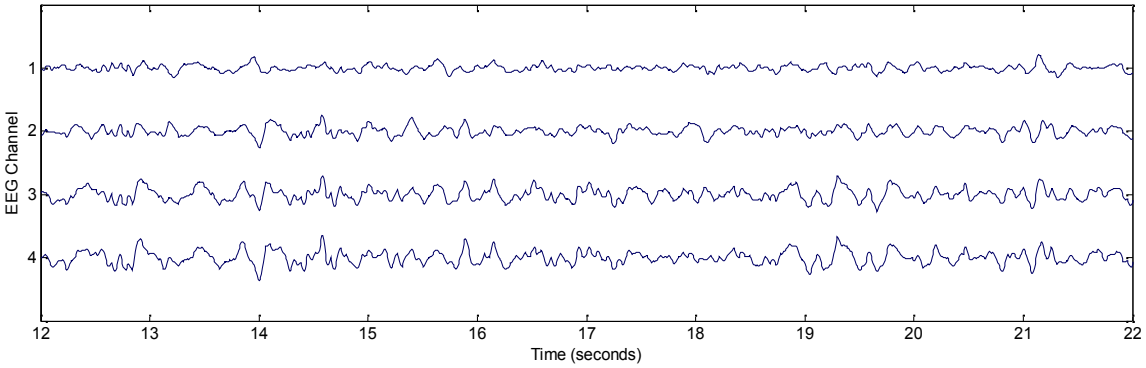


Figure 3.5: 10-second segment of EEG signals after being filtered

After being filtered, at each blood sampling point, a 40-second epoch of non-artifact EEG signals is extracted. Each epoch is labelled as Normal, Early Onset, Hypoglycemia or Recovery based on the corresponding phase of the study. Referring to Figure 3.1, each state is defined as shown in Table 3.1.

Table 3.1: Definition of patients' states

State	Corresponding Blood glucose level	Corresponding segment in Figure 3.1
Normal	> 3.9 mmol/l	ABB1
Early Onset	3.3 - 3.9 mmol/l	B1C
Hypoglycemia	≤ 3.3 mmol/l	CD
Recovery	> 3.3 mmol/l	DEF

Spectral analysis is applied to explore EEG signals in frequency domain. In signal processing, the Fourier Transform is the most common technique which provides the means of transforming a signal $x(t)$ defined in the time domain into $X(f)$ defined in the frequency domain:

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-i 2\pi f t} dt \tag{3.1}$$

In the case where both the time and frequency are discrete variables, the Discrete Fourier Transform (DFT) is applied:

$$X(mF) = \sum_{n} x(nT) e^{-inm 2\pi FT} \tag{3.2}$$

Considering the computational cost, the Fast Fourier Transform (FFT) is applied in this thesis to transform the EEG signals into frequency domain. Basically, FFT is a common technique of implementing the DFT with considerable savings in computational time. By using Fast Fourier Transform (FFT), the relationship between epoch length and frequency resolution is inversely proportional. A longer epoch length of signals in time domain leads

to a higher frequency resolution, allowing better identification of small shifts in EEG power spectrum. On the other hand, a longer epoch length results in a reduced number of data points available for analysing and an increase in time delay. This time delay is an essential factor for real-time systems, especially applications related to detecting health problems as in this thesis. Therefore, choosing a suitable epoch length plays an important role in analysing EEG signals. To do this, five different epoch lengths will be explored. The filtered 40-second epochs of EEG signals are subdivided into smaller segments of 5 different cases as follows:

(i) segments of 1-second

(ii) segments of 2-second

(iii) segments of 5-second

(iv) segments of 10-second

(v) segments of 20-second

By applying FFT, these segments are transformed into frequency domain which results in the power spectrum $P(f_i)$ for each segment. With the EEG sampling frequency of 128Hz, the frequency resolutions of the power spectrum are 1Hz, 0.5Hz, 0.2Hz, 0.1 Hz, and 0.05Hz, corresponding to epoch lengths of 1-second, 2-second, 5-second, 10-second and 20-second respectively. The power spectrum $P(f_i)$ is then subdivided into different frequency bands. In this thesis, the delta frequency band is not analysed because of its inherent characteristic of high artifact contamination. As a result, from the power spectrum of each EEG segment, three frequency bands will be obtained, including theta (θ: 4-8Hz), alpha (α: 8-13 Hz) and beta (β: 13-30Hz).

3.2.2.2 Feature extraction

The main aim of implementing feature extraction is to analyse the responses of various EEG parameters to hypoglycemic episodes and explore important features which can be used as inputs for the algorithm of hypoglycemia detection. This is an essential part in developing a system to detect hypoglycemia because valuable parameters can significantly enhance the success of the detection algorithm. However, this step has been skipped in previous studies of diagnosing and detecting health problems from EEG signals, including the application of detecting hypoglycemia (Juhl et al., 2010; Laione & Marques 2005). These studies usually stopped at employing the peak amplitude of each frequency bin of the EEG power spectra or the power level of frequency bands. This may lead to noisy input to the classification algorithm because of the highly unpredictable characteristic of EEG signals.

In this thesis, to characterize the spectrum within each band, EEG parameters are selected based on three main reasons: (i) the parameters have been proposed as valuable features in a variety of EEG research; (ii) the calculation of these parameters is straightforward and computationally efficient; (iii) the parameters give reliable results in the cases of analysing short EEG segments (1-second; 2-second; 5-second; 10-second; 20-second) thus being applicable in real-time patient monitoring systems, like in the application of detecting hypoglycemia.

As a result, from the power spectrum of each frequency band, four different EEG parameters are estimated as follows:

- ***Sub-band Power (P):*** The power level within each frequency band has been shown as a common feature in EEG research. Previous studies indicated that there were changes in the power level of theta and alpha bands caused by hypoglycemia occurrence (Bendtson et al. 1991; Bjørgaas et al. 1998; Howorka et al. 1996; Pramming et al. 1988). In this thesis, the power level within each frequency band is estimated from the power spectrum $P(f_i)$ by using a numerical integration technique (the trapezoidal rule).

- ***Centroid Frequency (CF):*** The centroid frequency of each frequency band can be interpreted as the center of gravity of the spectrum within each band (Dustman, Shearer & Emmerson 1999). It is estimated as the frequency which subdivides the area under the spectral curve within each band into identical parts.

$$CF = \frac{\sum_i f_i P(f_i)}{\sum_i P(f_i)} \tag{3.3}$$

The centroid frequency of alpha band has been drawing attention from a number of EEG studies in different areas. In many research areas, the alpha frequency activity was often represented by using peak alpha frequency (PAF) which measures the discrete frequency with the highest magnitude within the alpha range. Previous studies have indicated that PAF increased from infants to adults, and then started to decline with age (Niedermeyer 1999; Richard Clark et al. 2004; Stroganova, Orekhova & Posikera 1999). Centroid alpha frequency, which measures the center of gravity rather than the peak, within boundaries of the alpha band has been shown to be a more accurate measure of the distribution of alpha band than PAF (Klimesch 1997). Estimating the centroid frequency within the boundary of alpha frequency for each individual and calling it the individual alpha frequency (IAF), Klimesch and his group reported that this feature relates to response time and speed of processing information (Klimesch et al. 1996), as well as cognitive and memory performance (Klimesch 1999).

- ***Spectral Variance (SV):*** This feature is a measure of how the power spectrum is spread out within each frequency band (Dustman, Shearer & Emmerson 1999). This is a good indicator of changes in the power spectrum distribution of each frequency band.

$$SV = \frac{\sum_i f_i^2 P(f_i)}{\sum_i P(f_i)} - CF^2 \tag{3.4}$$

- ***Spectral Entropy (SE):*** This feature is a measure of the distribution of normalized power spectrum within a frequency range. This feature reflects the distribution in the power spectrum. It reaches to maximum when all frequencies in the power spectrum have the same power level. In the case that power spectrum concentrates in a smaller frequency range, the SE will decrease.

$$SE = -\frac{1}{\log N_f}\sum_f P_n(f_i)\log P_n(f_i) \tag{3.5}$$

where N_f is the number of frequency bins within each frequency band.

As a result, a total of 48 EEG features (4 different features x 3 frequency bands x 4 channels) are estimated for each epoch. The vector of extracted features for each epoch is considered as a data point. For comparison and classification purposes, four sets of data are extracted, corresponding to four states of patients during the study, including Normal, Early Onset, Hypoglycemia and Recovery. Referring to Figure 3.1, Normal is defined as segment ABB1, Early Onset is defined as segment B1C, Hypoglycemia is defined as segment CD and Recovery is defined as segment DEF. To reduce the variability in data, each final data point in each set is estimated as the average of two consecutive non-overlapping points.

3.2.2.3 Statistical analysis

After being extracted, statistics is applied to compare and determine the significance of changes in EEG parameters under the transition of patients' state during the glucose clamp study. First, descriptive analysis is carried out to assess each parameter's data distribution like mean, standard deviation, normality, skewness, etc. To compare EEG responses between three different states of Normal, Hypoglycemia and Recovery, analysis of differences between pairs of groups is performed using *t*-test for features with normal distribution and the nonparametric Wilcoxon test for features with non-normal distribution. The correlations between EEG parameters and actual BGLs during the study are also analysed by using a nonparametric ranking test.

Besides, statistical analysis is also implemented with the data from Early Onset data set which corresponds with BGL range of 3.3-3.9 mmol/l to explore the differences of EEG responses at two BGL thresholds of 3.3 mmol/l and 3.9 mmol/l.

In all analyses, probability values (*p*-values) less than 0.05 are considered to be significant.

3.2.3 Standard neural network for hypoglycemia detection

An artificial neural network (conventionally named a neural network) can be defined as a mathematical model inspired by the human nervous system which is used for modelling complex non-linear relationships between inputs and outputs or to find patterns in data. Generally, it is a network of simple processing elements (neurons) which exhibits complex global behaviour determined by the connections between the processing elements as well as between parameters of these elements. By learning from observed or training data, neural networks can adjust and generalise its structure during a training phase. This approach has been employed popularly in various complex biomedical applications as a powerful tool of classification and pattern recognition.

3.2.3.1 Neural network structure

Since the first form of neural network (the McCulloch-Pitts threshold neuron) was proposed in 1943 (McCulloch & Pitts 1943), a variety of neural network structures has been developed. Single-layer neural networks, with threshold activation functions, were presented by Rosenblatt in 1958 who called them perceptrons (Rosenblatt 1958). At the same time as Rosenblatt introduced the perceptron, Widrow and co-workers also worked along a similar direction to introduce the adaptive linear element (Widrow & Hoff 1960). This is a single processing unit with threshold non-linearity which is essentially the same as the perceptron. It became clear later that these kinds of single-layer neural networks are incapable of handling non-linear separation problems.

Until now, a variety of network topologies have been also introduced with the aim of enhancing the performance of function approximation and learning capability of neural networks. The multilayer feed-forward neural network is the most popular network topology which is constructed by multiple layers of perceptrons (Widrow & Lehr 1990). The feed-forward network mapping allows data to travel one way only from input layer, through hidden layers to output layer. The number of units (or neurons) in the input layer and output layer is determined by each unique application. The number of hidden layers and the number of neurons in each hidden layer are determined by different factors like the developer's experience, the application's requirement. Recently, different statistical theories (such as Bayesian theory) as well as evolutionary algorithms have been applied to develop frameworks to find the most proper structure of neural networks (Leung et al. 2003; Mackay 1992; Penny & Roberts 1999).

In this thesis, considering the final aim of developing a real-time detecting system which requires reducing the computational cost, a neural network with feed-forward three-layer topology is developed and employed as the classification unit. The structure of the neural network is shown in Figure 3.6. There are three layers in the network. The input layer includes features extracted from EEG signals. The output layer consists of one output node. The desired output (or target) is set at 1 in case of hypoglycemia and -1 in case of non-hypoglycemia.

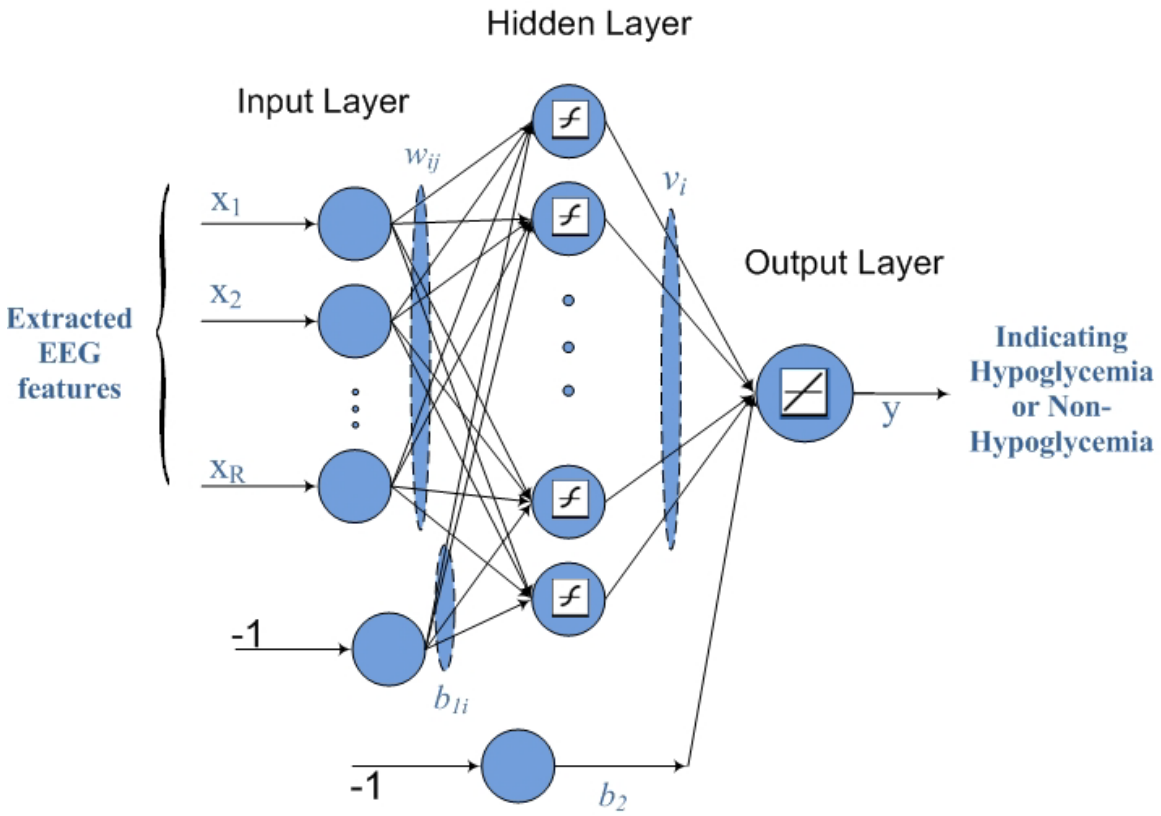


Figure 3.6: Neural network structure

From Figure 3.6, the input-output relationship of the developed neural network can be written as follows:

$$y = \sum_{i=1}^{S} v_i tansig\left[\sum_{j=1}^{R}\left(w_{ij}x_j - b_{1i}\right)\right] - b_2 \tag{3.6}$$

where

- S is the number of hidden nodes
- $v_i, i = 1,2 \ldots, S$, is the weight of the link between i-th hidden node and the output
- R is the number of inputs
- w_{ij}, $i = 1 \div S, j = 1 \div R$, is the weight of the link between i-th hidden node and the j-th input.
- b_1, b_2 are the biases for the hidden nodes and output nodes respectively.

- The total number of network parameters is calculated as:

$$\begin{aligned} \text{Number of parameter} &= (R+1)*S+S+1 \\ &= (R+2)*S+1 \end{aligned} \tag{3.7}$$

- *tansig* is the hyperbolic tangent sigmoid transfer function of hidden layer

$$tansig(a) = \frac{e^{a} - e^{-a}}{e^{a} + e^{-a}} \tag{3.8}$$

- Transfer function of output layer is a linear function: $y(a) = a$

3.2.3.2 The Levenberg-Marquardt algorithm for training neural network

One of the most important issues in developing neural network is the learning or training process of the network. Training a neural network essentially means finding a set of network parameters that optimise a cost function in order to achieve the best network performance. Most of the algorithms used in training artificial neural networks employ some forms of gradient descent. This is done by simply taking the derivative of the cost function with respect to the network parameters and then changing those parameters in a gradient-related direction.

The Levenberg-Marquardt algorithm is a well-known algorithm for training neural networks which estimates the second directional derivative of the cost function in order to direct the training process to a local minimum. It has become a standard technique for non-linear least-squares problems, widely adopted in various disciplines for dealing with data-fitting applications. In this chapter, the Levenberg-Marquardt algorithm is selected to be used as the network training algorithm because it is one of the fastest training methods and provides a steady convergence capability for the network training process. Details about the algorithm are provided as follows.

Let the input vector be *x* and its associated targeted output vector be *t* and suppose that the

neural network produces an actual output vector of y. An objective cost function which will be used to train the neural network is defined as the squared error $E(w) = (y - t)^2$. This function can be expressed in a quadratic form as follows:

$$\begin{aligned} E(w) &= \frac{1}{2}\sum_{n=1}^{N} \{y(x^n, w) - t^n\}^2 \\ &= \sum_{n=1}^{N} e^2(w) = e^T(w)\, e(w) \end{aligned} \tag{3.9}$$

where N is the number of data points of the training set.

The aim of the Levenberg-Marquardt algorithm is to compute the network weight vector w at which $E(w)$ is minimum. To do this, the weight vector w_{k+1} will be updated from the previous weight vector w_k based on the second-order Taylor series as follows:

$$E(w_{k+1}) = E(w_k + \Delta w_k) \approx E(w_k) + \frac{\partial E(w_k)}{\partial w}\Delta w_k + \frac{1}{2}\Delta w_k^T \frac{\partial^2 E(w_k)}{\partial w^2}\Delta w_k \tag{3.10}$$

A local minimum of the error function can be reached by taking the gradient of the function with respect to Δw_k and setting it equal to zero, leading to:

$$w_{k+1} = w_k + A_k^{-1} g_k \tag{3.11}$$

where $g_k = \nabla E(w)|_{w=w_k}$, and $A_k = \nabla^2 E(w)|_{w=w_k}$

Both the gradient and the second order of the error function would be obtained through the application of the chain rule and the multiplication rules.

First, the gradient would be:

$$\nabla E(w) = \frac{\partial E(w)}{\partial w} = 2\sum_{i=1}^{N} e(w)\frac{\partial e(w)}{\partial w} \tag{3.12}$$

In the matrix form:

$$\nabla E(w) = 2J^T(w)\,e(w) \tag{3.13}$$

where $J(w)$ is the Jacobian matrix which can be expressed as:

$$J(w) = \begin{bmatrix} \frac{\partial e_1(w)}{\partial w_1} & \frac{\partial e_1(w)}{\partial w_2} & \cdots & \frac{\partial e_1(w)}{\partial w_n} \\ \frac{\partial e_2(w)}{\partial w_1} & \frac{\partial e_2(w)}{\partial w_2} & \cdots & \frac{\partial e_2(w)}{\partial w_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial e_N(w)}{\partial w_1} & \frac{\partial e_N(w)}{\partial w_2} & \cdots & \frac{\partial e_N(w)}{\partial w_n} \end{bmatrix}$$

Next, the second order of the error function forms the Hessian matrix whose k, j element would be:

$$\left[\nabla^2 E(w)\right]_{k,j} = \frac{\partial^2 E(w)}{\partial w_k \partial w_j} = 2\sum_{i=1}^{N}\left\{\frac{\partial e(w)}{\partial w_k}\frac{\partial e(w)}{\partial w_j} + e(w)\frac{\partial^2 e(w)}{\partial w_k \partial w_j}\right\} \tag{3.14}$$

In the matrix form:

$$A = \nabla^2 F(w) = 2J^T(w)J(w) + 2S(w) \tag{3.15}$$

where $S(w) = \sum_{i=1}^{N} e_i(w)\nabla^2 e_i(w)$

It is usually reasonable to assume that errors in the training set are fairly independent and identically distributed around a mean of zero. Thus, the term $S(w)$ can be ignored, and the Hessian matrix can be approximated as follows:

$$\nabla^2 F(x) \approx 2J^T(w)J(w) \tag{3.16}$$

Substitute equation (3.15) and equation (3.16) into the equation (3.11), we obtain:

$$w_{k+1} = w_k - \left[J^T(w_k) J(w_k) \right]^{-1} J^T(w_k) e(w_k) \tag{3.17}$$

This update method is known as Gauss-Newton. Its advantage over the standard Newton's method is that it does not require calculation of second derivatives. However, the matrix Hessian $A = J^T J$ might not be invertible. In order to overcome this, an invertible approximate Hessian matrix is introduced:

$$G = A + \mu I \tag{3.18}$$

where I is a unit matrix and μ is the scaling factor that will be updated during the training process.

By increasing μ until a large enough number, the matrix G is certainly invertible, and this leads to the Levenberg – Marquardt algorithm:

$$w_{k+1} = w_k - \left[J^T(w_k) J(w_k) + \mu_k I \right]^{-1} J^T(w_k) e(w_k) \tag{3.19}$$

or

$$\Delta w_k = -\left[J^T(w_k) J(w_k) + \mu_k I \right]^{-1} J^T(w_k) e(w_k) \tag{3.20}$$

In summary, for the purpose of training neural networks, the procedure of the Levenberg – Marquardt algorithm which aims to minimize the cost function $E(w)$ can be presented as follows:

i. Compute $E(w_k)$

ii. Initialise μ_k with a small value

iii. Solve for w_{k+1} to compute $E(w_{k+1})$

iv. If $E(w_{k+1}) \geq E(w_k)$ then increase μ_k by a factor of γ (e.g $\gamma = 10$), then go to iii.

v. If $E(w_{k+1}) < E(w_k)$ then decrease μ_k by a factor of γ, then go to iii.

3.2.3.3 Cross-validation technique for training neural network

Generalisation is one of the most important factors when assessing the performance of neural network. In terms of classification, the generalisation of a network means that after being trained, the network could classify data from the same class as the learning data that it has never seen before. Ideally, the network is trained until a desired target of cost function is gained. However, this process may lead to a prominent shortcoming of training algorithms which is over-fitting (or overtraining). This is the situation when the network is over-trained (training time is too long, the training cost function is forced to be a very small value). In this case, the performance of neural network on the training data has still improved while the performance on unseen data becomes worse, leading to a poor ability for generalisation of the network (as shown in Figure 3.6).

In order to constrain the limitation of over-training network and enhance the network generalisation, in this thesis, the cross-validation technique is applied. To do this, the available data are divided into three separate subsets: a training set, a validation set and a testing set. The training set is used to train the network by updating the network weights and biases through minimizing a cost function (error function). During the training process, the error function on the validation set is also monitored. The validation error normally decreases during the initial phase of training, as does the training set error. However, when the network begins to over-fit the data, the error on the validation set typically begins to rise while the training error still decreases (the Over-fitting point on Figure 3.6). When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are returned as the final network parameters. The testing set is separated with the other two subsets and only used for testing the performance of the final network as a data set which is totally unseen by the training process.

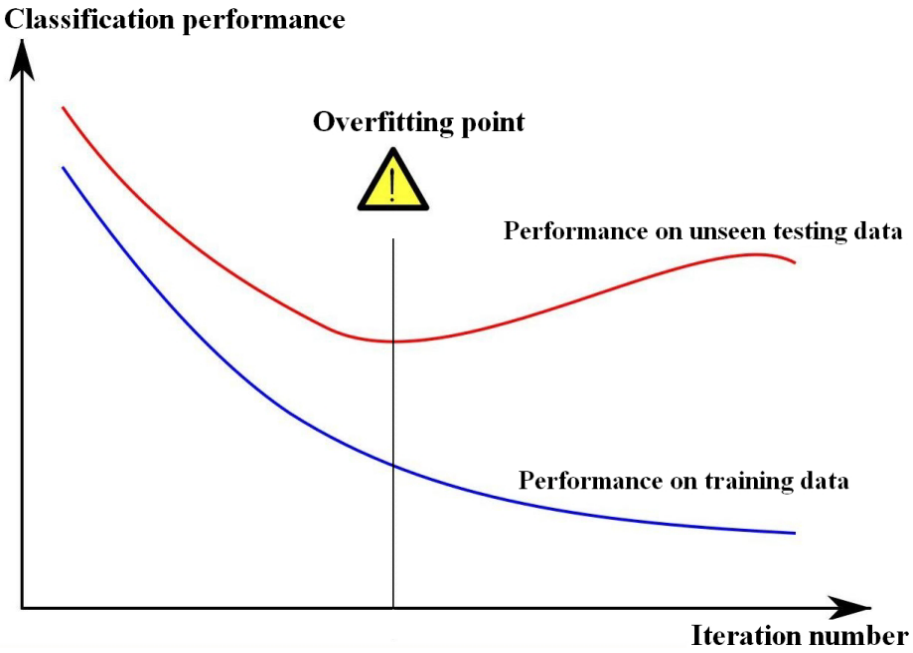


Figure 3.7: Neural network over-fitting

3.2.3.4 Definition of classification performance

Sensitivity and specificity

After being trained by Levenberg-Marquardt algorithm and cross-validation technique, the final structure and parameters (including w_{ij} ; b_{1i} ; v_i ; b_2 as presented in section 3.2.3.1) of the developed neural network are determined. The classification performance of the final network will be estimated by evaluating the sensitivity and specificity of the neural network on each data set. Specifically, for the application of detecting hypoglycemic episodes, two criteria of sensitivity and specificity are defined as follows:

$$Sensitivity = \frac{TP}{TP + FN}$$
$$Specificity = \frac{TN}{TN + FP} \tag{3.21}$$

where:

- True Positive (*TP*) is the number of hypoglycemic episodes which are correctly classified as hypoglycemia
- True Negative (*TN*) is the number of non-hypoglycemic episodes which are correctly classified as non-hypoglycemia.
- False Positive (*FP*) is the number of non-hypoglycemic episodes which are wrongly classified as hypoglycemia
- False Negative (*FN*) is the number of hypoglycemic episodes which are wrongly classified as non-hypoglycemia.

In effect, sensitivity and specificity represent two types of accuracy: sensitivity is the accuracy for actual hypoglycemic cases and specificity is the accuracy for actual non-hypoglycemic cases. These two parameters are important criteria for assessing performance of diagnosing and detecting human health problems. In most cases, the compromise between these two parameters needs to be considered in order to achieve unique requirements of each application.

In the application of hypoglycemia detection for T1DM patients, especially for patients with hypoglycemic unawareness, the sensitivity, which represents the rate of correctly detecting hypoglycemic episodes, is more important than the specificity. By selecting the linear transfer function $y(a) = a$ for neural network output layer (as presented in section 3.2.3.1), the trade-off between sensitivity and specificity can be adjusted in order to satisfy the mentioned demand of the application by means of plotting the Receiver Operating Characteristic curve for the combined training and validation set.

Receiver Operating Characteristic Curve

By definition, a Receiver Operating Characteristic curve (ROC curve) presents the compromise between the true positive rate versus false positive rate (equivalently, sensitivity versus 1–specificity) for different thresholds of the classifier output (Figure 3.8). That means each point on the curve is corresponding to one specific output threshold. Utilizing the characteristic of compromising between sensitivity and specificity for each output threshold, ROC curve is used to find the output threshold that can lead to the desired classification performance. It is noted that in the application of hypoglycemia detection, the sensitivity, which represents the rate of correctly detecting hypoglycemia episodes, is more important than the specificity. Therefore, in this thesis, based on the plotted ROC curve, a criterion for output threshold is set at the point producing classification sensitivity of 80%. This procedure leads to desired sensitivity and relatively low but reasonable specificity for the application of hypoglycemia detection.

The area under ROC curve (*AuC*) is also an important measure that presents the performance of the classification. This parameter can be used to compare the performance of different classifiers which are trained by using the same training set. The higher the *AuC*, the better the classifier. A random classification gives an *AuC* of 0.5, while an ideal classification gives an *AuC* of 1 (as shown in Figure 3.8). In this thesis, besides sensitivity and specificity, *AuC* is also used as a criterion for assessing classification performance.

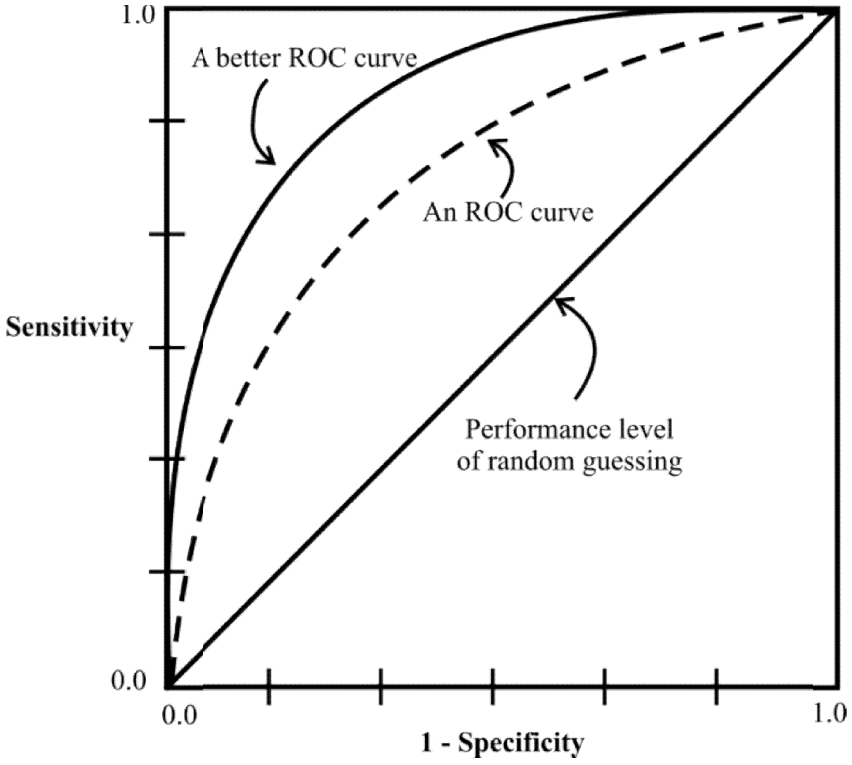


Figure 3.8: ROC curve

3.2.3.5 Selecting the number of hidden nodes for neural network

Determining the size of the hidden layer is an important issue when developing feed-forward multi-layer neural networks in real-life applications. It is obvious that a smaller network size will help reduce the computational cost which is one of the main factors to be considered when developing real-time systems. However, a neural network with a small structure may have less power in modelling the input-output function. On the other hand, a too big architecture may lead to a huge computation cost as well as over-fitting which causes bad fitting when totally unseen data are applied to the neural network. In practice, the size of the hidden layer or the number of hidden nodes is usually determined experimentally based on the dimension of the input and output spaces as well as on the different requirements of each application.

In this thesis, in order to determine the hidden layer size which can produce better performance, the number of hidden nodes is varied from 1 to 15. As a result, 15 different neural network architectures are developed and trained by using the same training set with the same training procedure using the Levenberg-Marquardt algorithm and cross-validation technique. ROC curve is plotted for each case in order to select the output threshold of each classifier. Comparisons between developed neural networks are made based on the classification performance of *AuC,* sensitivity and specificity. In order to make the comparison easier, the cut-off point that produces a sensitivity of 80% will be selected as the output threshold for each neural network. With the same sensitivity, the higher *AuC* and specificity lead to the higher performance of the neural network.

3.3 Results

3.3.1 Blood glucose profile of patients during the glucose clamp study

During the glucose clamp study which was implemented at the Princess Margaret Hospital for Children in Perth, blood glucose levels (BGLs) of five T1DM patients were collected with the blood sampling period of 5 minutes. Details of the study protocols are described in Chapter 3, section 3.2.1. The BGL profile of each patient is used as reference to determine the corresponding state of the patient at each blood sampling point. The definition of each state is provided in Table 3.1. The actual BGL profiles of five patients are plotted in Figure 3.9.

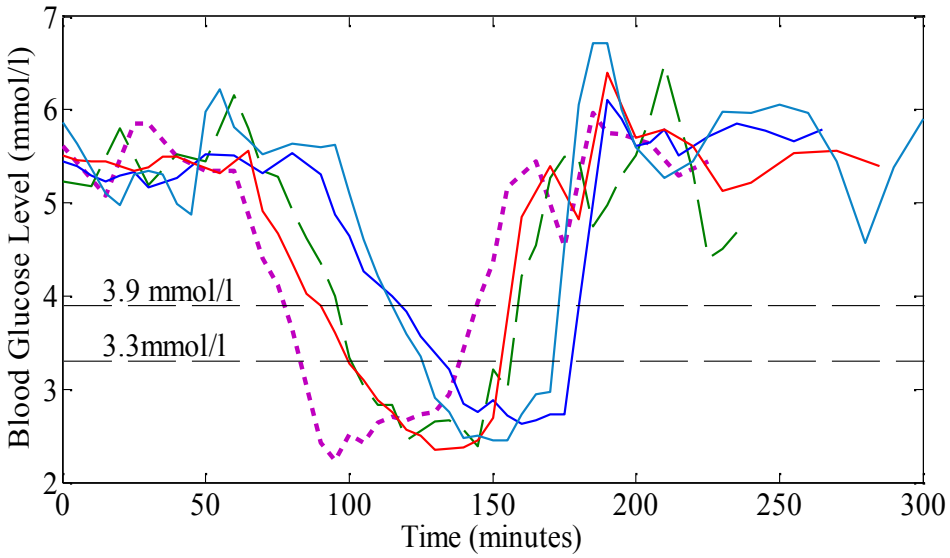


Figure 3.9: Blood glucose level profiles during the glucose clamp study of five T1DM patients

3.3.2 EEG feature extraction results

The main objective of implementing feature extraction is to analyse the responses of various EEG features or parameters in order to determine the most important features that significantly change under hypoglycemic conditions. These extracted EEG features later

will be employed as inputs of classification algorithms for detecting hypoglycemic episodes. In order to do this, three steps of analysis are carried out:

- First, group comparisons between three phases of the clinical hypoglycemia-associated study including Normal, Hypoglycemia and Recovery will be implemented. The correlation of each EEG parameter with the BGL transition during the insulin-induced study will be explored to reassure the results of group comparisons.

- Second, EEG parameters with 5 different epoch lengths of 1-second, 2-second, 5-second, 10-second and 20-second are analysed to find the most suitable epoch length of signals for the application of detecting hypoglycemia.

- Third, data at the BGL range of 3.3-3.9 mmol/l (which is named as the Early Onset phase of the study) will be analysed to figure out the responses of EEG parameters to the potentially early onset of hypoglycemia as well as to find a BGL threshold to distinguish between non-hypoglycemic and hypoglycemic states.

3.3.2.1 Responses of EEG parameters to different blood glucose levels during the glucose clamp study

As presented in section 3.2.2.2, at each EEG channel, 12 EEG parameters are extracted from the data acquired during the overnight glucose clamp study and categorized into three groups corresponding to three different phases of the study including Normal, Hypoglycemia and Recovery. Statistical results of the comparison on each parameter at each channel between three phases of Normal, Hypoglycemia and Recovery are presented in Table 3.3-3.5. Significant tests are reported in bold. The EEG responses of all five patients show significant changes during the study.

The centroid alpha frequency is shown to be the most significant feature which is highly correlated to BGLs during the study at all four EEG channels (p=0.001 at C3, p=0.004 at

C4 and $p<0.0001$ at O1 and O2). Group comparisons also indicate a significant decrease in the centroid alpha frequency under hypoglycemic conditions, and then a re-establishment under recovery conditions, as shown in Figure 3.10. There are significant changes in two other features of alpha bands which are spectral variance and spectral entropy at all channels. Under hypoglycemic conditions, decreases in these two features show that the spectrum of the alpha band tends to concentrate on a narrower range. Based on the mentioned results, it is established that during the occurrence of hypoglycemia, in the power spectrum of alpha band, there is a shift toward smaller frequencies as shown in Figure 3.11 where an example of alpha spectrums at two different states is presented.

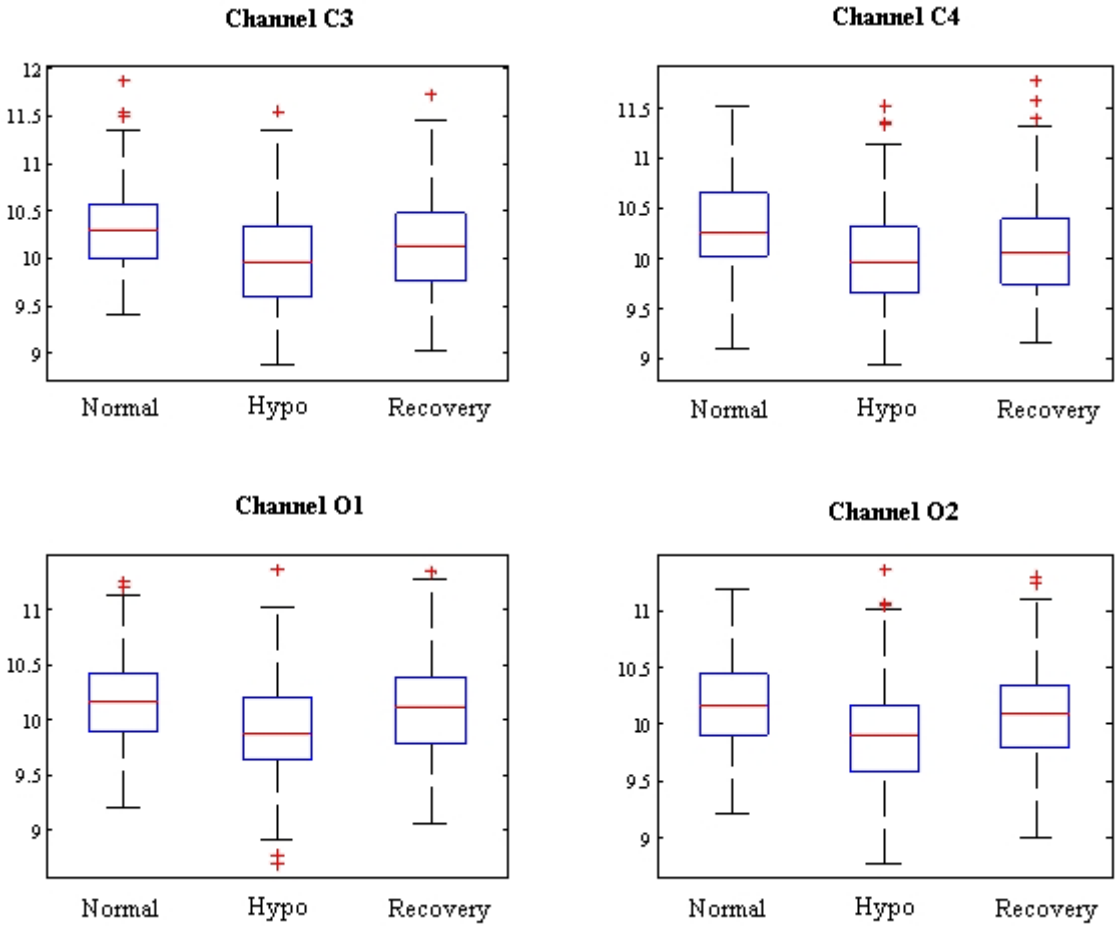


Figure 3.10: Comparison of centroid alpha frequency between three phases of the study at four EEG channels

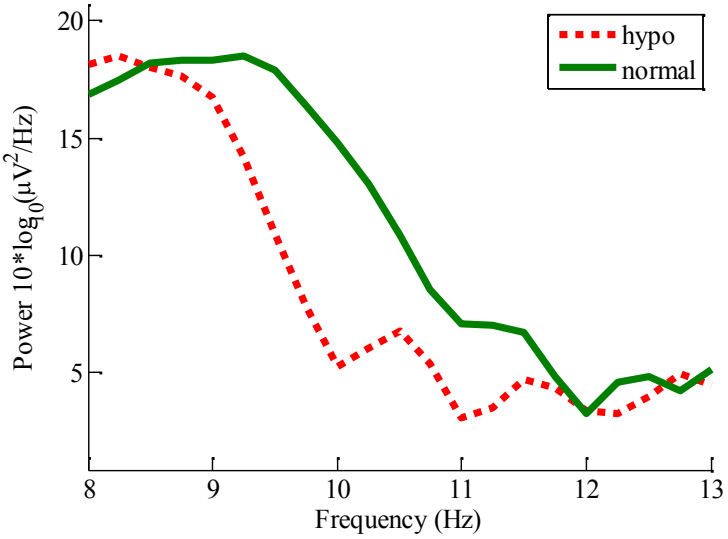


Figure 3.11: Example of changes in alpha spectrum

The statistical results also produce a slightly significant increase in centroid theta frequency at channel O1 and O2 as shown in Figure 3.12 (p=0.01 at O1 and 0.003 at O2). It is shown that the responses in theta band are more significant at channel O1 and O2 than channels C3 and C4.

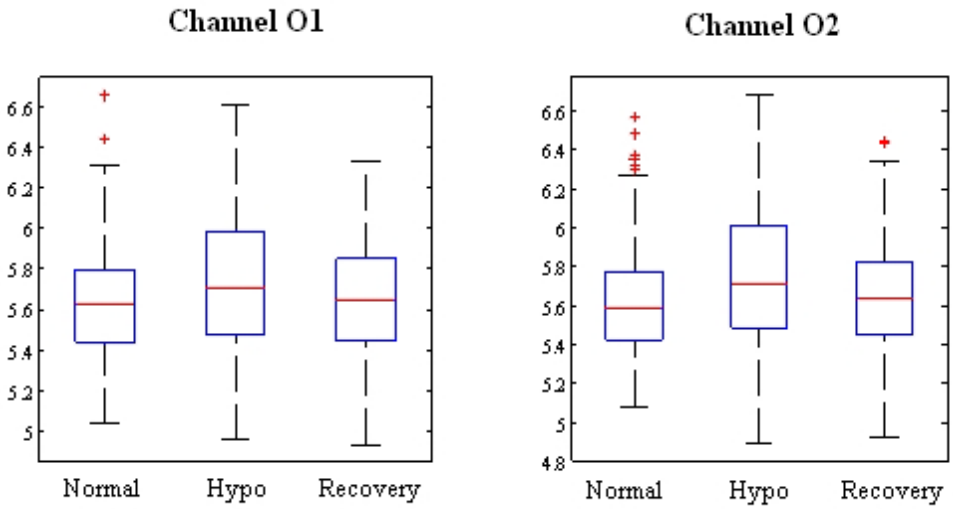


Figure 3.12: Comparison of centroid theta frequency between three phases of the study at channels O1 and O2

The correlation analysis indicates no change in power levels in all 3 frequency bands at all 4 channels except some slight changes in the power level of theta band at channel C3 (p=0.012); alpha band at channels O1 (p=0.026) and O2 (p=0.026); and beta band at channel C4 (p=0.003). However, these changes are not consistent with other channels, and with the group comparison results which provide no difference between pairs of groups. Moreover, when analysing data from each patient, it is shown that the changes in the power level are also not consistent with all patients (increasing in some patients, decreasing in the other patients). Based on these analyses, it is established that these changes in the power level at some channels are caused by body movements as well as changes in sleep stages of patients during night.

Comparisons between data from Normal group and Recovery group indicate similarities in most of the features at all four channels. These results demonstrate that during the study, EEG responses from five patients at all channels significantly change under hypoglycemic conditions and re-establish under recovery conditions. However, it should be noted that the difference between Normal group and Hypoglycemia group is slightly stronger than the difference between Recovery group and Hypoglycemia group. These results probably reaffirm conclusions from previous studies that the frequent exposure to hypoglycemia can cause alterations in EEG signals and lead to impairments to the brain functioning of patients with T1DM (Bjørgaas, Sand & Gimse 1996; Clarke et al. 2009; Cryer 2007; Hyllienmark et al. 2005).

The statistical results also indicate similarities in EEG responses between channels in the same brain area (i.e. O1 and O2 in occipital area; C3 and C4 in the central area). Consequently, it is established that there is no significant difference in EEG responses between channels lying in the same brain area but on different brain hemispheres (left or right sides).

Based on the above analyses, it is concluded that the responses of the alpha spectrum, especially the decrease in centroid alpha frequency under hypoglycemic conditions are the most significant findings of the EEG feature extraction step implemented in this thesis. This feature can be explained by the lack of vigilance or awareness which is a common

symptom normally happening in T1DM patients under the occurrence of hypoglycemic conditions (Howorka et al. 1996; Howorka et al. 2000). The changes in the alpha spectrum have been reported to be associated with the vigilance level of healthy individuals as well as people with various health problems, as shown in Table 3.2. In this context, vigilance can be defined as the readiness of a person to have appropriate behaviours, both in quality and quantity, in response to a given (internal or external) stimulus situation; as well as the ability to remain alert to that stimuli for an acceptable period of time (Howorka et al. 1996; Shi & Lu 2008).

Table 3.2: Reports of the correlation between the alpha spectrum and vigilance level

Research Group	Reported conclusions
Klimesch et al. (1990) Passero et al. (1995)	- Patients with Alzheimer's disease (AD) have lower IAF (or PAF) compared to age-matched controls
Klimesch et al. (1993)	- During increasing memory demands, individuals with lower memory performance decreased their IAF while controlled individuals held their IAF constant.
Angelakis et al. (2004)	- PAF is proposed as an indicator for cognitive preparedness - Subjects with traumatic brain injury have lower PAF than normal subjects during resting after a working memory task.
Billiot et al. (1997)	- In patients with chronic fatigue syndrome, PAF is negatively correlated with total fatigue and 'today fatigue' reports.
Jarm et al. (2007)	- The reduction of CAF is suggested as an indicator of physical fatigue (drivers' fatigue).

Individual alpha frequency (IAF), Peak alpha frequency (PAF), Centroid alpha frequency (CAF) are defined in 3.2.2.2

Table 3.3: Comparison between three phases of Normal, Hypoglycemia and Recovery using data from channel C3

Parameters	**Correlation to BGLs**		**Mean ± SD**			**Groups comparison *p*-values**		
	r	***p*-values**	**Normal (N)**	**Hypoglycemia (H)**	**Recovery (R)**	**N-H**	**H-R**	**N-R**
P-θ	-0.098	**0.012**	2.105 ± 1.328	2.141 ± 1.698	1.846 ± 1.165	0.526	0.272	0.051
P-α	-0.021	0.585	1.552 ± 1.333	1.611 ± 1.123	1.624 ± 1.659	0.202	**0.006**	0.143
P- β	-0.064	0.075	1.038 ± 0.633	1.238 ± 1.221	0.955 ± 0.719	0.452	0.056	0.115
CF-θ	-0.041	0.294	5.628 ± 0.302	5.659 ± 0.331	5.610 ± 0.290	0.379	0.265	0.876
CF-α	0.156	**0.001**	10.255 ± 0.460	9.993 ± 0.495	10.132 ± 0.510	**< 0.001**	**0.004**	**0.005**
CF- β	-0.074	0.055	17.411 ± 1.468	17.857 ± 1.554	17.687 ± 1.445	**0.014**	0.527	0.057
SV-θ	-0.032	0.411	1.132 ± 0.229	1.163 ± 0.227	1.149 ± 0.247	0.208	0.534	0.562
SV-α	0.057	0.143	1.844 ± 0.488	1.688 ± 0.550	1.768 ± 0.491	**0.002**	0.104	0.099
SV- β	-0.070	0.069	14.513 ± 5.506	15.470 ± 6.309	15.538 ± 5.370	**0.023**	0.663	**0.013**
SE-θ	-0.001	0.987	0.858 ± 0.032	0.859 ± 0.034	0.860 ± 0.038	0.672	0.454	0.246
SE-α	0.091	**0.019**	0.856 ± 0.044	0.830 ± 0.064	0.847 ± 0.051	**<0.0001**	**0.013**	0.060
SE- β	0.055	0.157	0.800 ± 0.076	0.812 ± 0.075	0.808 ± 0.078	0.251	0.649	0.106

Table 3.4: Comparison between three phases of Normal, Hypoglycemia and Recovery using data from channel C4

Parameters	**Correlation to BGLs**		**Mean ± SD**			**Groups comparison *p*-values**		
	r	***p*-values**	**Normal (N)**	**Hypoglycemia (H)**	**Recovery (R)**	**N-H**	**H-R**	**N-R**
P-θ	-0.067	0.071	2.059 ± 1.273	2.114 ± 1.300	1.991 ± 1.560	0.656	0.385	0.605
P-α	-0.038	0.305	1.634 ± 1.047	1.695 ± 1.084	1.754 ± 1.367	0.557	0.628	0.281
P- β	-0.113	**0.003**	1.205 ± 0.562	1.376 ± 0.996	1.166 ± 0.679	0.118	**0.041**	0.489
CF-θ	-0.034	0.377	5.594 ± 0.259	5.644 ± 0.323	5.630 ± 0.289	**0.046**	0.706	**0.009**
CF-α	0.111	**0.004**	10.240 ± 0.449	10.006 ± 0.478	10.104 ± 0.490	**< 0.0001**	**0.035**	**0.037**
CF- β	-0.055	0.153	17.278 ± 1.378	17.763 ± 1.589	17.637 ± 1.555	**0.042**	0.761	0.056
SV-θ	0.017	0.663	1.133 ± 0.239	1.123 ± 0.235	1.149 ± 0.234	0.884	0.360	0.197
SV-α	-0.039	0.311	1.814 ± 0.456	1.855 ± 0.491	1.806 ± 0.478	0.395	0.190	0.557
SV- β	-0.013	0.741	14.493 ± 5.625	14.993 ± 6.195	15.101 ± 5.461	0.573	0.939	0.487
SE-θ	0.076	0.051	0.862 ± 0.030	0.856 ± 0.034	0.862 ± 0.037	0.118	**0.014**	0.264
SE-α	0.096	**0.013**	0.856 ± 0.036	0.845 ± 0.045	0.853 ± 0.042	**0.004**	**0.049**	0.343
SE- β	0.069	0.076	0.798 ± 0.075	0.808 ± 0.075	0.805 ± 0.085	0.214	0.910	0.065

Table 3.5: Comparison between three phases of Normal, Hypoglycemia and Recovery using data from channel O1

Parameters	Correlation to BGLs		Mean ± SD			Groups comparison *p*-values		
	r	***p*-values**	**Normal (N)**	**Hypoglycemia (H)**	**Recovery (R)**	**N-H**	**H-R**	**N-R**
P-θ	-0.070	0.075	2.218 ± 1.507	2.125 ± 1.427	1.965 ± 1.205	0.690	0.249	0.121
P-α	-0.086	**0.026**	1.946 ± 2.076	2.221 ± 2.931	2.325 ± 3.839	**0.011**	0.071	0.316
P- β	-0.015	0.707	1.162 ± 0.737	1.164 ± 0.693	0.967 ± 0.517	0.719	0.103	0.109
CF-θ	-0.086	**0.013**	5.661 ± 0.288	5.739 ± 0.352	5.656 ± 0.294	**0.027**	**0.042**	0.767
CF-α	0.178	**<0.0001**	10.141 ± 0.394	9.876 ± 0.453	10.110 ± 0.434	**<0.0001**	**<0.0001**	0.398
CF- β	0.012	0.766	17.027 ± 1.122	17.068 ± 0.903	17.283 ± 1.209	0.424	0.066	0.013
SV-θ	-0.049	0.206	1.137 ± 0.233	1.199 ± 0.242	1.151 ± 0.262	**0.011**	0.057	0.579
SV-α	0.074	0.055	1.757 ± 0.470	1.542 ± 0.601	1.713 ± 0.488	**<0.001**	**0.010**	0.228
SV- β	0.058	0.132	13.935 ± 5.519	13.395 ± 5.180	13.516 ± 4.988	0.154	0.093	0.358
SE-θ	-0.007	0.851	0.858 ± 0.031	0.859 ± 0.034	0.861 ± 0.034	0.628	0.527	0.214
SE-α	0.202	**<0.0001**	0.857 ± 0.046	0.813 ± 0.081	0.852 ± 0.046	**<0.0001**	**<0.0001**	0.159
SE- β	0.84	**0.030**	0.789 ± 0.070	0.795 ± 0.060	0.799 ± 0.072	0.522	0.069	0.084

Table 3.6: Comparison between three phases of Normal, Hypoglycemia and Recovery using data from channel O2

Parameters	**Correlation to BGLs**		**Mean ± SD**			**Groups comparison *p*-values**		
	r	***p*-values**	**Normal (N)**	**Hypoglycemia (H)**	**Recovery (R)**	**N-H**	**H-R**	**N-R**
P-θ	-0.043	0.269	2.208 ± 1.727	2.047 ± 1.388	1.897 ± 1.271	0.908	0.334	0.259
P-α	-0.105	**0.010**	2.258 ± 2.618	3.363 ± 5.737	2.644 ± 4.212	0.146	0.033	0.471
P- β	-0.033	0.389	1.121 ± 0.733	1.201 ± 0.825	1.059 ± 0.612	0.922	0.050	0.137
CF-θ	-0.116	**0.003**	5.650 ± 0.275	5.754 ± 0.377	5.647 ± 0.290	**0.008**	**0.011**	0.695
CF-α	0.178	**< 0.0001**	10.133 ± 0.392	9.863 ± 0.454	10.083 ± 0.432	**<0.0001**	**<0.0001**	0.229
CF- β	0.003	0.93	16.982 ± 1.150	17.074 ± 0.952	17.268 ± 1.260	0.227	0.060	0.016
SV-θ	-0.076	**0.048**	1.133 ± 0.244	1.196 ± 0.247	1.143 ± 0.235	**0.008**	**0.042**	0.476
SV-α	0.122	**0.002**	1.752 ± 0.480	1.553 ± 0.576	1.731 ± 0.509	**<0.001**	**0.005**	0.329
SV- β	0.068	0.079	13.755 ± 5.493	13.467 ± 5.129	14.617 ± 5.052	0.445	**0.006**	0.054
SE-θ	-0.014	0.724	0.858 ± 0.031	0.856 ± 0.038	0.859 ± 0.033	0.766	0.807	0.495
SE-α	0.115	**0.004**	0.852 ± 0.047	0.818 ± 0.071	0.850 ± 0.048	**<0.001**	**<0.001**	0.730
SE- β	0.086	0.152	0.785 ± 0.070	0.793 ± 0.063	0.797 ± 0.080	0.261	0.076	0.074

3.3.2.2 Comparison using different spectral epoch lengths in hypoglycemia detection

In real-time applications using EEG signals to detect acute and spontaneous health problems like hypoglycemia, choosing a proper epoch length for processing signals plays an important role in determining the time delay of the detecting system. It is obvious that on the one hand, a longer epoch length provides more information about responses occurring inside the signal. However, on the other hand, it reduces the size of data set available for classification algorithm, and also most importantly increases the system time delay which can be counted from the point of starting collect data to the point of yielding final system response (i.e. in the case of hypoglycemia detection, the final system response is hypoglycemia or non-hypoglycemia).

In this section, in order to select an appropriate epoch length of signals, from the filtered 40-second EEG segments, EEG parameters with 5 different epoch lengths of 1-second, 2-second, 5-second, 10-second, and 20-second will be extracted and used to compare between Hypoglycemia phase and Normal phase. Accordingly, from the raw data set acquired from 5 T1DM patients during the glucose clamp study, for each case, the number of data points extracted is 940, 470, 188, 94 and 47, respectively. To make the comparison simpler, only results of four parameters from the alpha band at two channels C3 and O2 are presented. Similarities are recognized with other frequency bands as well as EEG channels.

It is shown that there are slight differences in the results of different cases. When comparing all 4 parameters between 2 groups, smaller epoch lengths with bigger data sizes produce smaller *p*-values which indicate more significant parameters. The results show that the smallest epoch length of 1-second, with corresponding frequency resolution of 1Hz, is sufficient to identify shifts in power spectrum. It is clear that a smaller epoch length produces a bigger data set, which is important in classification. However, a too small epoch length is not necessary for some applications such as hypoglycemia detection and can lead to higher computational cost. Based on advantages as well as disadvantages of each case, we establish that using 5-second epoch length is sufficient for spectral analysis which can produce a good-sized data set for classification.

Table 3.7: Statistical results of comparing epoch lengths

		1-second			2-second			5-second		
		Normal	Hypo	*p*-value	Normal	Hypo	*p*-value	Normal	Hypo	*p*-value
C3	*P-α*	2.669±3.957	2.709±3.600	0.782	1.413±1.210	1.621±1.583	0.011	1.552 ± 1.333	1.611 ± 1.123	0.202
	CF-α	9.805±0.552	9.591±0.599	< 0.0001	10.076±0.497	9.831±0.548	< 0.0001	10.255 ± 0.460	9.993 ± 0.495	< 0.001
	SV-α	1.578±0.491	1.485±0.495	0.001	1.691±0.479	1.613±0.530	0.039	1.844 ± 0.488	1.688 ± 0.550	0.002
	SE-α	0.797 ± 0.089	0.773 ± 0.103	< 0.0001	0.824 ± 0.065	0.800 ± 0.079	< 0.0001	0.856 ± 0.044	0.830 ± 0.064	<0.0001
O2	*P-α*	1.434±2.824	2.537 ± 6.006	0.093	1.755 ± 3.147	2.940 ± 5.373	0.105	2.258 ± 2.618	3.363 ± 5.737	0.146
	CF-α	9.728 ± 0.500	9.486 ± 0.562	< 0.0001	9.963 ± 0.459	9.737 ± 0.514	< 0.0001	10.133 ± 0.392	9.863 ± 0.454	<0.0001
	SV-α	1.503 ± 0.496	1.384 ± 0.496	< 0.0001	1.614 ± 0.491	1.494 ± 0.564	< 0.0001	1.752 ± 0.480	1.553 ± 0.576	<0.001
	SE-α	0.796 ± 0.088	0.762 ± 0.105	< 0.001	0.820 ± 0.064	0.790 ± 0.084	< 0.001	0.852 ± 0.047	0.818 ± 0.071	<0.001

		10-second			20-second		
		Normal	Hypo	*p*-value	Normal	Hypo	*p*-value
C3	*P-α*	1.600±1.072	1.933 ± 1.356	0.110	1.410±0.848	1.891±1.078	0.005
	CF-α	10.338±0.376	10.010±0.459	< 0.0001	10.300±0.362	10.084±0.446	0.001
	SV-α	1.957±0.465	1.774±0.539	0.008	2.033±0.365	1.808±0.576	0.019
	SE-α	0.877± 0.036	0.856 ± 0.058	0.012	0.893 ± 0.028	0.873 ± 0.049	0.034
O2	*P-α*	1.392 ± 1.926	3.173 ± 4.578	0.164	1.228 ± 1.553	2.722 ± 3.441	0.174
	CF-α	10.208 ± 0.369	9.904 ± 0.412	<0.0001	10.228 ± 0.347	9.969 ± 0.422	0.004
	SV-α	1.809 ± 0.450	1.586 ± 0.589	0.006	1.836 ± 0.484	1.650 ± 0.618	0.110
	SE-α	0.876 ± 0.381	0.839 ± 0.063	0.013	0.891 ± 0.028	0.860 ± 0.054	0.005

3.3.2.3 Determination of a blood glucose threshold for defining hypoglycemic state in T1DM patients

The determination of a blood glucose level (BGL) at which EEG parameters start to change plays an important role in terms of identifying the hypoglycemic state from EEG signals. As presented in Chapter 2, section 2.3, the definition of hypoglycemia by specifying a blood glucose threshold is different from study to study and depends crucially on the aim and available technique of each study. If this threshold is set at a high level, there is a possibility of early detection of hypoglycemic episodes which is valuable for patients. However, the specificity of the system may become very low, making it unreliable. Conversely, if it is set at a too low level, the system may give alarm too late and the patient will not have enough time to fix the situation until it becomes severe. With the purpose of developing a system to detect hypoglycemia from EEG signals, this threshold should be determined by finding the blood glucose level at which EEG parameters start to respond to the hypoglycemic condition.

Until now in this thesis, the BGL of 3.3 mmol/l has been used as the threshold for identifying episodes of hypoglycemia. In this section, the possibility of detecting an earlier onset of hypoglycemia will be verified by exploring the responses of EEG parameters at an earlier BGL threshold of 3.9 mmol/l. The BGL range of 3.3-3.9 mmol/l is labelled as Early Onset phase of the study, as defined in section 3.2.2.1, Table 3.1. By comparing data extracted from the Early Onset phase with data extracted from the Normal phase (BGLs > 3.9 mmol/l) and Hypoglycemia phase (BGLs < 3.3 mmol/l), the difference in EEG responses at two BGL values of 3.3 mmol/l and 3.9 mmol/l will be demonstrated.

The results of group comparisons are presented in Table 3.8 and 3.9. As shown in section 3.3.2.1, there are similarities in EEG responses between EEG channels at the same brain area. Consequently, in this section, only results from 2 channels C3 and O2 which are from two different brain sides and areas are reported. Similar responses are recognized at the other two channels C4 and O1. Statistical results show no significant difference between the Normal state and Early Onset state except some slight changes at channel O2. Comparison between the Early Onset group and the Hypoglycemia group produces

significant changes which are similar to the differences between Normal and Hypoglycemia groups. These results indicate that at the BGL area of 3.3-3.9 mmol/l, there are some slight changes in EEG parameters; however, these changes are not significant until the BGL is lower than 3.3 mmol/l. As a results, it is concluded that EEG responses to the onset of hypoglycemia only significantly occur when patients' BGLs fall to the threshold of 3.3 mmol/l.

Table 3.8: EEG responses when BGL=3.3-3.9 mmol/l at channel C3

EEG parameters		**Early Onset (EO)**	**Group comparisons**	
			N-EO	**EO-H**
C3	*P-θ*	1.960± 1.464	0.197	0.434
	P-α	1.509 ± 1.062	0.702	0.170
	P- β	1.154 ± 1.399	0.075	0.114
	CF-θ	5.632 ± 0.267	0.782	0.615
	CF-α	10.247 ± 0.268	0.904	<0.0001
	CF- β	17.408 ± 2.382	0.387	0.002
	SV-θ	1.161 ± 0.215	0.307	<0.0001
	SV-α	1.750 ± 0.337	0.076	0.403
	SV- β	17.086 ± 3.227	0.125	0.320
	SE-θ	0.862 ± 0.033	0.216	0.454
	SE-α	0.868± 0.030	0.038	<0.0001
	SE- β	0.760 ± 0.115	0.049	<0.001

Table 3.9: EEG responses when BGL=3.3-3.9 mmol/l at channels O2

EEG parameters		Early Onset (EO)	Group comparisons	
			N-EO	EO-H
O2	*P-θ*	1.943 ± 0.877	0.417	0.237
	P-α	2.344 ± 0.980	0.856	0.020
	P- β	1.281 ± 1.073	0.023	0.095
	CF-θ	5.697 ± 0.261	0.067	0.012
	CF-α	10.112 ± 0.326	0.851	<0.0001
	CF- β	17.055 ± 2.046	0.029	0.764
	SV-θ	1.176 ± 0.238	0.097	0.615
	SV-α	1.781 ± 0.353	0.745	<0.001
	SV- β	13.539 ± 4.799	0.653	0.026
	SE-θ	0.858 ± 0.026	0.544	0.954
	SE-α	0.845 ± 0.026	0.068	0.024
	SE- β	0.786 ± 0.059	0.111	0.043

3.3.3 Classification performance of hypoglycemia detection

3.3.3.1 Classification results of neural networks using all four EEG channels

As presented in section 3.2.3, with the powerful ability of modelling complex non-linear relationships between inputs and outputs, in this thesis, feed-forward neural networks are selected as classification units for the purpose of detecting hypoglycemia from EEG signals. In this section, the classification performance of developed neural networks will be presented to evaluate the capability of the detection of hypoglycemic episodes from EEG signals for patients with T1DM.

The structure of the developed neural network is presented in Figure 3.6, including an input layer, a hidden layer and an output layer. The size of the input layer is the number of extracted EEG features used as inputs. The number of hidden nodes in the hidden layer is selected to be the one giving the best classification performance as the procedure presented in section 3.2.3.4. There is only one output node in the output layer which indicates the state of hypoglycemia or non-hypoglycemia. The output threshold, which is used to distinguish between hypoglycemic state and non-hypoglycemic state, is determined by plotting the ROC curve for the trained neural network.

Based on the statistical results in section 3.3.2.1, it has been established that under hypoglycemic conditions, the decrease in centroid alpha frequency and the increase in centroid theta frequency are the most significant and consistent changes in EEG signals from 5 T1DM patients. Thus, in this section, these two features at all four channels are selected as inputs for classification. As a result, a neural network will be developed with the structure of 8 input nodes (2 features x 2 channels), S hidden nodes and 1 output node. Referring to equation 3.7, the total number of parameters of the developed neural network is estimated as:

$$\text{Number of parameter} = (8+2)*S+1 = 10S+1 \tag{3.22}$$

The developed neural network is trained by the Levenberg-Marquardt algorithm (as shown in section 3.2.3.2) and the cross-validation technique (as shown in section 3.2.3.3). To do

this, the overall data set acquired from 5 T1DM patients is separated into 3 different data sets including a training set, a validation set and a testing set. The training set and validation set are formed by randomly dividing a data set from 3 patients, named patient A, patient B and patient C. The size ratio of training set to validation set is 3:1. After being trained, the neural network will be tested by a data set from two previously unseen patients, named patient D and patient E. As a result, the number of data points for each data set is presented as shown in Table 3.10.

Table 3.10: Number of data points for training and testing neural network using the LM algorithm and the cross-validation technique

		Total	Hypoglycemia	Non-hypoglycemia
Training	*data from 3 patients A, B and C, randomly divided with ratio of 3:1*	**213**	**84**	**129**
Validation		**71**	**28**	**43**
Testing	*data from 2 patients D and E*	**144**	**76**	**68**

It is noted that selecting network architecture plays an important role in developing neural network to achieve desired classification performance. As presented in section 3.2.3.4, in order to determine the size of hidden layer which can produce better performance, the number of hidden nodes S is varied from 1 to 15, corresponding to 15 different network architectures. For comparison purposes, based on the ROC curve plotted for each case, the point that produces sensitivity of 80% is selected as output threshold to distinguish between hypoglycemia and non-hypoglycemia. Comparison between performances of 15 developed neural networks is given in Table 3.11, in terms of AuC, sensitivity and specificity on the combined training/validation data set. The number of network parameters is estimated for each network by using equation 3.22. The reported results are the best results of 20 running times.

Table 3.11: Comparison between classification performance of neural network structures of 8 input nodes, *S* hidden nodes and 1 output nodes

Number of hidden nodes *S*	Number of network parameters	**Classification performance**		
		AuC	Sen	Spe
1	11	0.69	80%	39%
2	21	0.70	80%	48%
3	31	0.72	80%	45%
4	41	0.74	80%	48%
5	51	0.74	80%	49%
6	61	0.76	80%	50%
7	71	0.77	80%	49%
8	**81**	**0.78**	**80%**	**52%**
9	91	0.78	80%	51%
10	101	0.76	80%	52%
11	111	0.77	80%	50%
12	121	0.79	80%	53%
13	131	0.78	80%	54%
14	141	0.76	80%	52%
15	151	0.79	80%	54%

Consequently, with the developed neural network of 8 inputs and 1 output, when $S = 1 \div 5$, it is recognized that the *AuC* is small, corresponding with small specificity (smaller than 50%). $S = 6 \div 8$ provide better results with $AuC = 0.77 \div 0.78$, and specificity higher than 50%. The best classification results we can achieve is *AuC* of 0.78, sensitivity of 80% and specificity of 60% when $S = 8$. $S = 9 \div 15$ produces similar results with no statistical enhancement compared to $S = 8$. As a result, we choose the final neural network structure of 8 hidden nodes.

The ROC curve for the neural network with 8 hidden nodes which produces the best classification results in 20 running times is plotted in Figure 3.13. The corresponding *AuC* for the combined training/validation data set is 0.77. With this ROC curve, the threshold to distinguish between the hypoglycemia and non-hypoglycemia states is set at -0.3764, which produce training/validation classification results of 80% sensitivity and 52% specificity. To show how well these results generalise to new data, the testing data from two entirely new patients (patient D and patient E) are applied to neural network. All classification results are presented in Table 3.12.

The classification results in Table 3.10 show that by using two features of centroid theta frequency and centroid alpha frequency at all four EEG channels of C3, C4, O1, O2 as inputs, a neural network with 8 hidden nodes performs acceptably. With 80% sensitivity and 52% specificity on the combined training/validation data set, and 71% sensitivity and 50% specificity on the testing set, it is demonstrated that hypoglycemia can be detected non-invasively and efficiently from EEG signals. Comparing to the results reported by other research groups using different methodologies (Bode et al. 2004; Laione & Marques 2005), it is noted that the sensitivity of 80% on the training set is desirable in terms of hypoglycemia detection. However, the specificity of 52% can be considered as a low level, leading to a high rate of non-hypoglycemic episodes which are wrongly classified and alarmed as hypoglycemic episodes, causing a huge amount of inconvenience for patients and their caregivers, especially during the night. It is obvious that more advanced methodologies should be explored in order to enhance the overall performance of the system.

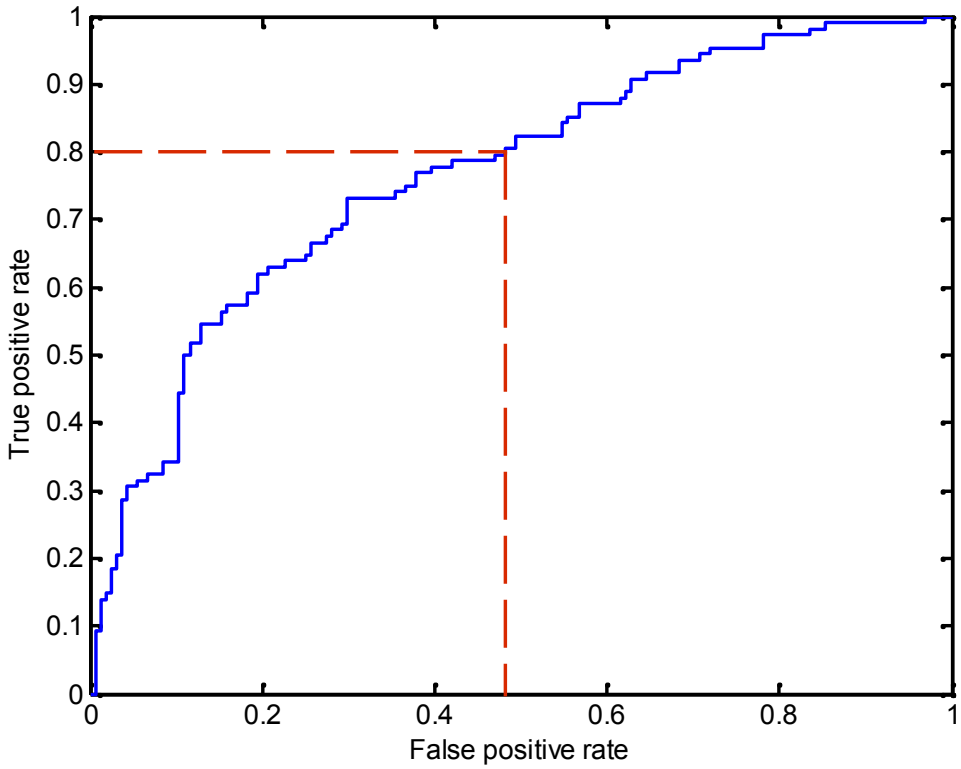


Figure 3.13: ROC curve plotted for the trained neural network using data from all 4 EEG channels

Table 3.12: Classification results of the developed neural network with 8 input nodes, 8 hidden nodes and 1 output node

	AuC	Combined training/ validation set		Testing set	
		Sen	**Spe**	**Sen**	**Spe**
Best results	0.77	80%	52%	71%	50%
Mean results	0.75	80%	50%	69%	43%

AuC: area under the ROC curve
Sen: sensitivity ; Spe: specificity

3.3.3.2 Classification results of neural networks using only two EEG channels

Considering the issue of developing a hypoglycemia detection device which can work in real-time and be applied in the real clinical environment, it is noted that reducing the number of EEG electrodes benefits the device in several ways, such as: (i) minimizing the cumbersome size of the sensor system, therefore creating more comfort for the user while wearing the device, especially during the night; (ii) decreasing the size of data needed to be transmitted and processed, leading to a smaller computational burden for the system, therefore limiting the delay in time and increasing the efficiency of the system; (iii) reducing the cost of the whole system; etc.

As presented in the previous section, when using two features of centroid theta frequency and centroid alpha frequency from all four EEG channels of C3, C4, O1, O2 as inputs to neural network (corresponding to 8 inputs), the classification results of the developed neural network demonstrate the potentiality of the proposed computational methodology for hypoglycemia detection from EEG signals. The aim of this section is to explore the possibility of reducing the number of needed EEG channels without significantly affecting the performance of the system. As shown in section 3.3.2.1, there are similarities in EEG responses between EEG channels at the same brain area. Thus, in this section, only EEG data from 2 channels C3 and O2 which are from two different sides and areas of the brain are used as inputs for classification.

For comparison purposes, the same procedure of developing and training neural networks by using the Levenberg-Marquardt algorithm and cross-validation technique as shown in section 3.3.3.1 will be applied. The same data sets as presented in Table 3.10 are used to train and test the performance of neural networks. Two features of centroid alpha frequency and centroid theta frequency are extracted from two channels C3 and O2 to be fed into neural networks as inputs. As a result, neural networks with the structure of 4 input nodes, S hidden nodes and 1 output node will be developed. The number of hidden nodes S is varied from 1 to 15, leading to 15 different network structures in order to find the hidden layer size which produces the best classification performance. Consequently, it is recognized that for this application, with 4 input nodes and 1 output node, the neural

network with 9 hidden nodes yields the best classification results. All the following reported results are corresponding with the network structure of 4 input nodes, 9 hidden nodes and 1 output node. Referring to equation 3.7, the total number of parameters of the developed neural network is equal to 55.

The ROC curve for the trained neural network with 9 hidden nodes is plotted in Figure 3.14. The corresponding *AuC* for the combined training/validation data set is 0.78. With this ROC curve, the threshold to distinguish between the hypoglycemia and non-hypoglycemia states is selected at -0.4649, which produce training/validation classification results of 80% sensitivity. To show how well these results generalise to new data, the testing data from two entirely new patients (patient D and patient E) are applied to neural network. All classification results are presented in Table 3.13, in which the reported results are the best and mean results of 20 running times.

Comparing to results of section 3.3.3.1, it is shown that using EEG data from only 2 channels of C3 and O2, the neural network structure with 4 input nodes, 9 hidden nodes and 1 output node produces acceptable classification results which are not significantly different from using EEG data from 4 channels of C3, C4, O1 and O2. The similarity has an important meaning as the number of EEG electrodes can be reduced to two, leading to a reduction of network structure from 81 parameters to 55 parameters. With the best classification results of 80% sensitivity and 53% specificity on the combined training/validation set and 71% sensitivity and 54% specificity on the testing set, it indicates that hypoglycemic episodes can be effectively detected from only two EEG channels. Once again, in order to pursue the proposed methodology to develop a hypoglycemic detection device that can be applied into the real clinical environment, the achieved results need to be enhanced further.

Table 3.13: Classification results of the developed neural network with 4 input nodes, 9 hidden nodes and 1 output node

	AuC	Combined training/ validation set		Testing set	
		Sen	**Spe**	**Sen**	**Spe**
Best results	0.75	80%	53%	71%	54%
Mean results	0.73	80%	52%	67%	42%

AuC: area under the ROC curve
Sen: sensitivity ; Spe: specificity

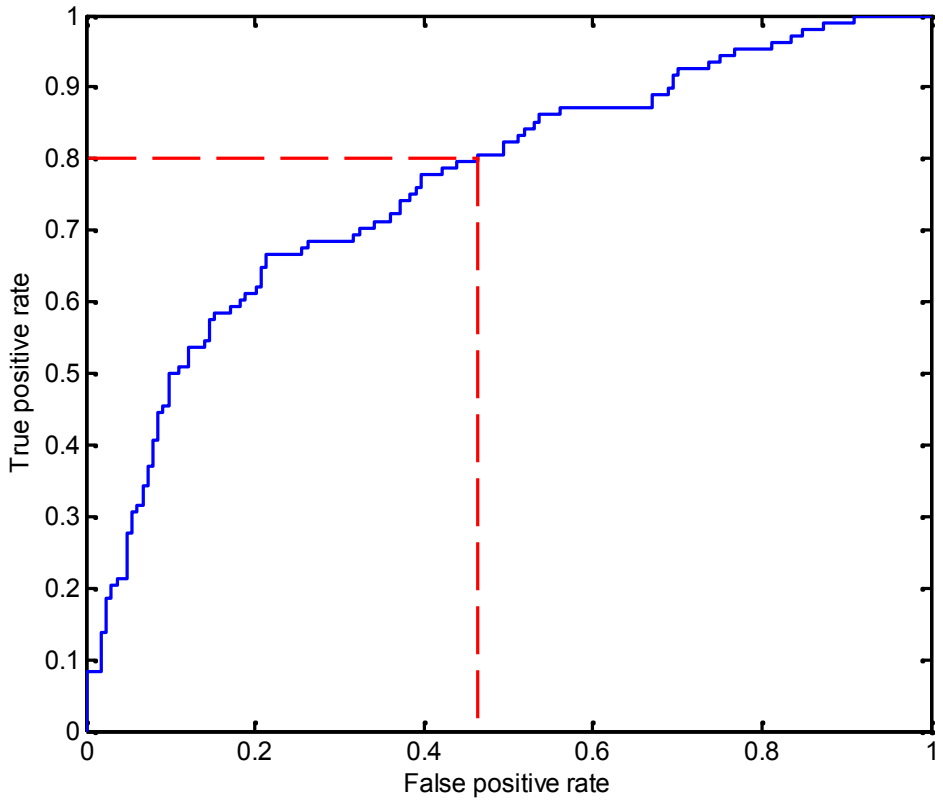


Figure 3.14: ROC curve plotted for the trained neural network using data from all 4 EEG channels plotted for the trained neural network using data from only two EEG channels of C3 and O2

The results demonstrate that there is a prominent difference between classification performances on the training data set and the unseen testing set. The mean results of 20 running times even show a clearer gap (80% sensitivity and 52% specificity on the combined training/validation data set, versus 67% sensitivity and 42% specificity on the testing set). This difference is unwanted but common in classification applications, especially when using EEG signals as input of classifiers. There are two main reasons which can be used to explain this issue:

- First, the worse results on a totally unseen testing set indicate a low ability of generalisation of the developed neural network. Because the cross-validation technique has been applied in this chapter to avoid the over-fitting situation, this problem can be caused by one of the well-known inherent limitations of Levenberg-Marquardt algorithm, which is trapping into local optimal. When the training process is trapped into one of the local optimum which is not the global one, the neural network can perform well on the training data set, but cannot generalise to another unseen data set.

- Second, when using EEG signals as inputs for any system, the difference between classification results on two different groups of subjects are predictable due to the high variability of EEG signals from person to person. This inherent characteristic of EEG has been mentioned in various EEG-related studies as a barrier to achieve their desired performance.

Because the ability to generalise to new users is essential for any health-care system, the gap between classification results on the training set and testing set needs to be overcome. As a result, the later chapters of this thesis will be dedicated to explore more advanced solutions to consecutively deal with those two main factors specified above which cause the difference in classification results on different groups of data.

3.4 Discussion

In this chapter, a computational methodology of detecting hypoglycemic episodes from EEG signals is presented, including two main tasks:

- Feature extraction: deriving and analysing various EEG parameters in order to find important features which significantly correlate to the transitions of patients' states during the study, from normal to hypoglycemic and then to recovery state.

- Classification: using extracted features as inputs, a classification algorithm based on standard neural network is developed for detecting hypoglycemic episodes.

Using data of 5 T1DM patients, collected during an overnight insulin-induced study, the results presented in this chapter show that by applying the proposed methodology, the occurrence of hypoglycemic episodes in patients with T1DM can be detected efficiently from EEG signals.

The feature extraction results indicate that under hypoglycemia conditions, there are significant changes in the theta and alpha bands of EEG signals. At all four EEG channels, the decrease in centroid alpha frequency is shown to be the most significant feature. This change can be related to the decrease of patients' vigilance which is a normal symptom for T1DM patients when hypoglycemia occurs. Under recovery conditions, these changes are shown to regain the normal state prior to hypoglycemia. However, this re-establishment does not happen consistently with all features in all patients. These results lead to a conclusion that hypoglycemia has the potential to make irrecoverable damage to the human brain. By analysing the data from the BGL range of 3.3-3.9 mmol/l, it is concluded that the mentioned responses to hypoglycemia only significantly occur when patients' BGLs fall to the threshold of 3.3 mmol/l.

Based on the results of feature extraction, centroid alpha frequency and centroid theta frequency at four channels are established to be used as inputs for classification. A standard neural network algorithm is developed. The Levenberg-Marquardt algorithm and cross-validation technique are applied for training the neural network. With the classification

results of 80% sensitivity and 52% specificity on the combined training/validation dataset, and 71% sensitivity and 50% specificity on the testing set, it is shown that using neural network is an effective method for classifying and detecting hypoglycemic episodes from EEG signals.

The classification results also indicate the efficiency of using two features of centroid theta frequency and centroid alpha frequency at four EEG channels of C3, C4, O1, O2 as inputs of neural network. The use of four EEG channels in diagnosing and detecting human health problems is considered to be reasonable. However, regarding the purpose of developing a real-time system for detecting hypoglycemia from EEG signals, a reduction in the number of EEG channels or EEG electrodes will lead to a variety of benefits. To do this, data from only two channels of C3 and O2 are explored to be fed into neural networks, leading to network structures of 4 input nodes (2 features x 2 channels). With the classification results of 80% sensitivity and 53% specificity on the combined training/validation dataset, and 71% sensitivity and 54% specificity on the testing set, it is shown that there is no significant difference between two cases of using two EEG channels and four EEG channels. As a result, it is concluded that by using two EEG channels of C3 and O2 as inputs, it is capable of providing desired performances for the application of detecting hypoglycemic episodes in patients with T1DM.

It is also noted that classification results on the testing set from totally unseen patients are poorer compared to the training set. This issue can be explained by two main reasons: (i) the poor generalisation of neural networks caused by the common limitation of trapping into local optimal of the Levenberg-Marquardt algorithm; (ii) the high variability of EEG signals from patient to patient which is an inherent characteristic of EEG-based applications. In order to enhance the overall classification performance, based on the developed standard neural network, more advanced strategy of training will be explored in the next chapters to deal with these two factors.

Chapter 4

Combining Genetic Algorithm and Levenberg-Marquardt Algorithm in Training Neural Network for Hypoglycemia Detection using EEG Signals

---

4.1 Introduction

As specified in previous chapters, the computational methodology of detecting hypoglycemia from EEG signals contains two main parts of EEG feature extraction and classification in which the classification plays an important role in determining the performance of the developed methodology. In terms of selecting classification algorithm for hypoglycemia detection, artificial neural networks have been employed popularly in biomedical areas as a powerful tool (Nguyen 2008). It has been recognized that neural networks can successfully classify complex situations and effectively model non-linear relationships between inputs and outputs. One of the most popular training techniques is the Levenberg-Marquardt (LM) algorithm which is based on the second-order gradient information of an error function in order to direct the training process to a local optimal (Hagan & Menhaj 1994). Genetic algorithm (GA) is a derivative-free global search

optimisation technique which is inspired by the natural evolution. This technique has been applied widely in evolving neural network models which can efficiently drive the training process to the global optimal (Montana & Davis 1989).

In Chapter 3, EEG signals of five T1DM patients from a glucose clamp study were acquired and analysed to find important spectral features to be used as inputs for a neural network - based classification unit. Being trained by the LM algorithm, the developed neural network produced acceptable classification results which indicate the possibility of classifying and detecting episodes of hypoglycemia from EEG signals. However, more advanced classification strategies need to be explored in order to improve the overall performance of the whole system.

This chapter aims to investigate different training algorithms for neural network in order to enhance the performance of the developed neural network-based classification unit. To do this, first, the GA algorithm is explored to assess its ability for directing the training process to the global solution. After that, a combination of GA and LM is proposed to utilize advantages as well as avoid limitations of each algorithm in training neural network. The GA algorithm will be used to locate the region of the global optimal consistently. The LM algorithm acts as a fine tuner to help the training process quickly converge toward the global solution. The main objective of this chapter is to demonstrate that by applying a properly combined strategy to train neural network, the performance of hypoglycemia detection using only two EEG channels can be improved markedly.

4.2 General description of genetic algorithm

Genetic algorithm (GA) is an evolutionary computational technique that has been widely applied to deal with complex optimisation problems where the number of variables is large and the analytical solutions are difficult to obtain. Utilizing operations inspired by the biological process of natural evolution where selection, mutation and crossover play a major role, GAs have been recognized to be a powerful technique for searching the globally

optimal solution over a domain. It has been applied in various areas such as fuzzy control (Belarbi & Titel 2000), modelling and classification (Setnes & Roubos 2000) which includes evolving neural network (Jagielska, Matthews & Whitfort 1999; Leung et al. 2003), etc.

4.2.1 Genetic algorithm

In terms of optimisation, genetic algorithms are gradient-free techniques that search for a global optimum over a population of possible solutions by evaluating a performance criterion (named fitness function). To do this, GAs simulate the process of natural evolution through coding and special operators. In genetic algorithm, the space of possible solutions of the problem first will be coded into a population of individuals or chromosomes. Each chromosome is composed of a set of genes in which each gene represents one variable of the problem.

A variety of coding methods have been introduced for GAs. Traditional binary coding is a common method which is very simple to understand and implement. However, it is not efficient when applied to multi-dimensional, high-precision or continuous problems where the bit strings can get very long, the search space may be enormous, and the algorithm performance will become very poor. Drawbacks of the binary coded GAs are overcome by introducing the floating-point coding method. In this situation, each chromosome is coded as a vector of real numbers which has the same length as the solution vector. In this way, a large searching domain can be handled more effectively. In this thesis, the floating-point coding method will be selected for implementing GAs.

The procedure of GA is presented in Figure 4.1. First a fitness function and termination conditions will be defined for the optimisation problem based on the requirement of each application. A population of chromosomes P is then initialised as follows:

$$P = [\mathbf{p}_1\ \mathbf{p}_2 \ldots \mathbf{p}_i \ldots \mathbf{p}_{n_p}] \tag{4.1}$$

where

- n_p denotes the number of chromosomes in the population or the population size. The population size is a user-defined parameter which affects the performance of GAs. Increasing n_p will increase the diversity of the search space and reduce the probability that GAs prematurely converge to a local optimum. However, it also increases the evolutionary time to converge to the optimal region which is an inherent limitation of GAs.

- $\mathbf{p}_i = [p_{i1}\ p_{i2} \cdots p_{ij} \cdots p_{in_g}]$, $j = 1 \div n_g$ where n_g denotes the number of genes in the chromosome. Each gene is corresponding to one variable of the optimisation problem. Therefore n_g is determined by the number of variables to be tuned during the evolutionary process.

- Each gene p_{ij} has a constrained range of $p_{j\min} \le p_{ij} \le p_{j\max}$

At each iteration (or generation) of the GA evolving procedure, each chromosome in the population will be evaluated by the defined fitness function:

$$fitness = f(\mathbf{p}_i) \tag{4.2}$$

The better chromosomes will produce higher fitness values. The population is then updated through a process of selection, crossover and mutation. By applying genetic operators to the population representing the current generation, a new generation will be created with a goal of improving the fitness. The selection chooses some chromosomes out of the population for reproduction based on fitness values of each chromosome in the population. The selected chromosomes undergo two genetic operators of crossover and mutation which are governed by the probabilities of crossover (μ_c) and mutation (μ_m). Basically, the crossover operation helps exchange information while the mutation operation helps alter characteristics of two selected chromosomes to generate offspring. The offspring after that are evaluated by the fitness function to replace the worst chromosomes in the old

population. As a result, after an iteration of selection, crossover and mutation, a new generation is generated and ready for the next iteration. This process is repeated until one of terminating conditions is met or when a predefined number of iteration is reached. Whatever the stopping criterion, the chromosome with highest fitness value at the last generation is taken as the solution to the problem.

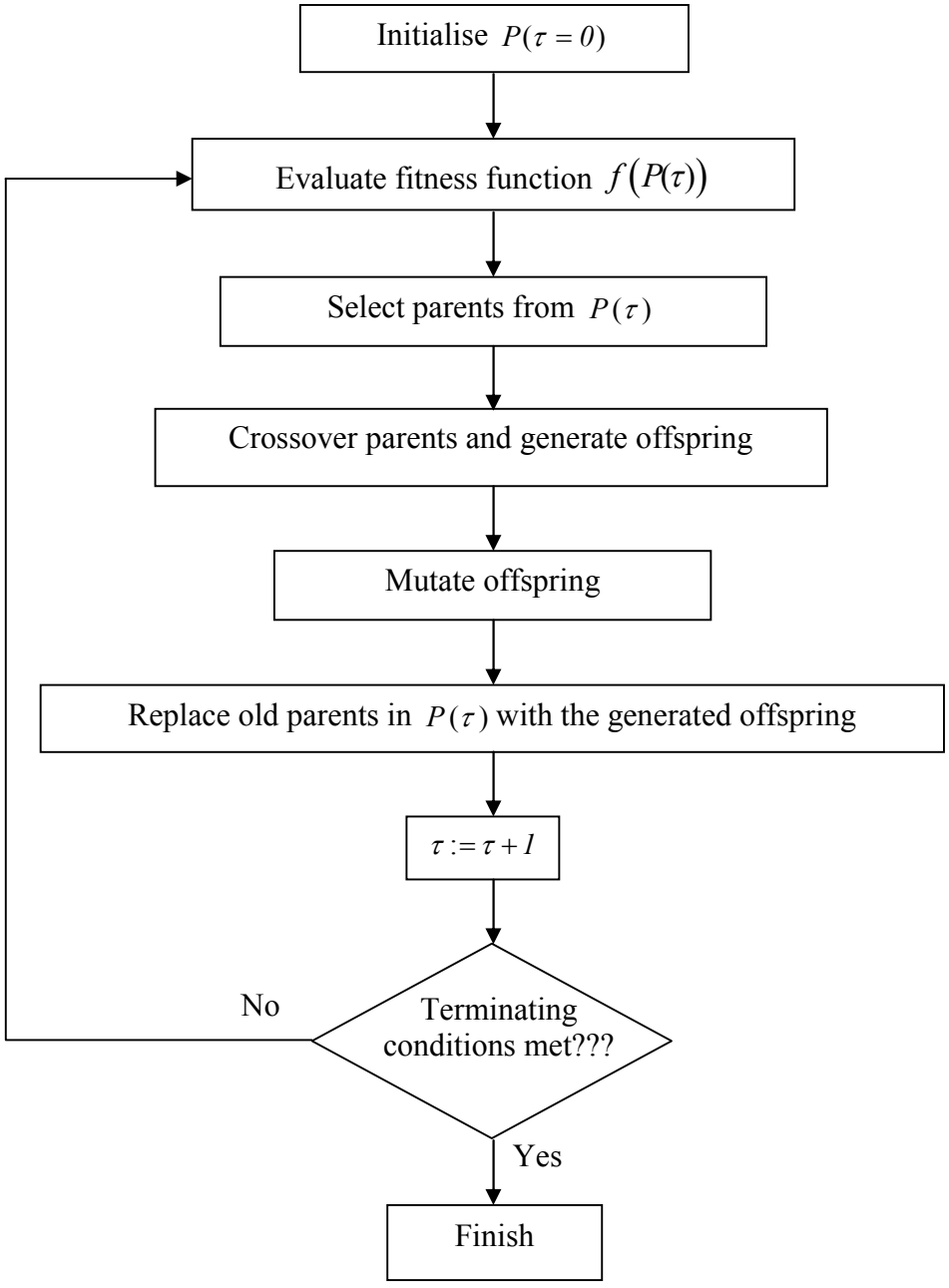


Figure 4.1: Procedure of implementing GA

4.2.2 Genetic operators

The facts of being able to explore the search space in parallel without requiring the optimised function to be differentiable or have any smooth properties make GA feasible for a wide range of optimisation problems. Clearly, the precision of the solution obtained depends on a variety of factors including chromosome representation, population initialisation, genetic operations, criteria for termination, and fitness function for evaluation. In this thesis, genetic operations are referred to three different operations including selection, crossover and mutation.

4.2.2.1 Selection operator

Selection is the process of randomly choosing two chromosomes out of the population for reproduction according to their fitness values. This operation attempts to apply pressure on the population in a similar manner to that of natural selection found in biological systems. The higher the fitness value, the more chance a chromosome will be selected, while poorer performing individuals will be rejected. There are several schemes for implementing the selection process. In a common approach, a probability of selection q_i will be assigned to each chromosome $\mathbf{p}_i$ based on its fitness value. As a result, a set of n_p probability values is produced for the population. The method of assigning probabilities to chromosomes and the rule of selecting two chromosomes out of the population depend on different techniques (e.g. Roulette wheel selection, tournament selection, ranking selection, etc.).

In this thesis, the normalized geometric ranking selection is applied as the selection operator. This is a ranking selection process based on a non-stationary penalty function which is a function of the generation number. As the number of generations increases, the penalty increases, leading to higher selective pressure to the GA to find the feasible solution. In general, a chromosome which has higher rank will have a higher chance to be selected. This method only requires the evaluation function to map the solutions to a partially ordered set, thus allowing for minimization and negativity. Unlike other methods

which are based on the real fitness values to assign the selection probability q_i to each chromosome, ranking methods in general assign q_i based on the rank of each solution when all solutions are sorted. The normalized geometric ranking selection defines the selection probability q_i to each chromosome by:

$$q_i = q'(1-q_{best})^{r-1} \tag{4.3}$$

where:

- $i = 1, 2, \ldots, n_p$ (n_p is the number of chromosomes in the population)
- q_{best} is the probability if selecting the best chromosome
- r is the rank of the chromosome of which the best has $r = 1$
- $q' = \dfrac{q_{best}}{1-(1-q_{best})^{n_p}}$ (4.4)

4.2.2.2 Crossover operator

The aim of implementing crossover is to share information between two parents (chromosomes) obtained by the selection process. This process combines the features of the two parent chromosomes to form two offspring, with the possibility that good chromosomes may generate better ones. The crossover operation is governed by a pre-defined probability of crossover μ_c which gives an expected number of chromosomes in the population that undergo the crossover ($=\mu_c.n_p$). Various methods have been introduced to implement crossover for floating-number genetic algorithm. In this thesis, Blend-α crossover is applied as the crossover operator (Eshelman & Schaffer 1993).

Let assume that $\mathbf{p}_1^\tau = [p_{11}^\tau \; p_{12}^\tau \; ... \; p_{1j}^\tau \; ... \; p_{1n_g}^\tau]$ and $\mathbf{p}_2^\tau = [p_{21}^\tau \; p_{22}^\tau \; ... \; p_{2j}^\tau \; ... \; p_{2n_g}^\tau]$ are two chromosomes that have been selected at the generation τ to be parents for reproduction. By applying the Blend- α crossover, the procedure to produce two offspring $\mathbf{o}_1^{\tau+1}$ and $\mathbf{o}_2^{\tau+1}$ are shown in Figure 4.2, where α is a positive constant.

For $i = 1 \div n_g$ **do**

Estimating:

$$d_i = \left| p_{1i}^\tau - p_{2i}^\tau \right|$$

$$X_i^1 = \min\left(p_{1i}^\tau, p_{2i}^\tau \right) - \alpha d_i$$

$$X_i^2 = \max\left(p_{1i}^\tau, p_{2i}^\tau \right) + \alpha d_i$$

Choosing uniform random real numbers u and v from the interval $\left[X_i^1, X_i^2 \right]$

$\mathbf{o}_1^{\tau+1} = u$;

$\mathbf{o}_2^{\tau+1} = v$

end

Figure 4.2: Procedure of implementing Blend-α crossover

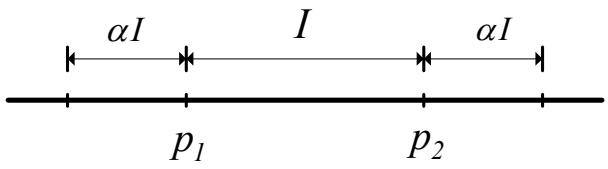


Figure 4.3: The Blend-α crossover

In the procedure of implementing Blend-α crossover, selecting α is one of the factors that determine the balance between exploration (finding completely new solution) and preservation (inheriting parents' features). When $\alpha = 0$, the Blend-α crossover is equivalent to the flat crossover which generates an offspring by uniformly picking the values of genes between (inclusively) the two parents' genes values. By this way, the searching process tends to be biased towards the center of the search space and can lead to an overemphasis on preservation. Eshelman & Schaffer (1993) showed that $\alpha = 0.5$ is the optimal choice in which the Blend-α crossover picks the offspring's genes values from points that lie on an interval which extends $0.5I$ on both sides of the interval I between the parents (as shown in Figure 4.3)

4.2.2.3 Mutation operator

Mutation is a genetic operator which alters the genes of chromosomes in the population with the aim that the features inherited from parents can be changed in the new generation. By doing this, the mutation help the searching process to escape from local minima's traps and also maintain the diversity of the population. Unlike the crossover which is aimed to explore new regions in the searching space, the mutation is supposed to exploit the already sampled regions. The level of exploitation is governed by a pre-defined probability of mutation μ_m which govern the number of genes that undergo the mutation ($=\mu_m.n_g.n_p$). Increasing μ_m tends to turn the searching process into random search so that when $\mu_m = 1,$ all genes will mutate.

There are different forms of mutation which have been introduced for the different kinds of genetic coding. In this thesis, for floating-number representation, non-uniform mutation will be investigated. With the fine-tuning capability, this mutation operator reduces limitations of applying random mutation (or uniform mutation) in floating-number GAs. Unlike the uniform mutation operator in which, at any generation, one gene is randomly selected and set at a random value between its upper and lower bounds, the non-uniform

mutation procedure depends on the generation number of the population.

Let's assume that at the generation τ, the gene p_{ij}^{τ} of the chromosome $\mathbf{p}_i^{\tau} = [p_{i1}^{\tau}\ p_{i2}^{\tau}\ ...\ p_{ij}^{\tau}\ ...\ p_{in_g}^{\tau}]$ will undergo the mutation operation. The resulting gene $o_{ij}^{\tau+1}$ of the offspring $\mathbf{o}_i^{\tau+1} = [o_{i1}^{\tau+1}\ o_{i2}^{\tau+1}\ ...\ o_{ij}^{\tau+1}\ ...\ o_{in_g}^{\tau+1}]$ is given by:

$$o_{ij}^{\tau+1} = \begin{cases} p_{ij}^{\tau} + \Delta\left(\tau, p_{j\max} - p_{ij}^{\tau}\right) & \text{if } r_d = 0 \\ p_{ij}^{\tau} + \Delta\left(\tau, p_{ij}^{\tau} - p_{j\min}\right) & \text{if } r_d = 1 \end{cases}, \tag{4.5}$$

where

- r_d is a random number equal to 0 or 1 only
- $p_{j\min}$ and $p_{j\max}$ are the lower and upper bounds of the gene or variable p_{ij}
- τ is the present generation of the population.
- The function $\Delta(\tau, y)$ returns a value in the range $[0, y]$ such that $\Delta(\tau, y)$ approaches to zero when τ increases. Because of this property, this function is able to search the space uniformly initially (when τ is small), and very locally at later stages. This strategy helps to increase the probability of getting closer to the convergent point than a random choice. The function $\Delta(\tau, y)$ is defined as follows.

$$\Delta(\tau, y) = y\left(1 - r^{\left(1 - \frac{\tau}{T}\right)^{\zeta_{num}}}\right) \tag{4.6}$$

where r is a uniform random number from [0, 1], T is the maximum generation number of the evolutionary process, and ζ_{num} is a system parameter that determines the degree of non-uniformity (or the degree of dependency on the iteration number of the mutation operator).

4.3 Genetic algorithm-based neural networks for hypoglycemia detection from EEG signals

4.3.1 Genetic algorithm-based neural networks

With the ability of globally searching through complex, multimodal solution spaces without requiring the fitness function to be differentiable, genetic algorithms have been widely applied in developing neural networks. Different approaches to evolve neural networks based on genetic algorithms have been introduced, including tuning network parameters, finding optimal network structure (number of hidden layers, number of hidden nodes in each layer), selecting input features, initialising network parameters, selecting learning rules, etc. Depending on unique attributes of each application, hybrid approaches are also explored to optimise both structure and parameters of neural networks, or combined sets of initial network parameters, number of hidden nodes, learning rules, etc.

With the final aim of the project which is developing a real-time hypoglycemia detecting system that is applicable in real-life use, the structure of neural network is expected to be minimally reduced. A small set of 4 features from 2 EEG channels has been extracted to feed into neural network. With the input layer of 4 input nodes and output layer of 1 input node, it has been demonstrated in Chapter 3 and 4 that the acceptable classification performance can be achieved with the number of hidden nodes in the hidden layers in the range of 1-16. Therefore, optimizing this factor is not considered in this thesis, as well as other factors of selecting input features or learning rules.

In this chapter, genetic algorithm is only explored as a training technique to optimise network parameters. Based on the capability of globally optimisation, a genetic algorithm is developed and applied to neural network with the aim of searching over the whole domain and directing the training process to the global optimal region. The main features of this process are presented as follows.

- ***Population initialising***

First, a population of chromosomes or individuals $P = [\mathbf{p}_1\ \mathbf{p}_2\ \ldots\ \ldots\ \mathbf{p}_{n_p}]$ will be generated in which each chromosome $\mathbf{p}_i = [p_{i1}\ p_{i2}\ \ldots\ p_{ij}\ \ldots\ p_{in_g}]$, $j = 1 \div n_g$ is a solution of the optimisation problem. With the aim of training neural network to obtain optimal network parameter, each chromosome in the population will be represented by the set of parameters of neural network.

Based on a given structure of a three-layer neural network which consists of N_{in} input nodes, N_{hid} hidden nodes and N_{out} output node (as shown in Figure 3.5), each chromosome in the population is expressed as follows:

$$[w_{ij}\ v_{ki}\ b_{1i}\ b_{2k}] \tag{4.7}$$

where

- $w_{ij}, i = 1 \div N_{hid}, j = 1 \div N_{in}$ is the weight of the link between i-th hidden node and the j-th input.
- $w_{ki}, i = 1 \div N_{hid}, k = 1 \div N_{out}$ is the weight of the link between i-th hidden node and k-th output
- b_{1i}, b_{2k} are the biases for the hidden nodes and output nodes respectively.

Each gene of the chromosome is one neural network parameter. As a result, the length of chromosome n_g, which is equal to the total number of neural network parameters, is calculated as follows.

$$\begin{aligned} n_g &= (N_{in} + 1) * N_{hid} + (N_{hid} + 1) * N_{out} \\ &= (N_{in} + N_{out} + 1) * N_{hid} + N_{out} \end{aligned} \tag{4.8}$$

- ***Fitness function***

To learn the input-output relationship, a fitness function needs to be defined and used during the evolution to estimate and compare the performance of chromosomes in the population. In this thesis, the function used to estimate the fitness of each chromosome or each solution is defined as follows:

$$f(chromosome) = \frac{1}{1 + E(\mathbf{w})} \tag{4.9}$$

$E(\mathbf{w})$ is the mean squared error function defined as:

$$E(\mathbf{w}) = \frac{1}{N * N_{out}} \sum_{n=1}^{N} \sum_{k=1}^{N_{out}} \{y_k(x^n, \mathbf{w}) - t_k{}^n\}^2 \tag{4.10}$$

where:

- $\mathbf{w}=[w_{ij}\ v_{ki}\ b_{1i}\ b_{2k}]$ is the vector of network parameters or the current chromosome

- $y_k(x^n, \mathbf{w})$ and t_k^n are the real k-th output and its corresponding target when feeding the *input* x^n into network; $k = 1 \div N_{out}$ where N_{out} is the number of output nodes in the output layer ; $n = 1 \div N$ where N is the number of data point of the training set.

- ***Genetic operation***

At each iteration (or generation) of the training process, the population is updated through a process of selection, crossover and mutation. The selection chooses some chromosomes out of the population for reproduction based on fitness values of each chromosome in the population. The selected chromosomes undergo two genetic operators of crossover and

mutation to generate offspring. The offspring after that are evaluated by the fitness function to replace the worst chromosomes in the old population to generate a new population.

- ***Terminating conditions***

The evolutionary process, including selection, crossover and mutation will be repeated until one of terminating conditions is met. The best chromosome in the final population will be considered as the final solution of the optimisation problem, which is the final set of neural network parameters.

4.3.2 Results of hypoglycemia detection using GA-based neural network

As results of the feature extraction part presented in Chapter 3, for hypoglycemia detection purpose, a feed-forward three-layer neural network will be developed. The structure of the developed neural network is as follows:

- The input layer contains 4 input nodes which are corresponding with extracted features from EEG signals (2 features of centroid theta frequency and centroid alpha frequency x 2 channels of C3 and O2).

- The hidden layer contains S hidden nodes. The number of hidden nodes S is varied from 1 to 16 to select the one that gives the best performance. As a result, it is recognized that for our application with 4 input nodes and 1 output node, $S = 9$ yields the best classification results. The following results are corresponding with a neural network of 9 hidden nodes.

- The output layer contains only 1 output node which indicates the state of hypoglycemia or non-hypoglycemia.

Consequently, referring to equation 4.8, the total number of parameters of the developed neural network is estimated as:

$$\text{Number of parameters} = (N_{in} + N_{out} + 1) * N_{hid} + N_{out}$$
$$= (4 + 1 + 1) * 9 + 1 = 55$$

For training the neural network by using the GA algorithm, the overall data set is separated into a training set and a testing set. The training set is used to evolve neural network's parameters, while the testing set is used to verify the performance of the trained neural network. The number of data points in each set is presented in Table 4.1. The training set is formed from the data of 3 patients, who have been called patient A, B and C. This set consists of 284 data points which includes 112 points of hypoglycemia. The testing set is formed from data of 2 patients, who have been called patient D and E. This set consists of 144 data points which include 76 points of hypoglycemia.

Table 4.1: Number of data points for training and testing neural network using GA

	Total	Hypoglycemia	Non_hypoglycemia
Training (*data from patient A, B and C*)	**284**	**112**	**172**
Testing (*data from patient D and E*)	**144**	**76**	**68**

A population of chromosome $P = [\mathbf{p}_1\ \mathbf{p}_2\ \ldots\ \ldots\ \mathbf{p}_{n_p}]$ with the population size n_p of 50 is initialised. Each chromosome $\mathbf{p}_i = [p_{i1}\ p_{i2}\ \ldots\ p_{ij}\ \ldots\ p_{in_g}]$, $j = 1 \div n_g$ is equivalent to a vector of network parameters as shown in equation 4.7, where each gene of the chromosome is equivalent to a network parameter. With the developed neural network of 4 input nodes, 9 hidden nodes and 1 output node, the chromosome length equals 55 which is the number of parameters needed to evolve during the training process. The upper and lower bounds for each gene or each network parameter are set at -3 and 3 respectively. The initial population is generated uniformly at random.

In this thesis, for the evolutionary process, three genetic operators are selected, including normalized geometric ranking selection, Blend-α crossover, non-uniform mutation. Details about the above operators are presented in section 4.2.2.1 – 4.2.2.3. For normalized geometric ranking selection, the probability of selecting the best chromosome q_{best} is set at 0.08. The probability of crossover μ_c and probability of mutation μ_m are set at 0.8 and 0.5 respectively.

The fitness function used to evaluate the neural network performance during the evolutionary process is presented as in equations 4.9 and 4.10, where the number of output nodes in the output layer $N_{out} = 1$ and the number of data point of the training set $N = 284$ as shown in Table 4.1.

To stop the training process, two terminating conditions are established. One is when the fitness function approaches a targeted value which corresponds to the point producing the mean squared error $E(\mathbf{w}) = 10^{-2}$. If the training process cannot get to that targeted value, the evolutionary process will be terminated at the maximum number of generations which is set at 2000 in this thesis. This second terminating condition is necessary because it helps to prevent the training process from one of the inherent characteristic of GAs which is slow convergence. The evolutionary process will be stopped whenever one of the terminating conditions is fulfilled. The best chromosome producing the highest value of fitness function in the final population is considered as the optimal set of parameters of the neural network.

Table 4.2: Summary of GA setup for training neural network

Parameters	
Population size	50
Gene representation	$[w_{ij}\ v_{ki}\ b_{1i}\ b_{2k}]$
Chromosome length	55
Parameter range	$-3 \le p_{ij} \le 3$
Selection method	Normalized geometric ranking selection
Probability of selecting the best chromosome q_{best}	0.08
Crossover operator	Blend-α crossover
Probability of crossover μ_c	0.8
Mutation operator	Non-uniform mutation
Probability of mutation μ_m	0.5
Terminating criteria	$E(\mathbf{w}) = 10^{-2}$ Maximum number of generations = 2000

A summary of GA's parameters set up in this thesis to implement training neural network is presented in Table 4.2. After training the neural network, a Receiver Operating Characteristic (ROC) curve is plotted for the training set. Details about ROC curve are provided in section 3.2.3.3. The area under the curve (*AuC*) is estimated to evaluate the classification performance of the developed neural network. Based on the ROC curve, the threshold of neural network's output to distinguish between hypoglycemia and non-hypoglycemia is set at the point producing classification sensitivity of 80%. The testing set is applied to test the classification performance of GA-trained neural network.

A comparison between classification results of neural networks trained by the GA algorithm and the LM algorithm is provided in Table 4.3. In this table, the presented classification results are the mean and best performance of 20 running times. Details of achieving classification results by using LM-based neural network are provided in Chapter 3, section 3.3.3.

Table 4.3: Comparison of classification results between GA and LM algorithms

	Training method	***AuC***	Training set		Testing set	
			Sen	**Spe**	**Sen**	**Spe**
Mean results	GA	0.70	80%	37%	83%	40%
	LM	0.73	80%	52%	67%	42%
Best results	GA	0.74	80%	42%	73%	41%
	LM	0.75	80%	53%	71%	54%

AuC: area under the ROC curve
Sen: sensitivity ; Spe: specificity

As shown in Table 4.3, the best classification results of 20 running times indicate that the neural network trained by the LM algorithm produces better performance than the GA algorithm. Based on the ROC curve which corresponds to each algorithm, the point that produces training sensitivity of 80% is selected as the output threshold for each case. In this way, with the same acceptable sensitivity of 80% for hypoglycemia detection, it is obvious that LM algorithm produces a more desired specificity of 53% compared to a specificity of 42% produced by GA algorithm.

However, when looking at the mean classification results of 20 running times, the positives as well as drawbacks of each algorithm are demonstrated. Even though the LM algorithm produces better mean classification performance on the training set (*AuC* of 0.73, specificity of 52%) compared to the GA algorithm (*AuC* of 0.70, specificity of 37%), the performance on the testing set produced by the LM algorithm (sensitivity of 67%, specificity of 42%) is significantly poorer than the GA algorithm (sensitivity of 83%, specificity of 40%).

The poor results on the testing set yielded by the LM algorithm can be explained by its inherent characteristic of premature convergence. Due to this feature, the LM algorithm is not effective in directing the training process to the global solution but is likely to be trapped in a local optimum. On the other hand, the GA algorithm produces consistent results in all running times which can be explained by its ability to direct the optimisation process to the region of optimal solution. However, the mean classification results of 80% sensitivity and 42% specificity yielded by the GA algorithm indicates a prominent limitation of GAs which is inefficient in fine tuning to approach to the local optimal. As a result, even though the GA algorithm can lead the training process to the global solution, the classification results attained are quite low.

4.4 Combination of genetic algorithm and Levenberg-Marquardt algorithm in training neural network for hypoglycemia detection from EEG signals

4.4.1 Procedure of the GA+LM algorithm for training neural network

The comparison between classification results of the neural networks trained by genetic algorithm (GA) and Levenberg-Marquardt (LM) algorithm separately shown in section 4.3.2, demonstrates benefits as well as limitations of each algorithm. In this section, in order to take advantage of positives and avoid negatives of each algorithm, a combination of GA and LM (named the GA+LM algorithm) is explored for training neural network.

To do this, the neural network will be trained by a procedure of two consecutive steps of ***global search*** and ***local search***. Details about implementing the GA+LM algorithm for neural network training are presented as follows:

- ***Global search:*** First GA is employed to evolve neural network's parameters in order to direct the training process to the region of the global solution. The implementation of training neural network by using GA is presented in details in section 4.3.1. The evolutionary process including selection, crossover and mutation is applied on a population of network parameters with the aim of optimizing the fitness function defined in equations 4.9 and 4.10. Whenever one of the terminating conditions is met, the evolutionary process is stopped. The best chromosome with the highest value of fitness function in the last updated population is considered as the final solution of the GA algorithm or the final set of network parameters produced by the ***Global search*** step.

- ***Local search:*** In order to overcome GA's drawbacks of inefficient fine tuning and slow convergent rate, after the ***global search*** by GA, a step of ***local search*** is

implemented. To do this, first, the final set of network parameters obtained by the GA algorithm will be set as initial parameters for the neural network. The LM algorithm is then employed on this parameter set to continue training the neural network. In this way, the LM algorithm acts as a fine tuner to help the training process quickly converge toward the global solution. Details about implementing the LM algorithm for training neural network are provided in Chapter 3, section 3.3.3. In brief, the LM algorithm estimates the second directional derivative of the performance function (the mean squared error function as shown in equation 4.10), in order to ensure the training process direct to a local optimal. To avoid overtraining which may cause bad generalisation, the cost function on a separate validation set is also monitored during the training process. When the validation error keeps increasing for a given number of iterations, the training is stopped. The parameters at that stopped iteration will be considered as the final neural network weights and biases.

4.4.2 Results of hypoglycemia detection using GA+LM-based neural network

As presented in section 4.3.2, a neural network with the structure of 4 input nodes, 9 hidden nodes and 1 output node is developed for the aim of classifying and detecting episodes of hypoglycemia from EEG signals. This neural network will be trained by a 2-step procedure of global search and local search. The global search is carried out using the GA algorithm while the local search is carried out using the LM algorithm. The number of data points used for each step is arranged as follows:

- For training the neural network by using the GA algorithm, the overall data set of 5 participated patients is separated into a training set formed from the data of 3 patients, called patient A, B and C and a testing set formed from the data of 2 patients, called patient D and E. The number of data points in each set is presented as in Table 4.1.

- For training the neural network by using the LM algorithm, in order to implement

early stopping, the above training set used for GA training is randomly subdivided into an LM-training set and a LM-validation set with a ratio of 3:1. The same testing set from 2 patients D and E will be used to verify the performance of the GA+LM-based neural network. The number of data points in each data set for LM training is presented in Chapter 3, Table 3.10.

After training the neural network by two steps of global search and local search as presented in section 4.4.1, the ROC curve will be plotted for the training set (the combined dataset of LM-training set and validation set) to estimate the *AuC* as well as to determine output threshold to distinguish between hypoglycemia and non-hypoglycemia.

Classification results are presented in Table 4.5. The reported results are the mean and best performance of 20 run times. The ROC curve for the case producing the best performance is plotted in Figure 4.3. For comparison, the ROC curve for the best performance produced by GA-based training is also plotted in this figure.

Table 4.4: Classification results of GA+LM-based neural network

	AuC	Training set		Testing set	
		Sen	**Spe**	**Sen**	**Spe**
Mean results	0.79	80%	57%	74%	52%
Best results	0.82	80%	61%	75%	60%

AuC: area under the ROC curve
Sen: sensitivity ; Spe: specificity

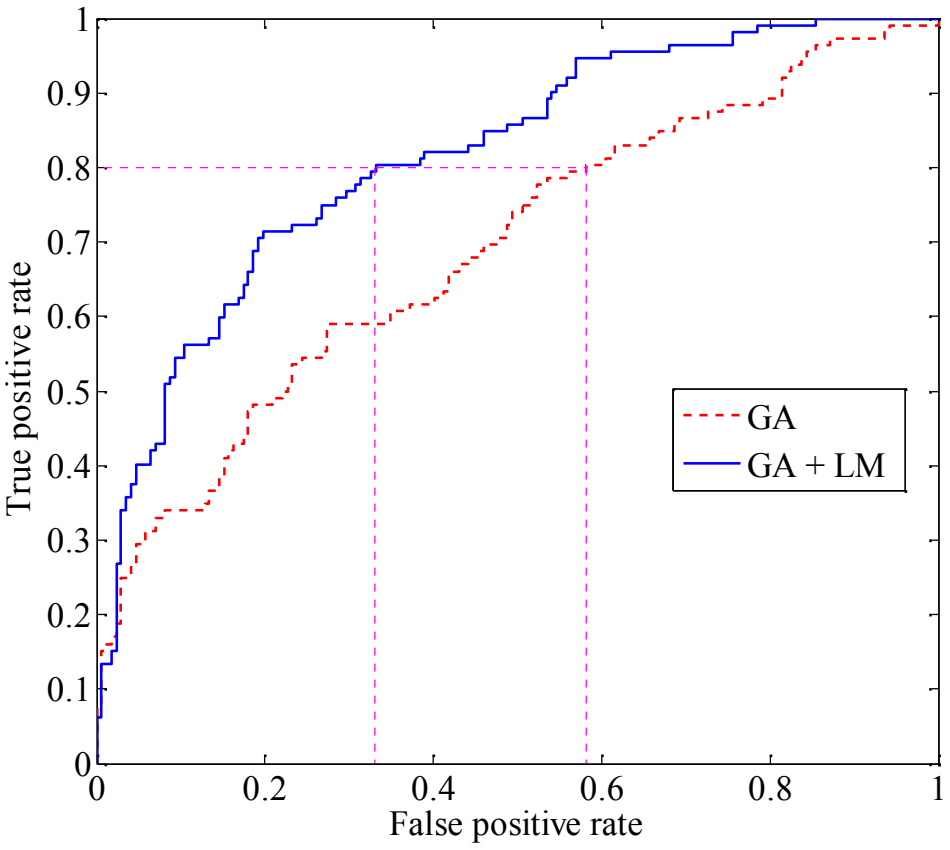


Figure 4.4: ROC curve

The classification results shown in Table 4.5 indicate that by combining GA and LM algorithms to train neural network, classification performance is enhanced markedly. The ROC curves plotted in Figure 4.2 demonstrates that the GA+LM algorithm (producing *AuC* of 0.82) outperforms the GA algorithm (producing *AuC* of 0.74) in training neural networks. As a result, on the same training set, the GA+LM-based neural network yields the best results of 80% sensitivity and 61% specificity which is obviously better compared to the best results yielded by the GA-based neural network (80% sensitivity and 42% specificity). In terms of hypoglycemia detection, the neural network trained by the GA+LM algorithm also produces considerable testing results 75% sensitivity and 60% specificity (which are the best testing results of 20 running times). The consistent mean testing results of 74% sensitivity and 52% specificity also show good generalisation of the neural network trained by the GA+LM algorithm.

It is demonstrated that by using the GA+LM algorithm for training neural network, the classification performance is significantly improved when compared with the GA algorithm and the LM algorithm. This improvement can be explained by the combination of benefits from each separate algorithm. The consistent and good generalisation of neural network profits from the ability of the GA algorithm to direct the training process to the region of optimal solution. This step helps the training process to avoid trapping into local optimal which is an inherent problem of derivative-based training algorithm. Taking this advantage of the GA algorithm, combined with the fine tuning capability of the LM algorithm, the GA+LM algorithm produces considerably enhanced classification performance.

4.5 Discussion

In this chapter, a combination of genetic algorithm (GA) and Levenberg-Marquardt (LM) algorithm (named GA+LM algorithm) is investigated in training neural network with the aim of improving the classification performance for the hypoglycemia detecting algorithm. As a result of the feature extraction carried out in Chapter 3, four EEG parameters from two non-invasive EEG channels of C3 and O2 are used as inputs for a neural network-based classification unit. The developed neural network is trained by a procedure of two consecutive steps including a step of global search and a step of local search. The global search, based on the GA algorithm, aims to drive the training process to the region of the global optimal. Using the solution yielded by the GA algorithm as the initial set of network parameters, the training process continues implementing the next step of local search by the LM algorithm. This local search step acts as a fine tuner helping the training process (currently staying at the global optimal area) to get closer to the final global optimal.

When comparing with the classification results yielded by the neural network trained by the LM algorithm in 20 running times, the proposed GA+LM algorithm produce much better mean results on the testing set (74% sensitivity and 52% specificity versus 67% sensitivity and 42% specificity). This outperformance of the GA+LM algorithm compared to the LM algorithm shows the efficiency of the GA-based global search step in consistently directing the training process to the global region.

On the other hand, the comparison between the classification performance produced by the neural network trained by the GA algorithm and the GA+LM algorithm indicates benefits of the LM-based local search step in training neural network. Although the GA algorithm is effective in global searching, the mean classification results produced by the GA algorithm (80% sensitivity and 37% specificity on the training set) show its failure to finely tune and converge the training process to the final global optimal. By implementing one more step of

LM-based local search, the classification performance is improved significantly up to 80% sensitivity and 57% specificity (mean results on the training set of 20 running times).

In conclusion, by utilizing the global search ability of GA and the local search ability of LM in training neural network, it is demonstrated that classification results of hypoglycemia detection from EEG signals can be enhanced remarkably up to 80% sensitivity and 61% specificity on the training set, and 75% sensitivity and 60% specificity on the testing set.

Chapter 5

Adaptive Strategy of Classification for Detecting Hypoglycemia using EEG Signals

---

5.1 Introduction

The severity and frequency of hypoglycemia in patients with type 1 diabetic mellitus (T1DM) lead to the necessity of developing devices that can detect hypoglycemic episodes and give an alarm to patients to take action before their conditions become severe and cannot be corrected. Using EEG parameters as inputs of the computational detecting methodology, it has been shown in previous chapters that the classification algorithm plays an important role in determining the performance of hypoglycemia detection. In this thesis, the artificial neural network is selected to be used as the classification unit. The combination of genetic algorithm (GA) and the Levenberg-Marquardt (LM) algorithm (GA+LM algorithm) is demonstrated to be an effective technique for training neural network. As shown in Chapter 4, compared to other techniques, this training method produces significantly enhanced classification results of hypoglycemia detection (best classification results of 80% sensitivity and 61% specificity on the training set; 75% sensitivity and 60% specificity on the testing set).

The difference in classification performance between the training set and the testing set can be explained by several causes, which include the generalisation ability of the developed neural network and the variability of EEG signals from person to person. With the proposed classification method of GA+LM-based neural network in Chapter 4, it is shown that the training process is successfully directed to the global optimal area by the GA algorithm, therefore enhancing the generalisation of the neural network. Meanwhile, the high variability is an inherent characteristic of EEG signals which raises considerable difficulties for most EEG-based applications in achieving desired performance.

The main objective of this chapter is to deal with the aforementioned second problem causing the difference in classification results when the neural network is trained and tested with EEG data from different groups of patients. To do this, an adaptive strategy of training neural network will be explored in order to allow the classifier to customise itself to new EEG patterns from new individual users. It will be demonstrated that by applying a properly adaptive strategy of classification, the effect of EEG variability from person to person can be eliminated, therefore improving the performance of detecting hypoglycemia from EEG signals.

Lastly, based on the performance of different algorithms and strategies of training neural networks which have been explored so far in this and previous chapters, the final network training procedure for hypoglycemia detection using EEG signals proposed by this thesis will be presented. This training procedure is separated into two sequential stages including the GA+LM algorithm implemented in conjunction with the adaptive strategy. Classification results of this training procedure provided at the end of this chapter are considered as the final performance of the computational methodology of hypoglycemia detection using only 2 EEG channels developed by this thesis.

5.2 An adaptive strategy for neural network training in the application of detecting hypoglycemia from EEG signals

5.2.1 An adaptive strategy for training neural network to enhance the generalisation of classification

Artificial neural networks have been employed popularly in the biomedical area as a powerful tool of classification and pattern recognition. It has been recognised that using neural networks is a successful method in classifying complex situations. It can effectively model non-linear relationships between inputs and outputs. When assessing the performance of a neural network, the generalisation is one of the most important factors. In terms of classification, the generalisation ability of a neural network means that after being trained, the network could perform well on data of the same class as the learning data that it has never seen before.

There are two main reasons for the limited generalisation ability of a neural network which are over-fitting and trapping into local optimal. Over-fitting is the problem which occurs when the training process lasts too long and the training error function (or the cost function) is forced to be a very small value. In this situation, the network will perform very well on that particular training set because it has memorised the training samples but it cannot learn to adapt to new situations. Trapping into local optimal is also a common but inherent problem of some training algorithms which leads to poor generalisation of neural network. Due to being trapped into one of the local optimal, the classification results can be high on the training data but not acceptable on the unseen testing data.

In this thesis, the cross-validation technique (as presented in Chapter 3, section 3.2.3.3) has been implemented together with the LM algorithm to help the training process avoid over-fitting while the GA algorithm (as presented in Chapter 4, section 4.3) is employed to help the training process direct to global optimal area instead of trapping into one of the local optimal. As a result, using EEG data from 5 T1DM patients, trained on a data set of 3

patients and tested on another unseen data set of 2 patients, classification results are demonstrated to be enhanced up to 80% sensitivity and 61% specificity on the training set; 75% sensitivity and 60% specificity on the testing set. This classification results show a good ability to generalise the trained neural network on data from totally unseen individuals. It is also shown that although the generalisation ability of neural network is enhanced by applying the cross-validation technique and genetic algorithm, there is still a gap between the classification performance on the training set and the testing set. This situation can be explained by the inherent variability of EGG signals from person to person. As this is a natural characteristic of EEG signals, it cannot be overcome totally by training algorithms.

In many previous EEG-associated works, it has been noted that EEG patterns considerably vary from individual to individual. This fact leads to considerable difficulties in generalising an EEG-based system to new users. That is the reason why in most of the studies using EEG signals as inputs for developing health-related applications (such as a brain-computer interface), researchers normally apply some specific strategies in order to overcome that inherent factor of EEG signals.

One common method proposed by many studies is to individualise the system for each subject. In this way, when the system is exposed to a new subject, data from the subject will be collected and used to train (or teach) the system to generalise itself to his/her individual EEG patterns. Generally, this is the most effective way to eliminate the impact of EEG variability from person to person and can considerably enhance the performance of the system on each subject. However, individualising any system to each user is normally accompanied by complicated procedures of initialising and operating, due to the requirement of collecting data and on-line training the system the first time it is used. Moreover, the fact that the system needs to be built to be able to generalise itself to each individual, with complex requisites for designing and manufacturing, obviously makes its cost higher. As a result, these drawbacks make the method less appealing and desirable for the purpose of developing real-life systems that can be commercialised on the market.

Unlike the above mentioned method, in this chapter, in order to reduce the impact of signal

variability on the overall performance of hypoglycemia detection, a strategy of training neural network adaptively will be explored. Instead of implementing the whole process of network training online for each new user, the training process is only allowed to adjust itself to new EEG patterns of an unseen subject. This is accomplished by employing a step of adaptive training in order to enhance the classification performance. To do this, the system is initially trained and validated with data from a group of subjects. After achieving the neural network's optimal structure as well as parameters, an adaptive training step is applied. The previously-trained neural network is then updated by an adaptive-training process with a small baseline set of data taken from an unseen subject. This further-trained neural network is then tested with the testing data from that subject to validate its performance.

In this section, in order to demonstrate the ability of the proposed adaptive strategy in terms of enhancing the generalisation of the classification, first it will be applied to the standard neural network trained by the LM algorithm. As shown in Chapter 3, section 3.3.3.2, using the centroid alpha frequency and centroid theta frequency at two channels of C3 and O2 as inputs, a neural network structure of 4 input nodes, 8 hidden nodes and 1 output node is developed. This neural network is trained by the LM algorithm using data from 3 patients and then tested using a testing set formed from two totally unseen patients. This approach is shown to yield acceptable classification results for the purpose of hypoglycemia detection (80% sensitivity and 53% specificity on the combined training/validation dataset, and 71% sensitivity and 54% specificity on the testing set). It has been concluded that the difference between classification performance on the combined training/validation set and the testing set is partly caused by the inconsistency of EEG signals from patient to patient.

For the purpose of adaptive training, based on optimal network parameters obtained by the training process, the trained neural network is then updated by an adaptive-training step using a small baseline set of data taken from two unseen patients. By doing this, network parameters will be further adjusted in order to be adaptable to their EEG patterns. To evaluate the performance of this further-trained neural network, the same testing data from those two patients that were used to test the developed network in Chapter 3 will be applied.

5.2.2 Classification results

As presented in Chapter 3, section 3.3.3.2, for training and testing the developed neural network of 4 input nodes, 8 hidden nodes and 1 output node, the overall data set acquired from 5 T1DM patients, who participated in the overnight glucose clamp study, is separated into 3 different data sets including a training set, a validation set and a testing set. The training set and validation set are formed by randomly dividing the data set from 3 patients who have been called patient A, B and C. The size ratio of training set to validation set is 3:1. The testing set is formed from the data of two other patients who have been called patient D and E. Details about the number of data points for each data set is provided in Chapter 3, Table 3.10.

In this section, in order to implement the step of adaptive training, another data set, named the baseline data set, is formed by taking 30 data points from the baseline phase of each patient D and E (as defined in Chapter 3, section 3.2.1). Based on the corresponding blood glucose levels, all data points in this baseline phase are classified as being in a non-hypoglycemic state. This baseline data set from two patients D and E is then blended into the data set from 3 patients A, B, and C to form a new data set, called the adaptive training data set which will be used to update the network in the step of adaptive training. Consequently, the adaptive training data set consists of 344 data points in which there are 112 points of hypoglycemia. By further training the network using the adaptive training data set that includes a small part of data from patients D and E, the previously trained network can adjust itself to the new EEG patterns of the unseen patients.

As a result of the network training process, by using the LM algorithm and the cross-validation technique implemented in Chapter 3, the final neural network structure with optimal network parameters has been derived. For the adaptive training purpose, the previously achieved optimal network parameters are set as initial weights and biases of a neural network with the same structure of 4 input nodes, 8 hidden nodes and 1 output node. Afterwards, the adaptive training step will be implemented on this neural network, using the adaptive training data set. The adaptive training procedure is also implemented by the LM algorithm and the cross-validation technique. To carry out the cross-validation

technique, the adaptive training data set is then randomly subdivided into a training set and a validation set with the ratio of 3:1.

The network is trained until one of the terminating conditions is met. At this step of adaptive training, the terminating conditions are set up as follows:

$$\begin{cases} E(\mathbf{w}) = 10^{-3} \\ E(\mathbf{w}) \text{ on the validation set keeps increasing for 200 iterations} \end{cases}$$

where $E(\mathbf{w})$ is the error function which is used as a cost function for the training process defined in Chapter 3, section 3.2.3.2. After the adaptive training step is stopped, the ROC curve for the further-trained neural network will be plotted to re-estimate the output threshold to distinguish between hypoglycemia and non-hypoglycemia states by selecting the point that produces the classification result of 80% sensitivity on the adaptive training data set.

Table 5.1: Classification results of the developed neural network with 4 input nodes, 9 hidden nodes and 1 output node

		Training set		Testing set	
		Sen	**Spe**	**Sen**	**Spe**
Original Training	Best results	80%	53%	71%	54%
	Mean results	80%	52%	67%	42%
Adaptive Training	Best results	80%	58%	74%	55%
	Mean results	80%	54%	70%	51%

Sen: sensitivity ; Spe: specificity

To show how well the neural network updated by the step of adaptive training generalises to two testing patients D and E, the testing data set is applied to the further-trained neural network. All classification results are presented in Table 5.1, in which the reported results are the best and mean results of 20 running times. For comparison purposes, classification results yielded by the original training process implemented in Chapter 3 are also presented in this table.

The classification results of 20 running times indicates that the neural network which is further-trained by the adaptive strategy produces better performance compared to the originally trained network. It is shown that by applying partly individual training and allowing the neural network to adjust itself to the EEG patterns of each subject, the mean classification results of 20 running times on the two testing patients are markedly enhanced from 67% sensitivity and 42% specificity up to 70% sensitivity and 51% specificity. It is widely known that the LM algorithm can drive the training process to be trapped into one of the local optimal solutions which leads to the low network ability to generalise. Consequently, the aforementioned mean classification performance indicates that the impact of the variability of EEG signals from subject to subject on classification performance can be eliminated by implementing the adaptive training strategy, therefore enhancing the generalisation of the overall system.

As shown in Table 5.1, it is noted that there is a slight difference between the classification results on the training set, between the neural network trained by the original training process with the adaptive training process (mean results of 80% sensitivity and 52% specificity versus 80% sensitivity and 54% specificity). This insignificant difference can be simply explained by the addition of 60 baseline data points from 2 patients D and E to the original training data set when implementing the step of adaptive training. Because the purpose of this adaptation step is tuning the developed neural network's parameters obtained by the original training process in order to adapt the neural network to EEG patterns of the two new patients, this similarity of classification performance on the training set is reasonable and predictable.

The classification performance obtained in this section implies that in the application of

hypoglycemia detection from EEG signals, the inherent drawback of the EEG which is the substantial signal variability from person to person can be considerably limited by implementing the adaptive training strategy to update the LM-trained neural network. Combined with the classification performance yielded by the advanced training method of GA+LM algorithm proposed in Chapter 4, it is suggested that a combination of different training strategies can successfully enhance the generalisation ability of the neural network-based classification unit for hypoglycemia detection from EEG signals. This combination will be explored in the next section in a way that it can overcome limitations of training algorithms as well as EEG signals with the aim of improving the ability of the overall system to generalise to new subjects.

5.3 Implementation of GA+LM algorithm in conjunction with adaptive strategy in training neural network for hypoglycemia detection from EEG signals

5.3.1 Procedure of the GA+LM+Adaptive algorithm for training neural network

Thus far in this thesis, artificial neural networks have been demonstrated to be an effective method of classification for the application of detecting episodes of hypoglycemia from EEG signals for patients with T1DM. In order to improve the performance of neural networks for hypoglycemia detection, especially the generalisation of networks, different training techniques and strategies have been explored. Combining the global search ability of genetic algorithm (GA) with the local search ability of the Levenberg-Marquardt (LM) algorithm, the GA+LM algorithm proposed in Chapter 4, section 4.4 is shown to be effective in directing the network training process to the global optimal, therefore significantly enhancing the classification performance of the developed neural network. Meanwhile, classification results presented in section 5.2 indicate that the adaptive strategy

of training network is a capable approach for dealing with the variability of EEG signals from person to person in order to enhance the ability to generalise to a totally new subject of the whole system.

In this section, with the aim of utilizing advantages of both approaches, the GA+LM algorithm will be implemented in conjunction with the adaptive strategy for training neural network as the final neural network training procedure proposed by this thesis for hypoglycemia detection using EEG signals. Using the EEG data acquired from five patients who participated in the overnight glucose clamp study, the final network training procedure (named the GA+LM+Adaptive algorithm) which consists of two consecutive stages will be applied. Details about each stage of the proposed GA+LM+Adaptive algorithm for neural network training is provided as follows:

- ***Stage 1:*** The developed neural network is initially trained by the GA+LM algorithm using a training set which is formed from the data of three patients who have been called patient A, B and C. In brief, this stage aims to guide the training process to the global optimal solution by two sequential steps. The first step of GA training helps to direct the training process to the region of the global optimal without trapping into local optimal solutions, while the second step of LM training acts as a fine tuning tool, managing to drive the training process closer to the final optimal solution. Details about using the GA+LM algorithms to train neural network are provided in section 4.4.1.

- ***Stage 2:*** In the beginning of this stage, the optimal set of network parameters obtained by the GA+LM algorithm in ***Stage 1*** will be set as initial parameters for the developed neural network. Then the second stage of adaptive training will be applied. Details about the adaptive strategy of training neural network are provided in section 5.2. Briefly, the trained neural network obtained by ***Stage 1*** will be adaptively updated by a further-training step in order to help the network customise itself to EEG patterns of new users. This step of further training is carried out by utilizing a small baseline set of data taken from two unseen patients (who have been called patient D and E) added to the original training set to form a new adaptive

training set. In this stage, the LM algorithm is applied to train the network. Whenever one of the LM terminating conditions is fulfilled, the stage of adaptive training will be stopped. The set of network parameters at the last training iteration will be acquired as the final optimal network solution of the GA+LM+Adaptive algorithm for training neural network proposed by this thesis. The same process of plotting ROC curve for the final neural network will be applied to estimate the output threshold to distinguish between hypoglycemia and non-hypoglycemia. Based on this threshold, the testing data from two patients D and E will be applied to verify the performance of the final neural network.

A summary of the procedure to implement the GA+LM+Adaptive algorithm for training neural network is presented in Figure 5.1-5.2. In this procedure, terminating conditions defined for each stage are presented as in Table 5.2. At each step of the training process, whenever one of the terminating condition is fulfilled, the training will be stopped and the set of network parameters at the terminating iteration will be considered as the optimal parameters set produced by that training step. The set of network parameters obtained at the final terminating iteration of the second stage will be considered as the final optimal solution of the network training process.

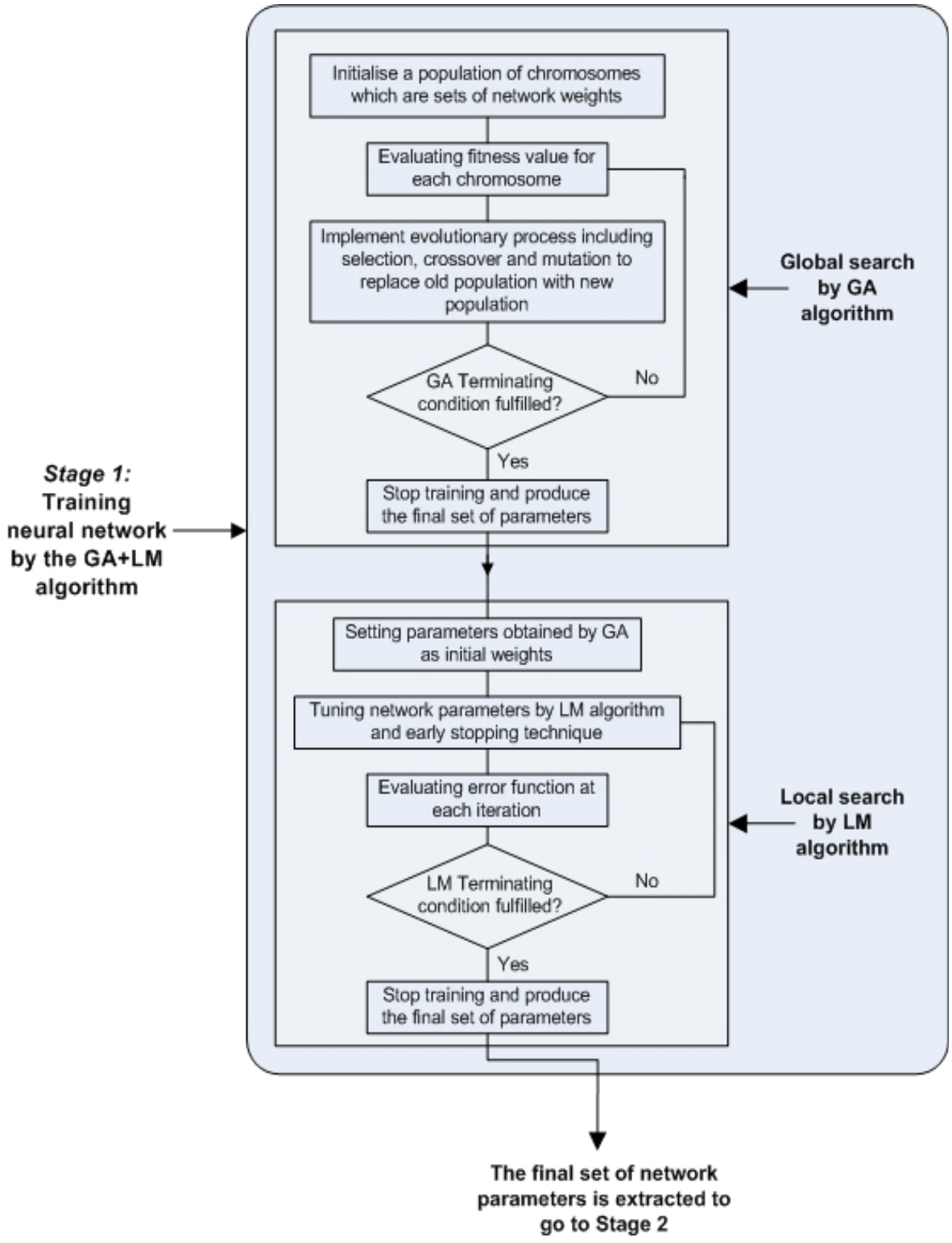


Figure 5.1: Procedure of implementing the GA+LM+Adaptive algorithm for training neural network

Stage 1

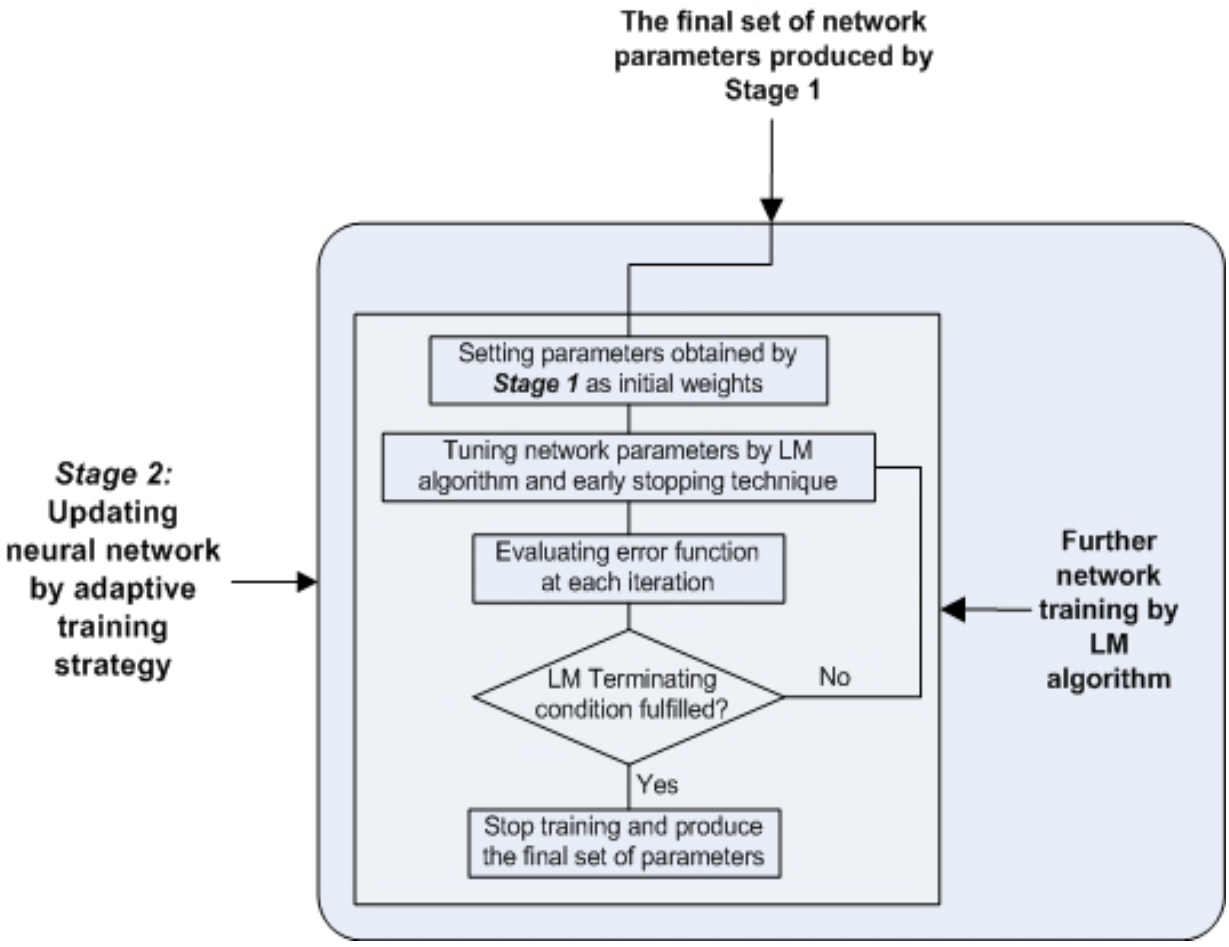


Figure 5.2: Procedure of implementing the GA+LM+Adaptive algorithm for training neural network

Stage 2

Table 5.2: Terminating conditions setting for training neural networks by the GA+LM+Adaptive algorithm

Stage of training		**Terminating conditions**
Stage 1	Global search by GA algorithm	$E(\mathbf{w}) = 10^{-2}$ Maximum number of generations = 2000
	Local search by LM algorithm	$E(\mathbf{w}) = 10^{-3}$ $E(\mathbf{w})$ on the validation set keeps increasing for 200 iterations
Stage 2	Adaptive training by LM algorithm	$E(\mathbf{w}) = 10^{-4}$ $E(\mathbf{w})$ on the validation set keeps increasing for 200 iterations

$E(\mathbf{w})$: the error function which is used as cost function for the training process, defined in Chapter 3, section 3.2.3.2.

5.3.2 Classification results

In this section, for classification purposes, a neural network structure of 4 input nodes (including 2 EEG features of centroid theta frequency and centroid alpha frequency at two channels C3 and O2), 9 hidden nodes and 1 output node is developed as the classification unit. With the aim of training the developed neural network using the GA+LM+Adaptive algorithm, data from 5 T1DM patients who participated in the overnight hypoglycemia-associated study are grouped into different data sets corresponding to each stage of the training process. The number of data points used for training and validating in each stage is summarized in Table 5.3. Details about forming data sets for each stage are provided in sections 3.3.3.1, 4.3.2 and 5.2.2.

The developed neural network is trained by the GA+LM+Adaptive algorithm which includes two consecutive stages of the GA+LM algorithm implemented in conjunction with the adaptive training strategy. The optimal solution of network parameters produced by the first stage is set as the initial network parameters ready for the start of the second stage. Whenever one of the terminating conditions is fulfilled, the final set of network parameters produced by the second stage is obtained and considered as the final optimal solution of the network training process.

After being trained by the GA+LM+Adaptive algorithm, the ROC curve of the trained neural network will be plotted for the adaptive training set (the combined dataset of Adaptive-LM-training subset and Adaptive-LM-validation subset, referring to Table 5.3) to estimate the *AuC* as well as to determine the cut-off point which produces a classification sensitivity of 80% on the adaptive training set. The ROC curve which corresponds with the best classification results in 20 running times is provided in Figure 5.3, producing the *AuC* of 0.79 and the cut-off point of -0.3013. Using this cut-off point as output threshold to distinguish between hypoglycemia and non-hypoglycemia, classification performance on each dataset will be determined. All final classification results are presented in Table 5.3. The reported results are the mean and best performance of 20 running times.

Table 5.3: Number of data points for training and testing neural network using GA+LM+Adaptive algorithm

				Total	Hypo	Non-hypo
Training	**Stage 1** *(data from patients A, B and C)*	**Global search**		**284**	**112**	**172**
		Local search	LM - Training	**213**	**84**	**129**
			LM - Validation	**71**	**28**	**43**
	Stage 2 *(data from patients A, B and C + baseline data from patients D and E)*	Adaptive-LM-Training		**258**	**84**	**174**
		Adaptive-LM-Validation		**86**	**28**	**58**
Testing *(data from patients D and E)*				**144**	**76**	**68**

Hypo: Hypoglycemia
Non-Hypo: Non-hypoglycemia

Table 5.4: Classification results of the neural network with 4 input nodes, 9 hidden nodes and 1 output node trained by the GA+LM+Adaptive algorithm

		Training set			Testing set	
		AuC	**Sen**	**Spe**	**Sen**	**Spe**
GA + LM	Best results	0.82	80%	61%	75%	60%
	Mean results	0.79	80%	57%	74%	52%
GA + LM + Adaptive	Best results	0.79	80%	60%	78%	62%
	Mean results	0.78	80%	57%	76%	59%

AuC: Area under the curve
Sen: sensitivity ; Spe: specificity

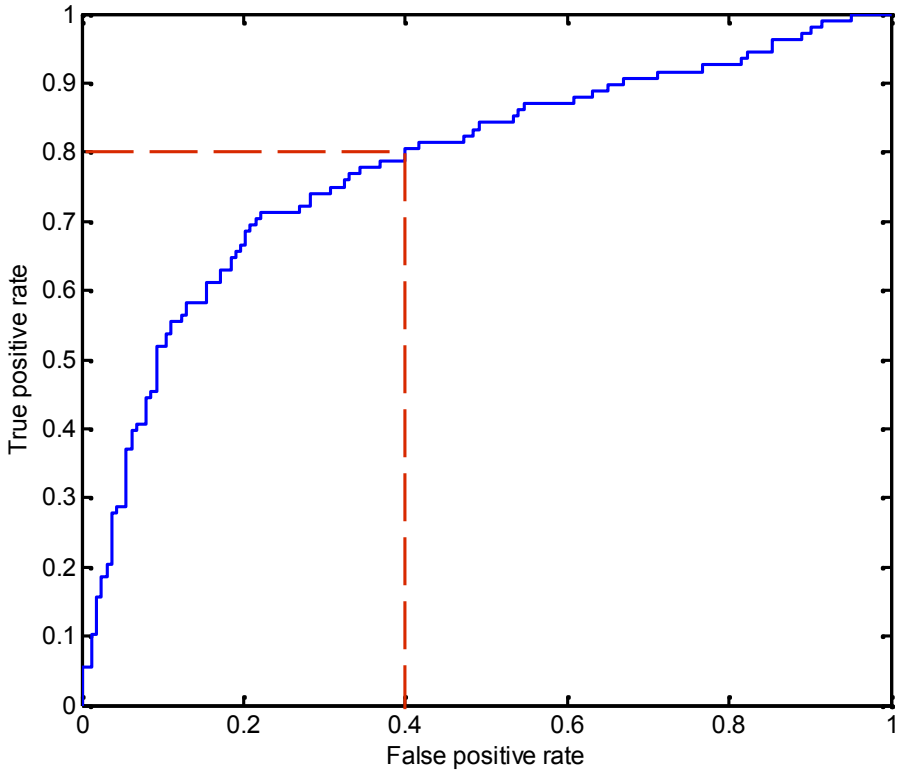


Figure 5.3: ROC curve

Classification results presented in Table 5.4 show that the GA+LM+Adaptive algorithm is an effective method of training neural network in the application of hypoglycemia detection. In all 20 running times of network training, it is recognised that by implementing one more stage of adaptive training after the first stage of the GA+LM algorithm, the classification performance is enhanced considerably. With the same network structure of 4 input nodes, 9 hidden nodes and 1 output node, compared to the performance yielded by the GA+LM algorithm implemented in Chapter 4, the neural network trained by the GA+LM+Adaptive algorithm produces improved mean results of 76% sensitivity and 59% specificity on the testing set. With the mean performance on the training set of 80% sensitivity and 57% specificity, these results show that the developed neural network has a good capability of generalising to unseen data from new subjects.

The enhancement of generalisation generated by applying GA+LM+Adaptive algorithm for training neural network can be explained by the combination of advantages of each training stage. With the ability of the GA+LM algorithm to consistently direct the training process to the global optimal solution, the ***Stage 1*** helps to overcome the limitations of trapping into local optimal as well as over-fitting of gradient-based training algorithms. With the ability of the adaptive training stage to allow the network to customise itself to new EEG patterns of unseen testing data, the ***Stage 2*** helps to overcome the variability of EEG signals from subject to subject.

The best classification performance yielded by the final methodology of classification presented in this section (80% sensitivity and 60% specificity on the training set and 78% sensitivity and 62% specificity on the testing set) demonstrates that nocturnal hypoglycemic episodes can be successfully detected for patients with T1DM using EEG signals from only two EEG channels. With the use of a neural network with the structure of only 4 input nodes, 9 hidden nodes and 1 output as the classification unit, combined with an efficient strategy for training the developed network adaptively, there is the potential to pursue the proposed computational methodology to develop a device for the purpose of detecting hypoglycemia for T1DM patients from EEG signals that can be used in the real clinical environment.

5.4 Discussion

In this chapter, an adaptive training strategy is investigated for training the neural network developed in previous chapters with the aim of eliminating the impacts of variability in EEG signals from person to person on classification performance of hypoglycemia detection. It has been shown that the inconsistency of EEG patterns between different individuals is one of the reasons causing the dissimilarity in classification results between the training set and testing set from unseen subjects. For the purpose of enhancing the ability to generalise to new users of the developed neural network, the strategy of training network adaptively allows the neural network to customise itself to the distinctive EEG patterns of each unseen patient by a further training step. Setting the optimised set of network parameters yielded by the original training step as initial weights and biases, the training process continues implementing the step of adaptive training to update the previously obtained neural network. This adaptive training step is based on an adaptive training data set, including the original training set plus a small baseline non-hypoglycemic data set from unseen testing subjects.

The classification results presented in this chapter demonstrate that the proposed adaptive strategy of training neural networks is an effective approach which remarkably improves the classification performance on unseen subjects. Due to the well-known drawback of trapping into local optimal of the LM algorithm, it has been shown in chapter 3 that the limited mean classification results of 20 running times on the testing set of 67% sensitivity and 42% specificity are predictable. Consequently, with the results of 70% sensitivity and 51% specificity yielded by implementing one more step of adaptive training as shown in section 5.2, it is concluded that the influence of EEG variability on classification performance can be reasonably eliminated.

It should be noted that in terms of hypoglycemia detection, the adaptive strategy of training neural network proposed in this chapter is dissimilar from the calibration technique that has been used widely in most devices that are currently available on the market, which aim to monitor patients' blood glucose concentration and detect hypoglycemic episodes. Because the main purpose of calibration is to help a system learn about physiological responses of a

patient's body to changing conditions throughout the day, usually this technique requires users to carry out the calibration frequently during its operating period, even several times in a day (for example, the Guardian Real-Time Continuous Glucose Monitoring System from Medtronic requires the first calibration at 2 hours after sensor insertion, the second calibration within the next 6 hours after the first, and then every 12 hours during the operating period). Meanwhile, the adaptive strategy proposed here aims to help the neural network training process customise itself to the distinctive EEG patterns of a new user, therefore requiring implementation only once on the first time exposure to the device during their normal, non-hypoglycemic state

As a result of exploring various techniques for training neural networks that have been carried out in Chapters 4 and 5 with the aim of enhancing the generalisation of the neural network-based classification unit, a procedure named the GA+LM+Adaptive algorithm has been introduced in this thesis. This includes two sequential stages of the GA+LM algorithm in conjunction with the adaptive training strategy, implemented as the final neural network training method. Using data from 5 T1DM patients who participated in the overnight insulin-induced hypoglycemia-associated study, the proposed method is applied to train a neural network structure of 4 input nodes (including centroid theta frequency and centroid alpha frequency at two channels C3 and O2), 9 hidden nodes and 1 output node (indicating hypoglycemic or non-hypoglycemic state). Compared to other approaches that have been implemented throughout the thesis (including LM algorithm, GA algorithm, GA+LM algorithm), the final GA+LM+Adaptive algorithm is shown to yield the best classification results.

The mean classification results of 20 running times (80% sensitivity and 57% specificity on the training set and 76% sensitivity and 59% specificity on the testing set) indicate that the proposed GA+LM+Adaptive strategy is an effective neural network training procedure which successfully enhances the generalisation of the developed network. By utilizing the GA algorithm capability of global searching and the LM algorithm capability of fine tuning, and by combining with the adaptive training strategy which allows the neural network to adapt itself to EEG patterns of new subjects, it is shown that the GA+LM+Adaptive algorithm successfully directs the training process to the global optimal

solution without trapping into local optimal, as well as significantly eliminating the variability of EEG signals from person to person.

Lastly, with the best classification performance of 80% sensitivity and 60% specificity on the training set and 78% sensitivity and 62% specificity on the testing set, it can be concluded that the final methodology of classification proposed by this thesis successfully performs in the application of hypoglycemia detection for T1DM patients using only two EEG channels. It should be noted that due to the importance of the rate of correctly identifying hypoglycemic episodes, the classification results of 80% sensitivity and 60% specificity, which are approximately equivalent on both training data and testing data, are reasonable and desirable for the purpose of monitoring and detecting episodes of hypoglycemia for patients with T1DM.

Chapter 6

Conclusion and Future Work

6.1 Discussion and conclusion

Hypoglycemia, or the state of abnormally low blood glucose level (BGL), is the most common but dangerous complication of the intensive insulin therapy for patients with type 1 diabetes mellitus (T1DM). Hypoglycemia impacts life quality of all T1DM patients, limits their intellectual as well as physical activities, and potentially causes irreversible severe effects, such as cognitive impairments, seizures, coma, and even death. A study in 2004 reported that severe hypoglycemia (defined as episodes in which patients need assistance to re-establish the normal BGL) happens in one third of 1076 self-reported participants with an incidence rate of 1.3 episodes/patient-year (Pedersen-Bjergaard et al. 2004). Nocturnal hypoglycemia is especially dangerous because sleep reduces and obscures early warning symptoms, so that an initially mild episode may become severe. It was reported previously that almost 50% of all episodes of severe hypoglycemia occur at night during sleep (Group 1991). Because of its severity and prevalence, intensive research has been devoted to the development of systems that can detect the onset of hypoglycemic episodes, and then give an alarm to provide enough time for patients and their caregivers to take action.

Under the occurrence of hypoglycemia, the human brain is one of the first affected organs. Because it cannot synthesize as well as store this primary metabolic fuel, the brain depends on a continuous supply of glucose and is vulnerable to any glucose deprivation (Cryer, Davis & Shamoon 2003). Since the electroencephalogram (EEG) is directly related to the metabolism of brain cells, a failure of cerebral glucose supply can cause early changes in EEG signals.

The core objective of this thesis is to introduce a computational methodology of detecting nocturnal hypoglycemia non-invasively from EEG signals for T1DM patients. Two main tasks have been implemented throughout the thesis: (i) analysing EEG signals to extract important features that significantly change under the transition from non-hypoglycemic state to hypoglycemic state; (ii) classifying and detecting hypoglycemic episodes, using EEG features extracted from the previous task as inputs. It is established by this thesis that episodes of hypoglycemia can be detected non-invasively and efficiently by an advanced neural network-based classification algorithm, using two features of centroid theta frequency and centroid alpha frequency as inputs of the algorithm.

The proposal of applying two EEG features of centroid theta frequency and centroid alpha frequency in detecting hypoglycemia is an important contribution of this thesis. The decrease in centroid alpha frequency has been used as a sign of reduction of vigilance in subjects with some other health problems (fatigue, Alzheimer, etc.). It has been shown in this thesis that in all 5 T1DM patients who participated in the glucose clamp study, under hypoglycemic conditions, there is a significant decrease in centroid alpha frequency and there is also a slight increase in centroid theta frequency. Because the lack of vigilance is also a common symptom happening in T1DM patients under hypoglycemic conditions, the two aforementioned changes in centroid alpha frequency and centroid theta frequency can be considered as signs of the early onset of hypoglycemia (occurring at BGL of 3.3 mmol/l).

Using the two aforementioned features as inputs, a neural network-based classification unit with different training strategies is explored to develop a computational algorithm for detecting the onset of hypoglycemia.

- The Levenberg-Marquardt (LM) algorithm is one of the most popular techniques for training neural networks which is based on the second-order gradient information of an error function in order to direct the training process to a local optimal. Standard neural networks trained by the LM algorithm were shown in this thesis to produce acceptable classification results which demonstrate the potential of the proposed methodology for hypoglycemia detection. With the aim of reducing computational burden for the system, a smaller number of EEG electrodes will be preferred. Thus, it is proposed by this thesis to use only EEG data from two channels C3 and O2, which are from two different sides and areas of the brain are used, as inputs for classification. As a result, a neural network with a structure of 4 input nodes (2 features x 2channels), 9 hidden nodes and 1 output node (indicating the state of hypoglycemia or non-hypoglycemia) is developed. This neural network, trained by the LM algorithm provides the best classification performance of 80% sensitivity and 53% specificity on the combined training/validation set and 71% sensitivity and 54% specificity on the testing set. These classification results demonstrate that hypoglycemic episodes can be effectively detected by the proposed methodology of using EEG signals and neural networks. Nevertheless, based on the developed neural network, more advanced algorithms for network training need to be explored in order to enhance the overall performance, as well as to improve the generalisation ability of the trained network on new data from unseen subjects (i.e. testing set).

- In order to overcome a well-known inherent drawback of the LM algorithm which is potentially driving the training process to be trapped in a local optimum, a more advanced technique, named GA+LM algorithm was explored. For the purpose of neural network training, this algorithm implements two consecutive steps of global search and local search. Based on a genetic algorithm (GA), the global search helps to direct the training process to the region of the global optimal without trapping in a local optimal solution. Using the solution yielded by the GA algorithm as the initial set of network parameters, the local search step acts as a fine tuner helping the training process (currently staying at the global optimal area)

to get closer to the final global optimal. By utilizing advantages of both GA algorithm and LM algorithm, the GA+LM algorithm is shown to produce markedly enhanced classification performance up to 80% sensitivity and 61% specificity on the training set, and 75% sensitivity and 60% specificity on the testing set.

- In order to limit effects of the variability in EEG signals from subject to subject which led to considerable difficulties in generalising the system to new users, an adaptive strategy of training neural network was applied in this thesis. The training process first set the optimised set of network parameters yielded by the original training step as initial network weights and biases and updated the obtained network by one more step of adaptive training. This step was implemented by using an adaptive training data set which included a small baseline non-hypoglycemic data set from unseen testing subjects. By doing this, the network was allowed to adapt itself to the new EEG patterns of testing subjects, therefore enhancing the generalisation ability of the developed neural network as well as the overall performance of the whole system.

- Utilizing advantages of each training strategy explored throughout the thesis, a final training algorithm, named the GA+LM+Adaptive algorithm was proposed for the purpose of classifying hypoglycemia from EEG signals in T1DM patients. The algorithm consisted of two sequential training stages. Stage 1 implemented the GA+LM algorithm to help the training process direct to the global optimal solution. Using the optimised network parameters yielded by Stage 1 as initial network weights and biases, Stage 2 implemented the adaptive training using the adaptive training dataset to help the training process customise the network to the distinctive EEG patterns of each new subject. With the same data from 5 T1DM patients who participated in the overnight hypoglycemia-induced study, the network trained by the GA+LM+Adaptive was shown to produce significantly enhanced classification performance up to 80% sensitivity and 60% specificity on the training set and 78% sensitivity and 62% specificity on the testing set. These classification results, which are considered as the final performance yielded by this

thesis, demonstrate that hypoglycemia can be effectively detected from EEG signals by the proposed classification method of using neural network trained by the GA+LM+Adaptive algorithm.

6.2 Future work

In this thesis, a computational method for early detection of nocturnal hypoglycemia using parameters extracted from non-invasive EEG signals was demonstrated. It should be noted that for the purpose of monitoring BGLs and detecting hypoglycemia for T1DM patients, using EEG signals as inputs is a new research area. While devices using other kinds of physiological signals or parameters (e.g. skin temperature, skin impedance, heart rate, cardiac parameters, etc.) have been explored, developed and commercialised much earlier, systems that aim to detect hypoglycemia from EEG signals are still under research and not available on the market. As a new area of research, it can be established that the results achieved by this thesis are acceptable and comparable to other methods of detection. With the potential results yielded by the proposed methodology, there are some possible research directions that would be explored in future works in order to enhance the performance of the overall system.

As shown in this thesis, four different EEG parameters are analysed to find their correlations with the occurrence of hypoglycemia during night in five participated T1DM patients. The theta centroid frequency and alpha centroid frequency are proved to be two important features which significantly changes under the onset of hypoglycemia. For the purpose of detecting hypoglycemia from EEG signals, finding significant EEG features which can be used as inputs of the classification algorithm plays one of the most essential roles in determining the effectiveness of the system. It is expected that there are other EEG parameters that also significantly respond to the occurrence of hypoglycemia. Thus, in the future, different EEG parameters can be continuously explored to find other features that can enhance the performance of the system.

In terms of classification algorithms, standard multi-layer feed-forward neural networks have been shown to be effective classification units for detecting hypoglycemia. In order to improve the accuracy of hypoglycemia detection, in the future, more advanced algorithms can be explored. To do this, there are several options to be investigated.

- First, neural networks with more complicated structures can be developed to enhance the classification efficiency. However, it should be noted that complex network structures will lead to the burden of cumbersome computational cost which may slow down the system operation when implementing the network in real-time. Because this research aims to develop the computational methodology for a detecting system that can effectively work in real-life, this direction has not been explored by this thesis.

- Second, based on the developed standard neural network, different techniques to train and optimise the network structure as well as parameters can be investigated in order to enhance the effectiveness of the developed neural network. In this thesis, three strategies of LM algorithm, GA+LM algorithm and GA+LM+Adaptive algorithm have been introduced for the purpose of training neural network. It is shown that the GA+LM+Adaptive algorithm remarkably enhances the performance of the developed neural network in the application of detecting hypoglycemia. It is expected that more advanced training techniques will help drive the training process closer to the global optimal solution, hence improving the overall classification performance. Furthermore, finding the optimised network structure (i.e. the number of hidden nodes) is also a potential direction which can be implemented in the future.

One of the limitations of this current study is the shortage of data. The data set collected from five participated T1DM patients is sufficient to establish that the onset of hypoglycemia induces early changes in EEG signals which can be detected by using the proposed computational methodology. However, in order to apply this method to develop a hypoglycemia detecting system that can perform in real clinical environments, more data will be needed to validate the achieved results of this thesis. To do this, other clinical

studies with more participants can be implemented in future, in which EEG signals from other brain areas will also be acquired with the aim of comparing and finding the most suitable positions to place EEG electrodes for the detecting system.

Furthermore, it is noted that the data set used in this thesis was acquired from a glucose clamp study, which involved a procedure of inducing hypoglycemia, rather than natural hypoglycemia. Previous works indicated that physiological responses to induced hypoglycemia and natural occurring or spontaneous hypoglycemia are dissimilar. In future work, natural hypoglycemia associated studies would be carried out to validate the possibility of the proposed methodology in detecting the onset of hypoglycemia that occurs spontaneously in the real life.

In conclusion, based on the methodology developed in this thesis, future works will be carried out to pursue the final purpose of developing the real-time system that can efficiently and continuously monitor patients' conditions and alert them as well as their caregivers when the onset of hypoglycemia is detected.

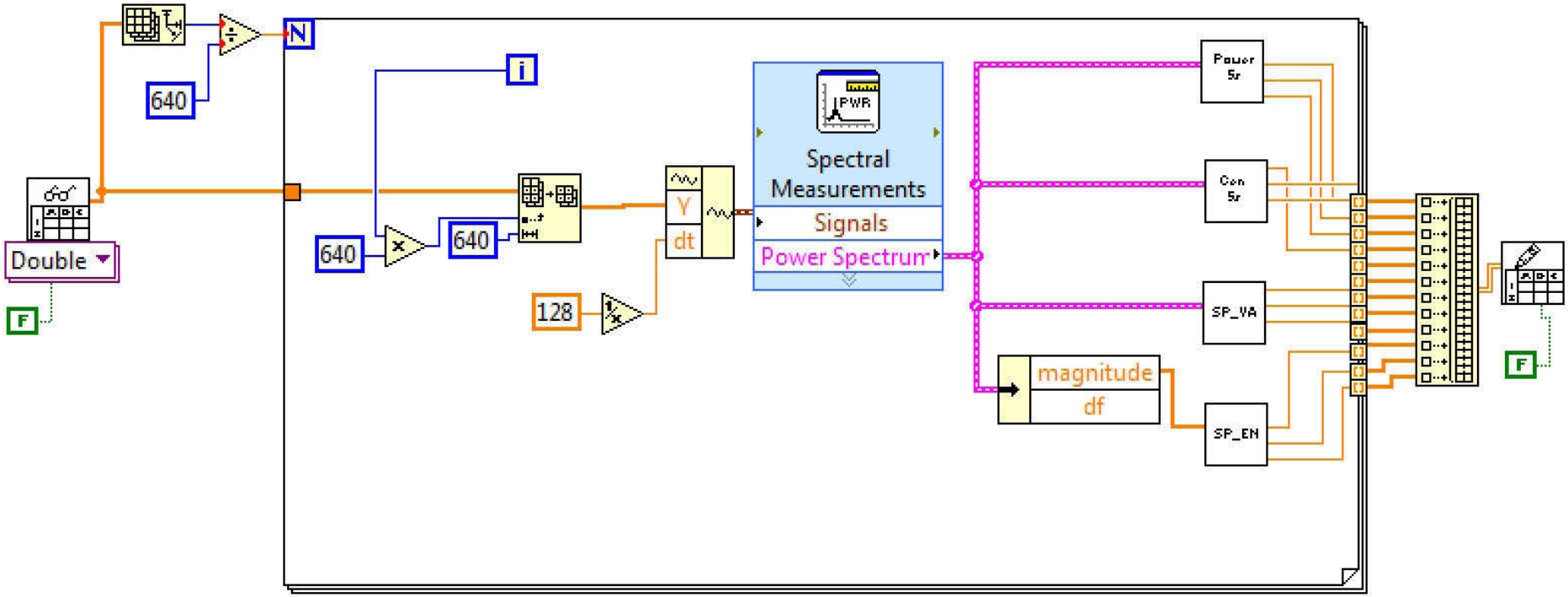
Appendix A

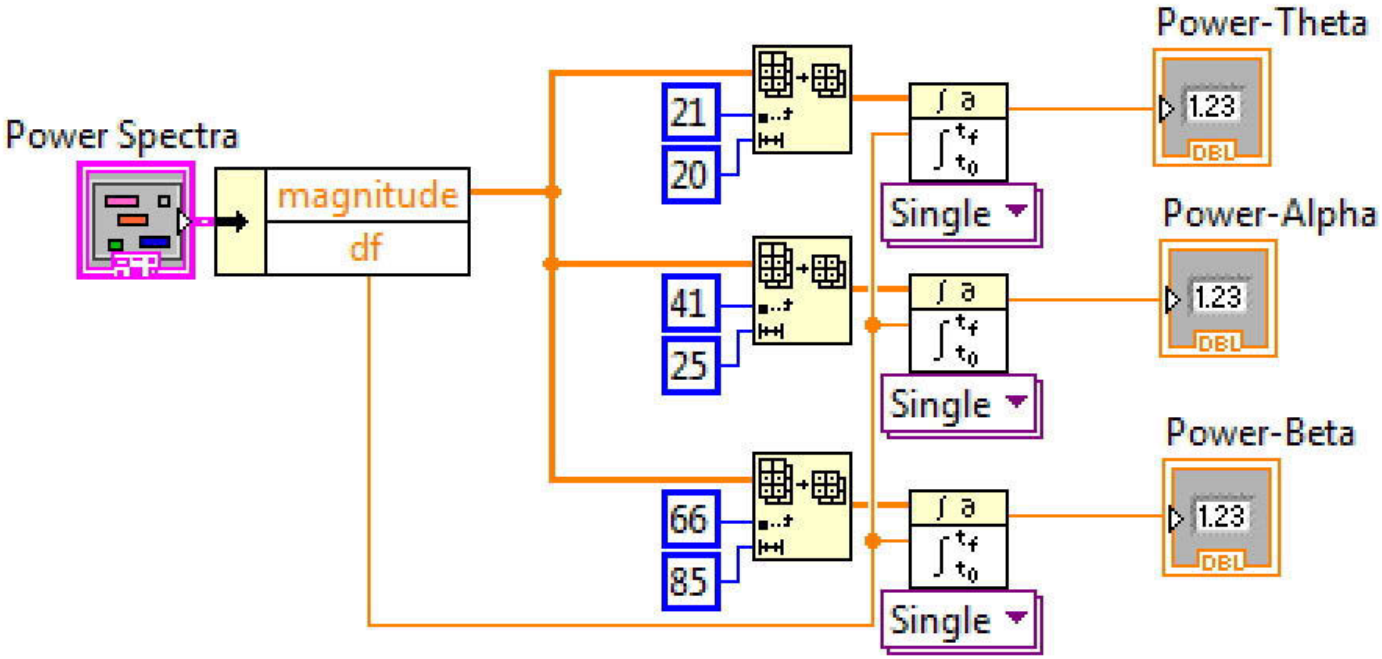
Programming Implementation

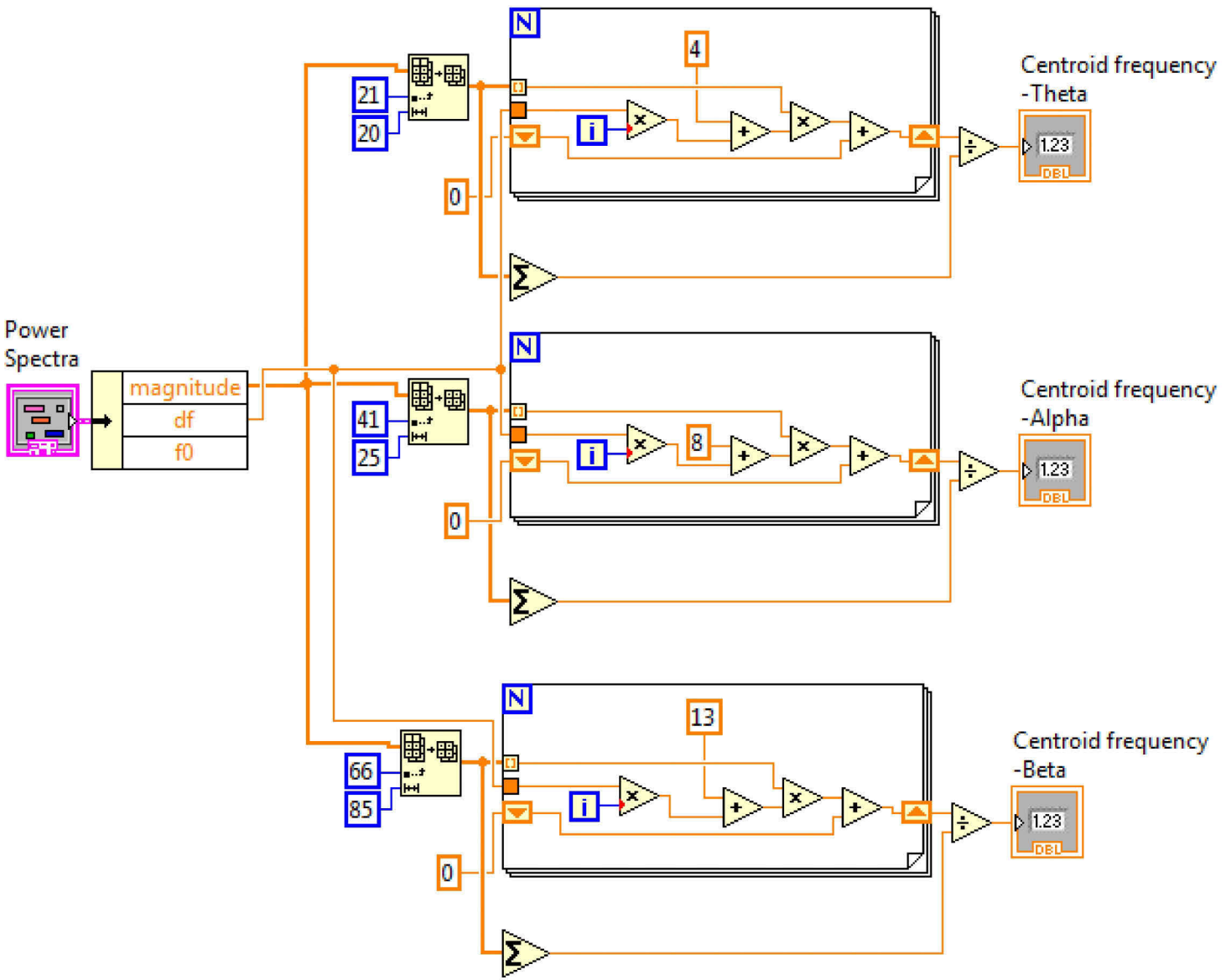
A.1. Feature extraction using Labview

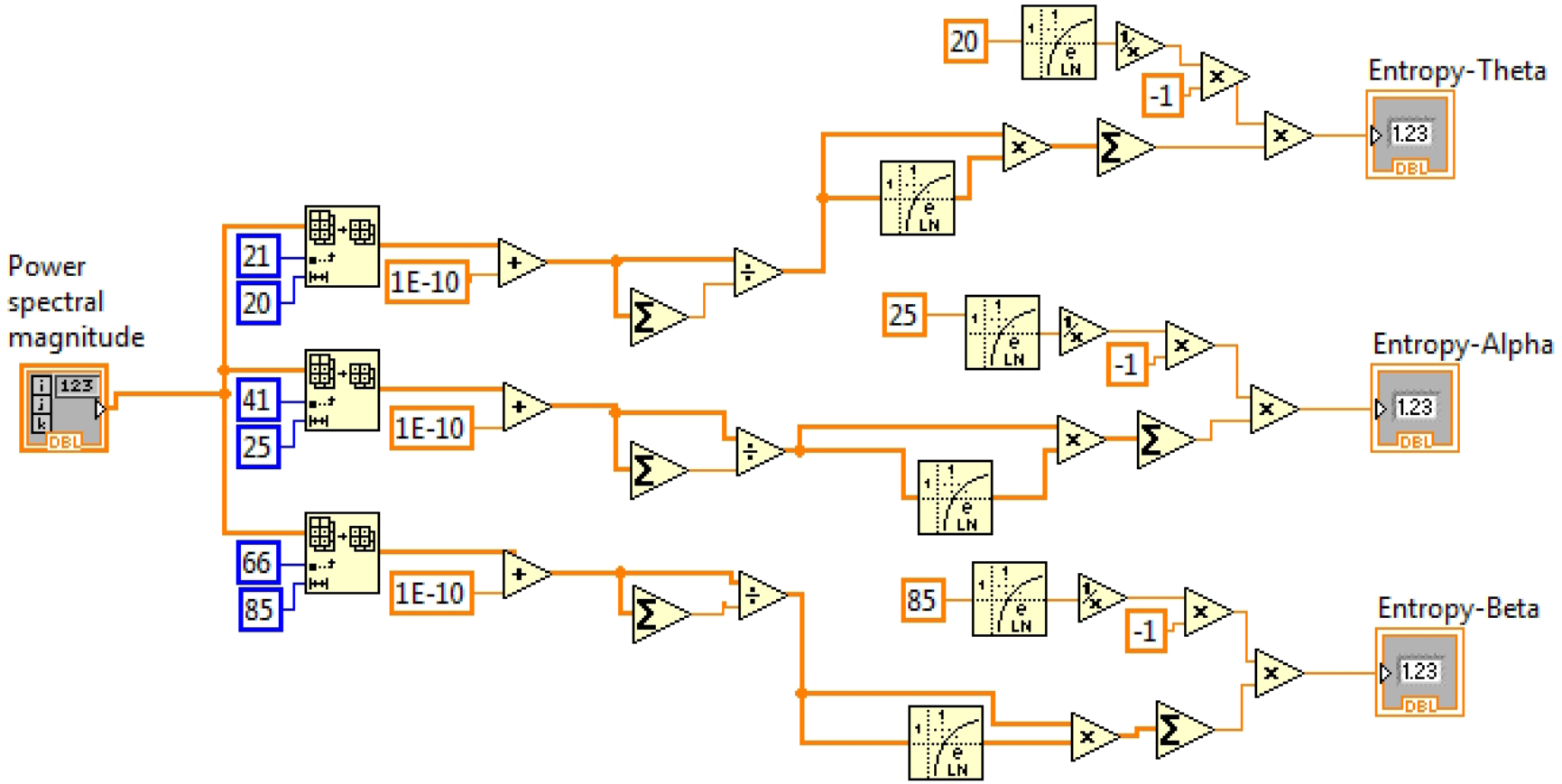
In this thesis, signal processing and feature extraction is implemented in Labview. The following works are built up for the purpose of estimating four EEG parameters of sub-band power, centroid frequency, spectral variance and spectral entropy within each frequency band. The data read into the program is filtered EEG signals. The signal filtering task is implemented by another Labview program. Non-artifact EEG data, after being read into the program, is processed by the provided program for the following purposes:

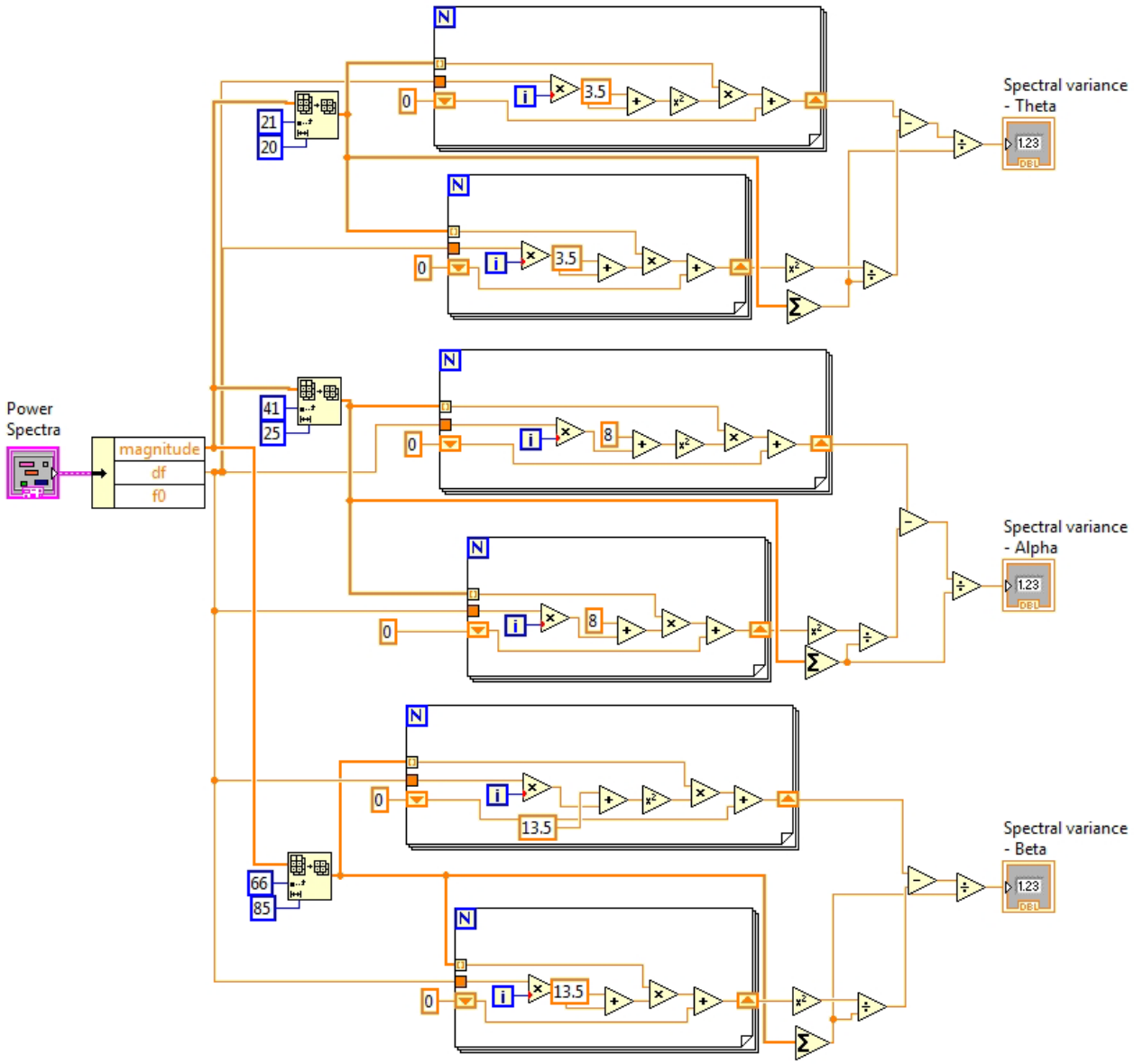
- Segmenting non-artifact data into 5-second epochs
- Transforming each epoch into frequency domain
- Dividing the power spectra of each epoch into three frequency bands of theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz)
- Extracting four aforementioned EEG parameters from each frequency band of each 5-second epoch of signal


640
N
i
640
640
128
Y
dt
PWR
Spectral Measurements
Signals
Power Spectrum
Double
F
Power Sr
Cen Sr
SP_VA
SP_EN
magnitude
df
F


Power Spectra
magnitude
df
21
20
41
25
66
85
Single
Single
Single
Power-Theta
Power-Alpha
Power-Beta
1.23
DBL


Power
Spectra
magnitude
df
f0
21
20
0
4
i
Centroid frequency
-Theta
1.23
DBL
41
25
0
8
i
Centroid frequency
-Alpha
1.23
DBL
66
85
0
13
i
Centroid frequency
-Beta
1.23
DBL


Power
spectral
magnitude
21
20
1E-10
41
25
1E-10
66
85
1E-10
20
-1
Entropy-Theta
25
-1
Entropy-Alpha
85
-1
Entropy-Beta
1.23


Power
Spectra
magnitude
df
f0
21
20
41
25
66
85
0
3.5
8
13.5
Spectral variance
- Theta
Spectral variance
- Alpha
Spectral variance
- Beta
1.23
DBL

A.2. Classification using Matlab

The following work is built up, based on the neural network toolbox in Matlab, for the following purposes:

- Develop a multi-layer feed-forward neural network to detect hypoglycemic episodes, using 4 EEG parameters (centroid theta and centroid alpha at two channels C3 and O2 as inputs) as inputs
- The developed neural network is trained by the LM+GA+Adaptive algorithm proposed by this thesis, implementing the LM+GA algorithm in conjunction with the adaptive strategy.
- The overall data set collected from 5 participated patients are divided into three different set of training set, validations set and testing set. The training and validation set are formed from data of 3 patients, named patients A, B, and C, randomly divided with the ratio of 3:1. The testing set is formed from data of 2 patients, named D and E.
- With the aim of implementing the adaptive training, the adaptive training set is formed by adding 30 baseline non-hypoglycemic data points from each patient D and E into the original training set.

```
% This m-file implements training neural network by the
LM+GA+Adaptive algorithm, in which training and validation sets
are randomly divided from data of 3 patients A, B, C  with ratio
of 3:1. Testing set is formed from data of 2 patients D and E.
It's used to develop and train neural network as the final
procedure proposed by Bich Lien Nguyen's thesis.



% Loading data

clear all;
close all;
load normal_5s; load hypo_5s;load addpD; load addpE;

% Deriving centroid theta frequency and centroid alpha frequency
at two channels C3 and O2

hypo2=[hypo_5s(4:5,:);hypo_5s(40:41,:)];
normal2=[normal_5s(4:5,:);normal_5s(40:41,:)];
addpD=[add2(4:5,:);add2(40:41,:)];
addpE=[add4(4:5,:);add4(40:41,:)];
% normal2=nor_5sN;hypo2=hypo_5sN;
data=[normal2 hypo2 addpD addpE];

% Processing data

for i=1:4
    data2(i,:)=mapminmax(data(i,:),-1,1);
end
normal2=data2(:,1:240); hypo2=data2(:,241:428);
add_datapD=data2(:,429:458);add_datapE=data2(:,459:488);

hypo_pA=hypo2(:,1:36);hpA=size(hypo_pA,2);
hypo_pD=hypo2(:,37:76);hpD=size(hypo_pD,2);
hypo_pB=hypo2(:,77:112);hpB=size(hypo_pB,2);
hypo_pE=hypo2(:,113:152);hpE=size(hypo_pE,2);
hypo_pC=hypo2(:,153:188);hpC=size(hypo_pC,2);

normal_pA=[normal2(:,1:40)];npA=size(normal_pA,2);
normal_pD=[normal2(:,41:68)];npD=size(normal_pD,2);
normal_pB=[normal2(:,69:132)];npB=size(normal_pB,2);
normal_pE=[normal2(:,133:180)];npE=size(normal_pE,2);
normal_pC=[normal2(:,181:239)];npC=size(normal_pC,2);

% Forming training set from data of 3 patients A, B and C

ntrain=size([normal_pA normal_pC normal_pB],2);
htrain=size([hypo_pA hypo_pC hypo_pB],2);
```

```
input = [normal_pA normal_pC normal_pB hypo_pA hypo_pC hypo_pB];
target = [-ones(1,ntrain) ones(1,htrain)];

% Forming testing set from data of 2 patients D and E

testpDE=[normal_pD normal_pE hypo_pD hypo_pE];

% Forming adaptive training set

retraindata=[add_datapD add_datapE input];
retraintarget =[-ones(1,size([add_datapD add_datapE],2)) target];
retraindata1=[add_datapD add_datapE];
retraintarget1 =[-ones(1,size([add_datapD add_datapE],2))];
retrainpDE=-ones(1,size([add_datapD add_datapE],2));


% Developing a feed-forward neural network

net = newff(input,target,9);

% Setting bounds for network parameters

bounds=ones(55,1)*[-3 3];



% Training the set of network parameters by Genetic algorithm

% Setting up GA parameters

% Initializing population

num=50;  % number of populations
numVars = size(bounds,1);       % Number of variables
rng = (bounds(:,2)-bounds(:,1))'; % The variable ranges'

xZomeLength = numVars+1;    %Length of string is numVar + fit
intpop          = zeros(num,xZomeLength);    %Allocate the new
population

intpop(:,1:numVars)=(ones(num,1)*rng).*(rand(num,numVars))+(ones(n
um,1)*bounds(:,1)');

for i=1:num
[intpop(i,:)intpop(i,xZomeLength)]=gaNNfit(intpop(i,:),net,input,t
arget);
end

opts = [1e-6 1 1];
```

```
% Default termination information

termOps=[2000];
termFN='maxGenTerm';

% Mutation operator

mutFNs=['nonUnifMutation'];
% mutOps=[2 25 3];
% mutFNs=['boundaryMutation multiNonUnifMutation nonUnifMutation
unifMutation'];
    mutOps=[4 0 0;6 termOps(1) 3;4 termOps(1) 3;4 0 0];


% Crossover operator

xOverFNs=['arithXover'];
% xOverFNs=['blendXover1'];
xOverOps=[2 0;2 3;2 0];

%  Selection operator

selectFN=['normGeomSelect'];
selectOps=[0.2];

xOverFNs=parse(xOverFNs);
mutFNs=parse(mutFNs);

startPop=intpop;
xZomeLength  = size(startPop,2);    %Length of the
xzome=numVars+fittness
numVar       = xZomeLength-1;       %Number of variables
popSize      = size(startPop,1);    %Number of individuals in the
population
endPop       = zeros(popSize,xZomeLength); %A secondary population
matrix
c1           = zeros(1,xZomeLength);    % An individual
c2           = zeros(1,xZomeLength);    % An individual
numXOvers    = size(xOverFNs,1);   % Number of Crossover operators
numMuts      = size(mutFNs,1);     % Number of Mutation operators
epsilon      = opts(1);     %  Threshold for two fitness to differ
oval         = max(startPop(:,xZomeLength)); % Best value in start
population
bFoundIn     = 1;           % Number of times best has changed
done         = 0;           % Done with simulated evolution
gen          = 1;           % Current Generation Number
collectTrace = (nargout>3); % Should we collect info every gen
floatGA      = opts(2)==1;  % Probabilistic application of ops
display      = opts(3);     % Display progress
```

```
while(~done)
    %Elitist Model
  [bval,bindx] = max(startPop(:,xZomeLength)); %Best of current
population
  best =  startPop(bindx,:);

  if collectTrace
    traceInfo(gen,1)=gen;                  %current generation
    traceInfo(gen,2)=startPop(bindx,xZomeLength);   %Best fittness
    traceInfo(gen,3)=mean(startPop(:,xZomeLength)); %Avg fittness
    traceInfo(gen,4)=std(startPop(:,xZomeLength));
  end

  if ( (abs(bval - oval)>epsilon) || (gen==1)) %If we have a new
best sol
    if display
      fprintf(1,'\n%d %f\n',gen,bval);        %Update the display
    end
      bPop(bFoundIn,:)=[gen startPop(bindx,:)];%Update bPop Matrix
    bFoundIn=bFoundIn+1;                %Update number of changes
    oval=bval;                          %Update the best val
  else
    if display
      fprintf(1,'%d ',gen);   %Otherwise just update num gen
    end
  end


% Running ga

  for i=1:numXOvers,
      for j=1:xOverOps(i,1),
    a = round(rand*(popSize-1)+1);  %Pick a parent
    b = round(rand*(popSize-1)+1);  %Pick another parent
    xN=deblank(xOverFNs(i,:));      %Get the name of crossover
function
    [c1 c2] = feval(xN,endPop(a,:),endPop(b,:),bounds,[gen
xOverOps(i,:)]);

    if c1(1:numVar)==endPop(a,(1:numVar))
      c1(xZomeLength)=endPop(a,xZomeLength); % Make sure we
created a new solution before evaluating
    elseif c1(1:numVar)==endPop(b,(1:numVar))
      c1(xZomeLength)=endPop(b,xZomeLength);
    else
%     [c1 c1(xZomeLength)]=' evalFN '(c1,[gen evalOps]);
      [c1 c1(xZomeLength)]=gaNNfit(c1,net,input,target);
    end
    if c2(1:numVar)==endPop(a,(1:numVar))
      c2(xZomeLength)=endPop(a,xZomeLength);
    elseif c2(1:numVar)==endPop(b,(1:numVar))
```

```
      c2(xZomeLength)=endPop(b,xZomeLength);
    else
%     [c2 c2(xZomeLength)]=' evalFN '(c2,[gen evalOps])
      [c2 c2(xZomeLength)]=gaNNfit(c2,net,input,target);
    end

    endPop(a,:)=c1;
    endPop(b,:)=c2;
      end
    end

    for i=1:numMuts,
      for j=1:mutOps(i,1),
    a = round(rand*(popSize-1)+1);
    c1 = feval(deblank(mutFNs(i,:)),endPop(a,:),bounds,[gen
mutOps(i,:)]);
    if c1(1:numVar)==endPop(a,(1:numVar))
      c1(xZomeLength)=endPop(a,xZomeLength);
    else
%     [c1 c1(xZomeLength)]=' evalFN '(c1,[gen evalOps]);
      [c1 c1(xZomeLength)]=gaNNfit(c1,net,input,target);
    end
    endPop(a,:)=c1;
      end
    end

    gen=gen+1;
    done=feval(termFN,[gen termOps],bPop,endPop); %See if the ga
is done
    startPop=endPop;            %Swap the populations
    [bval,bindx] = min(startPop(:,xZomeLength)); %Keep the best
solution
    startPop(bindx,:) = best;       %replace it with the worst
end

[bval,bindx] = max(startPop(:,xZomeLength));
if display
  fprintf(1,'\n%d %f\n',gen,bval);
end

x=startPop(bindx,:);
bPop(bFoundIn,:)=[gen startPop(bindx,:)];

if collectTrace
  traceInfo(gen,1)=gen;         %current generation
  traceInfo(gen,2)=startPop(bindx,xZomeLength); %Best fittness
  traceInfo(gen,3)=mean(startPop(:,xZomeLength)); %Avg fittness
end
```

```
% Updating neural network with the optimised parameters set
obtained by GA

net.IW{1,1}=reshape(best(1:36),9,4);
net.LW{2,1}=best(37:45);
net.b{1,1}=best(46:54)';
net.b{2,1}=best(55);


% Implementing the second step of local search by LM algorithm

% Setting parameters for training network by the Levenberg-
Marquardt algorithm and cross-validation technique

net.layers{1}.transferFcn='tansig';
net.layers(BCI2000).transferFcn='purelin';
%
net.performFcn='mse';
net.trainFcn='trainlm';
net.inputs{1}.processFcns={'removeconstantrows'};
net.outputs(BCI2000).processFcns={'removeconstantrows'};
net.dividefcn='dividerand';
net.trainParam.max_fail = 200;
net.trainParam.mu = 0.002;      %  Initialize Mu
net.trainParam.mu_dec=0.8 ;       %  Mu decrease factor
net.trainParam.mu_inc=1.2;       %  Mu increase factor
net.trainParam.min_grad=1e-100;
%
net.divideParam.trainRatio=3/4;
net.divideParam.valRatio=1/4;
net.divideParam.testRatio=0;

% Training network by the LM algorithm

net1=train(net,input,target);


% Updating neural network with the optimised parameters set
obtained by LM+GA algorithm

net2=train(net1,retraindata,retraintarget);

% Plotting ROC curve for the final optimised network

y1=sim(net2,input);
[X,Y,THRE,AUC,OPTROCPT,SUBY,SUBYNAMES] = perfcurve([-
ones(1,ntrain) ones(1,htrain)],[y1],'1');
plot(X,Y)
xlabel('False positive rate'); ylabel('True positive rate')
```

```
% Acquiring cut-off point of neural network based on ROC curve
index=find(Y>=0.80,1)+1;
Threshold=THRE(index);


% Estimating network performance

% Calculate classification results on the combined training-
validation set

sizedata=size(input,2);
a=0;a1=0;
b=0;b1=0;
%
for m=1:sizedata
        if m<=(sizedata-htrain)
            a=a+1;
            if y1(1,m)<=Threshold
                a1=a1+1;
            end
        end
        if m>(sizedata-htrain)
            b=b+1;
            if y1(1,m)>Threshold
                b1=b1+1;
            end
        end
end
%
spe_train=a1/a;
sen_train=b1/b;
acc_train=(a1+b1)/(a+b);


% Calculate classification results on the testing set of patients
D and E

l=0;l1=0;
h=0;h1=0;
%
y_pDE=sim(net1,testpDE);
%
for k=1:length(y_pDE)
        if k<=(length(y_pDE)-hpD-hpE)
            h=h+1;
            if y_pDE(1,k)<=Threshold
                h1=h1+1;
            end
        end
        if k>(length(y_pDE)-hpD-hpE)
            l=l+1;
```

```
            if y_pDE(1,k)>Threshold
                l1=l1+1;
            end
        end
end
%
spe_test_pDE=h1/h;
sen_test_pDE=l1/l;
acc_test_pDE=(h1+l1)/(h+l);
```

Bibliography

BCI2000, *User Tutorial:EEG Measurement Setup*, viewed 31 May 2013 <http://www.bci2000.org/wiki/index.php/User_Tutorial:EEG_Measurement_Setup>.

Belarbi, K. & Titel, F. 2000, 'Genetic algorithm for the design of a class of fuzzy controllers: an alternative approach', *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 4, pp. 398-405.

Bendtson, I., Gade, J., Rosenfalck, A.M., Thomsen, C.E., Wildschiødtz, G. & Binder, C. 1991, 'Nocturnal electroencephalogram registrations in Type 1 (insulin-dependent) diabetic patients with hypoglycaemia', *Diabetologia*, vol. 34, no. 10, pp. 750-756.

Bjørgaas, M., Sand, T. & Gimse, R. 1996, 'Quantitative EEG in Type 1 diabetic children with and without episodes of severe hypoglycemia: a controlled, blind study', *Acta Neurologica Scandinavica*, vol. 93, no. 6, pp. 398-402.

Bjørgaas, M., Sand, T., Vik, T. & Jorde, R. 1998, 'Quantitative EEG during controlled hypoglycaemia in diabetic and non-diabetic children', *Diabetic Medicine*, vol. 15, no. 1, pp. 30-37.

Blom, L., Dahlquist, G., Nystrom, L., Sandstrom, A. & Wall, S. 1989, 'The Swedish childhood diabetes study--social and perinatal determinants for diabetes in childhood', *Diabetologia*, vol. 32, no. 1, pp. 7-13.

Bluestone, J.A., Herold, K. & Eisenbarth, G. 2010, 'Genetics, pathogenesis and clinical interventions in type 1 diabetes', *Nature*, vol. 464, no. 7293, pp. 1293-1300.

Bode, B., Gross, K., Rikalo, N., Schwartz, S., Wahl, T., Page, C., Gross, T. & Mastrototaro, J. 2004, 'Alarms Based on Real-Time Sensor Glucose Values Alert Patients to Hypo- and Hyperglycemia: The Guardian Continuous Monitoring System', *Diabetes Technology & Therapeutics*, vol. 6, no. 2, pp. 105-113.

Borch-Johnsen, K., Joner, G., Mandrup-Poulsen, T., Christy, M., Zachau-Christiansen, B., Kastrup, K. & Nerup, J. 1984, 'Relation between breast-feeding and incidence rates of insulin-dependent diabetes mellitus. A hypothesis', *Lancet*, vol. 2, no. 8411, pp. 1083-1086.

Clarke, W., Jones, T., Rewers, A., Dunger, D. & Klingensmith, G.J. 2009, 'Assessment and management of hypoglycemia in children and adolescents with diabetes', *Pediatric Diabetes*, vol. 10, pp. 134-145.

Cole, B.T., Roy, S.H. & Nawab, S.H. 2011, 'Detecting freezing-of-gait during unscripted and unconstrained activity', *Conf Proc IEEE Eng Med Biol Soc*, vol. 2011, pp. 5649-5652.

Craig, D.A. & Nguyen, H.T. 2007, 'Adaptive EEG Thought Pattern Classifier for Advanced Wheelchair Control', *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2544-2547.

Cryer, P.E. 2007, 'Hypoglycemia, functional brain failure, and brain death', *The Journal of Clinical Investigation*, vol. 117, no. 4, pp. 868-870.

Cryer, P.E., Davis, S.N. & Shamoon, H. 2003, 'Hypoglycemia in diabetes', *Diabetes Care*, vol. 26, no. 6, pp. 1902-1912.

Daneman, D. 2006, 'Type 1 diabetes', *The Lancet*, vol. 367, no. 9513, pp. 847-858.

Davis, S. & Alonso, M.D. 2004, 'Hypoglycemia as a barrier to glycemic control', *J Diabetes Complications*, vol. 18, no. 1, pp. 60-68.

Dustman, R.E., Shearer, D.E. & Emmerson, R.Y. 1999, 'Life-span changes in EEG spectral amplitude, amplitude variability and mean frequency', *Clinical Neurophysiology*, vol. 110, no. 8, pp. 1399-1409.

Ercal, F., Chawla, A., Stoecker, W.V., Hsi-Chieh, L. & Moss, R.H. 1994, 'Neural network diagnosis of malignant melanoma from color images', *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 9, pp. 837-845.

Eshelman, L.J. & Schaffer, J.D. 1993, 'Real-coded genetic algorithms and interval-schemata', *Foundation of Genetic Algorithms 2*, ed. D. Whitley, Morgan Kaufmann., pp. 187-202<citeulike-article-id:1045422>.

Flood, T.M., Brink, S.J. & Gleason, R.E. 1982, 'Increased incidence of type I diabetes in children of older mothers', *Diabetes Care*, vol. 5, no. 6, pp. 571-573.

Group, D.R. 1991, 'Epidemiology of severe hypoglycemia in the diabetes control and complications trial', *The American Journal of Medicine*, vol. 90, no. 1, pp. 450-459.

Group, D.R. 1993, 'The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus', *New England Journal of Medicine*, vol. 329, no. 14, pp. 977-986.

Group, T.D.R.i.C.N.D.S. 2004, 'Accuracy of the GlucoWatch G2 Biographer and the Continuous Glucose Monitoring System During Hypoglycemia: Experience of the Diabetes Research in Children Network', *Diabetes Care*, vol. 27, no. 3, pp. 722-726.

Hagan, M.T. & Menhaj, M.B. 1994, 'Training feedforward networks with the Marquardt algorithm', *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989-993.

Handojoseno, A.M., Shine, J.M., Nguyen, T.N., Tran, Y., Lewis, S.J. & Nguyen, H.T. 2012, 'The detection of Freezing of Gait in Parkinson's disease patients using EEG signals

based on Wavelet decomposition', *Conf Proc IEEE Eng Med Biol Soc*, vol. 2012, pp. 69-72.

Henk-Jan Aanstoot, B.J.A., Denis Daneman, Thomas Danne, Kim Donaghue, Francine Kaufman, Rosangela R. Réa and Yasuko Uchigata 2007, 'Chapter One: Diabetes in children: epidemiology', *Pediatric Diabetes*, vol. 8, no. Supplement 8, pp. 10-18.

Howorka, K., Heger, G., Schabmann, A., Anderer, P., Tribl, G. & Zeitlhofer, J. 1996, 'Severe hypoglycaemia unawareness is associated with an early decrease in vigilance during hypoglycaemia', *Psychoneuroendocrinology*, vol. 21, no. 3, pp. 295-312.

Howorka, K., Pumprla, J., Saletu, B., Anderer, P., Krieger, M. & Schabmann, A. 2000, 'Decrease of vigilance assessed by EEG-mapping in Type I diabetic patients with history of recurrent severe hypoglycaemia', *Psychoneuroendocrinology*, vol. 25, no. 1, pp. 85-105.

Hyllienmark, L., Maltez, J., Dandenell, A., Ludvigsson, J. & Brismar, T. 2005, 'EEG abnormalities with and without relation to severe hypoglycaemia in adolescents with type 1 diabetes', *Diabetologia*, vol. 48, no. 3, pp. 412-419.

Jagielska, I., Matthews, C. & Whitfort, T. 1999, 'An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems', *Neurocomputing*, vol. 24, no. 1-3, pp. 37-54.

Juhl, C.B., Højlund, K., Elsborg, R., Poulsen, M.K., Selmar, P.E., Holst, J.J., Christiansen, C. & Beck-Nielsen, H., 'Automated detection of hypoglycemia-induced EEG changes recorded by subcutaneous electrodes in subjects with type 1 diabetes--The brain as a biosensor', *Diabetes Research and Clinical Practice*, vol. 88, no. 1, pp. 22-28.

Juhl, C.B., Højlund, K., Elsborg, R., Poulsen, M.K., Selmar, P.E., Holst, J.J., Christiansen, C. & Beck-Nielsen, H. 2010, 'Automated detection of hypoglycemia-induced EEG changes recorded by subcutaneous electrodes in subjects with type 1 diabetes--The brain as a biosensor', *Diabetes Research and Clinical Practice*, vol. 88, no. 1, pp. 22-28.

Karahaliou, A.N., Boniatis, I.S., Skiadopoulos, S.G., Sakellaropoulos, F.N., Arikidis, N.S., Likaki, E.A., Panayiotakis, G.S. & Costaridou, L.I. 2008, 'Breast Cancer Diagnosis: Analyzing Texture of Tissue Surrounding Microcalcifications', *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 6, pp. 731-738.

King, L.M., Nguyen, H.T. & Lal, S.K. 2006, 'Early driver fatigue detection from electroencephalography signals using artificial neural networks', *Conf Proc IEEE Eng Med Biol Soc*, vol. 1, pp. 2187-2190.

Klimesch, W. 1997, 'EEG-alpha rhythms and memory processes', *International Journal of Psychophysiology*, vol. 26, no. 1-3, pp. 319-340.

Klimesch, W. 1999, 'EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis', *Brain Research Reviews*, vol. 29, no. 2-3, pp. 169-195.

Klimesch, W., Doppelmayr, M., Schimke, H. & Pachinger, T. 1996, 'Alpha Frequency, Reaction Time, and the Speed of Processing Information', *Journal of Clinical Neurophysiology*, vol. 13, no. 6, pp. 511-518.

Klonoff, D.C. 2001, 'The need for hypoglycemia detection and prevention in Type 1 diabetes', *Diabetes Technology & Therapeutics*, vol. 3, no. 4, pp. 567-570.

Laione, F. & Marques, J. 2005, 'Methodology for hypoglycaemia detection based on the processing, analysis and classification of the electroencephalogram', *Medical and Biological Engineering and Computing*, vol. 43, no. 4, pp. 501-507.

Leung, F.H.F., Lam, H.K., Ling, S.H. & Tam, P.K.S. 2003, 'Tuning of the structure and parameters of a neural network using an improved genetic algorithm', *Neural Networks, IEEE Transactions on*, vol. 14, no. 1, pp. 79-88.

Luzheng, B., Xin-An, F. & Yili, L., 'EEG-Based Brain-Controlled Mobile Robots: A Survey', *Human-Machine Systems, IEEE Transactions on*, vol. 43, no. 2, pp. 161-176.

Mackay, D. 1992, 'A practical Bayesian framework for backpropagation networks', *Neural Computation*, vol. 4, pp. 448-472.

McCulloch, W. & Pitts, W. 1943, 'A logical calculus of the ideas immanent in nervous activity', *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115-133.

Montana, D.J. & Davis, L. 1989, 'Training feedforward neural networks using genetic algorithms', paper presented to the *Proceedings of the 11th International Joint Conference on Artificial Intelligence - Volume 1*, Detroit, Michigan.

Nguyen, A.V., Nguyen, L.B., Su, S. & Nguyen, H.T. 2012, 'Development of a Bayesian neural network to perform obstacle avoidance for an intelligent wheelchair', *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* San Diego, USA, pp. 1884-1887.

Nguyen, A.V., Su, S. & Nguyen, H.T. 2011, 'Development of a Bayesian recursive algorithm to find free-spaces for an intelligent wheelchair', *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, Massachusetts, USA, pp. 7250-7253.

Nguyen, H.T. 2008, 'Intelligent technologies for real-time biomedical engineering applications', *International Journal of Automation and Control*, vol. 2, no. 2, pp. 274-285.

Nguyen, H.T., Ghevondian, N. & Jones, T.W. 2009, 'Real-time detection of nocturnal hypoglycemic episodes using a novel non-invasive hypoglycemia monitor', *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3822-3825.

Nguyen, L.B., Nguyen, A.V., Sai Ho, L. & Nguyen, H.T. 2013a, 'Analyzing EEG signals under insulin-induced hypoglycemia in type 1 diabetes patients', *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* Osaka, Japan, pp. 1980-1983.

Nguyen, L.B., Nguyen, A.V., Sai Ho, L. & Nguyen, H.T. 2013b, 'Combining genetic algorithm and Levenberg-Marquardt algorithm in training neural network for hypoglycemia detection using EEG signals', *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Osaka, Japan, pp. 5386-5389.

Niedermeyer, E. 1999, 'The Normal EEG of the Waking Adult', in, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, Lippincott Williams & Wilkins, Baltimore MD, pp. 149-173.

Organization, W.H. 2012, *Diabetes*, viewed February 2013 <http://www.who.int/mediacentre/factsheets/fs312/en/index.html>.

Patterson, C.C., Carson, D.J., Hadden, D.R., Waugh, N.R. & Cole, S.K. 1994, 'A case-control investigation of perinatal risk factors for childhood IDDM in Northern Ireland and Scotland', *Diabetes Care*, vol. 17, no. 5, pp. 376-381.

Pedersen-Bjergaard, U., Pramming, S., Heller, S.R., Wallace, T.M., Rasmussen, Å.K., Jørgensen, H.V., Matthews, D.R., Hougaard, P. & Thorsteinsson, B. 2004, 'Severe hypoglycaemia in 1076 adult patients with type 1 diabetes: influence of risk markers and selection', *Diabetes/Metabolism Research and Reviews*, vol. 20, no. 6, pp. 479-486.

Penny, W.D. & Roberts, S.J. 1999, 'Bayesian neural networks for classification: how useful is the evidence framework?', *Neural Networks*, vol. 12, no. 6, pp. 877-892.

Poretsky, L. & Ali, O. 2010, 'Type 1 Diabetes Mellitus: Epidemiology, Genetics, Pathogenesis, and Clinical Manifestations', in, *Principles of Diabetes Mellitus*, Springer US, pp. 181-201.

Pramming, S., Thorsteinsson, B., Stigsby, B. & Binder, C. 1988, 'Glycaemic threshold for changes in electroencephalograms during hypoglycaemia in patients with insulin dependent diabetes', *British Medical Journal (Clinical research ed.)*, vol. 296, no. 6623, pp. 665-667.

Richard Clark, C., Veltmeyer, M.D., Hamilton, R.J., Simms, E., Paul, R., Hermens, D. & Gordon, E. 2004, 'Spontaneous alpha peak frequency predicts working memory performance across the age span', *Int J Psychophysiol*, vol. 53, no. 1, pp. 1-9.

Rosenblatt, F. 1958, 'The perceptron: a probabilistic model for information storage and organization in the brain', *Psychol Rev*, vol. 65, no. 6, pp. 386-408.

Rosenbloom, A.L. & Hanas, R. 1996, 'Diabetic ketoacidosis (DKA): treatment guidelines', *Clin Pediatr*, vol. 35, no. 5, pp. 261-266.

Rowan, A.J. & Tolunsky, E. 2003, *Primer of EEG : With a Mini-atlas*, Elsevier, Philadelphia, PA, USA.

Sadasivan, P.K. & Narayana Dutt, D. 1994, 'Minimization of EOG artefacts from corrupted eeg signals using a neural network approach', *Computers in Biology and Medicine*, vol. 24, no. 6, pp. 441-449.

Setnes, M. & Roubos, H. 2000, 'GA-fuzzy modeling and classification: complexity and performance', *IEEE Transactions on Fuzzy Systems*, vol. 8, no. 5, pp. 509-522.

Shi, L.C. & Lu, B.L. 2008, 'Dynamic clustering for vigilance analysis based on EEG', *Conf Proc IEEE Eng Med Biol Soc*, p. 4649089.

Sigurdsson, S., Philipsen, P.A., Hansen, L.K., Larsen, J., Gniadecka, M. & Wulf, H.C. 2004, 'Detection of skin cancer by classification of Raman spectra', *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 10, pp. 1784-1793.

Skladnev, V.N., Ghevondian, N., Tarnavskii, S., Paramalingam, N. & Jones, T.W. 2010, 'Clinical evaluation of a noninvasive alarm system for nocturnal hypoglycemia', *J Diabetes Sci Technol*, vol. 4, no. 1, pp. 67-74.

Smith, S.J.M. 2005, 'EEG in the diagnosis, classification, and management of patients with epilepsy', *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. suppl 2, pp. ii2-ii7.

Sovik, O. & Thordarson, H. 1999, 'Dead-in-bed syndrome in young diabetic patients', *Diabetes Care*, vol. 22, no. 2, pp. B40-42.

Strachan, M.W.J. 2007, 'Frequency, Causes and Risk Factors for Hypoglycaemia in Type 1 Diabetes', in, *Hypoglycaemia in Clinical Diabetes*, John Wiley & Sons, Ltd, pp. 49-81.

Stroganova, T.A., Orekhova, E.V. & Posikera, I.N. 1999, 'EEG alpha rhythm in infants', *Clinical Neurophysiology*, vol. 110, no. 6, pp. 997-1012.

Subasi, A. & Kiymik, M.K. 2010, 'Muscle fatigue detection in EMG using time-frequency methods, ICA and neural networks', *J Med Syst*, vol. 34, no. 4, pp. 777-785.

Tallroth, G., Lindgren, M., Stenberg, G., Rosen, I. & Agardh, C.D. 1990, 'Neurophysiological changes during insulin-induced hypoglycaemia and in the recovery period following glucose infusion in Type 1 (insulin-dependent) diabetes mellitus and in normal man', *Diabetologia*, vol. 33, no. 5, pp. 319-323.

Towle, V.L., Bolanos, J., Suarez, D., Tan, K., Grzeszczuk, R., Levin, D.N., Cakmur, R., Frank, S.A. & Spire, J.-P. 1993, 'The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy', *Electroencephalography and Clinical Neurophysiology*, vol. 86, no. 1, pp. 1-6.

Tribl, G., Howorka, K., Heger, G., Anderer, P., Thoma, H. & Zeitlhofer, J. 1996, 'EEG Topography during Insulin-Induced Hypoglycemia in Patients with Insulin-Dependent Diabetes mellitus', *European Neurology*, vol. 36, no. 5, pp. 303-309.

Warren, R.E. & Frier, B.M. 2005, 'Hypoglycaemia and cognitive function', *Diabetes, Obesity and Metabolism*, vol. 7, no. 5, pp. 493-503.

Welfare, A.I.o.H.a. 2011, *Incidence of insulin-treated diabetes in Australia 2000–2009*, viewed February 2013 <http://www.aihw.gov.au/diabetes/incidence/#t1>.

Widrow, B. & Hoff, M. 1960, 'Adaptive Switching Circuits', *1960 {IRE} {WESCON} Convention Record, Part 4*, {IRE}, pp. 96-104<citeulike-article-id:2510415 http://isl-www.stanford.edu/~widrow/papers/c1960adaptiveswitching.pdf>.

Widrow, B. & Lehr, M.A. 1990, '30 years of adaptive neural networks: perceptron, Madaline, and backpropagation', *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1415-1442.

Wolpert, H.A. 2007, 'Use of continuous glucose monitoring in the detection and prevention of hypoglycemia', *J Diabetes Sci Technol*, vol. 1, no. 1, pp. 146-150.

Woten, D.A., Lusth, J. & El-Shenawee, M. 2007, 'Interpreting Artificial Neural Networks for Microwave Detection of Breast Cancer', *Microwave and Wireless Components Letters, IEEE*, vol. 17, no. 12, pp. 825-827.

Yach, D., Hawkes, C, Gould, CL, Hofman, KJ 2004, 'The global burden of chronic diseases: Overcoming impediments to prevention and control', *JAMA*, vol. 291, no. 21, pp. 2616-2622.