NON-INVASIVE DETECTION OF HYPERGLYCAEMIA IN TYPE 1 DIABETIC PATIENTS USING ELECTROCARDIOGRAPHIC SIGNALS

By

Linh Lan Nguyen

A thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy

University of Technology Sydney

Faculty of Engineering and Information Technology

2014

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

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

I also certify that this 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

Production Note:
Signature removed prior to publication.

This thesis is especially dedicated to my dearest family

for their endless love, care and encouragement …

Acknowledgements

Firstly of all, I would like to thank Vietnamese Ministry of Education and Training (MOET) and University of Technology Sydney (UTS) for granting me a University of Technology Sydney – Vietnam International Education Development scholarship (UTS-VIED) without which I would not be able to carry out this research. I also would like to thank Hanoi University of Science and Technology for giving me an opportunity to attend the PhD program.

My great appreciation goes to my principal supervisor Professor Hung Nguyen for his great intellectual guidance and supervision during my PhD candidature. He gave me the opportunity to perform research work in a very practical and interesting scientific area. His advice and support have contributed enormously to the quality of my research and have made my PhD challenging, inspiring and enjoyable.

I would like to extend my gratitude to my co-supervisor Dr Steven Su for his consultation and encouragement. I am also extremely grateful to Lesley Nguyen for her enthusiasm in proofreading and correcting my thesis. I would like to sincerely thank Dr Nejhdeh Ghevondian for his thoughtful advice and ideas throughout the time I carried out this research.

I appreciate all my colleagues and staff in Centre for Health Technologies who supported me in my study with encouragement and confidence during the past few years.

Last but not least, I would like to express my heartfelt gratitude to my parents, my brother Hoang and especially my husband Dai and my son Phong for their great love, understanding and support over the years that enabled me to complete this thesis.

Table of contents

List of figures

List of tables

List of Abbreviations

ANN	:	Artificial neural network
AUC	:	Area under the curve
BGL	:	Blood glucose level
CGF	:	Conjugate gradient back propagation with Fletcher-Reeves updates
CGMS	:	Continuous glucose monitoring system
ECG	:	Electrocardiography
FFT	:	Fast Fourier transform
GA	:	Genetic algorithm
GDX	:	Gradient descent with momentum
gm	:	Geometric mean
HRV	:	Heart rate variability
KNN	:	K-nearest neighbour
LDA	:	Linear Discriminant Analysis
LM	:	Levenberg Marquardt
LM-NN	:	Neural network using Levenberg Marquardt algorithm
MSE	:	Mean squared error
NIR	:	Near-infrared light
NN	:	Neural network
PC	:	Principle component
PCA	:	Principal component analysis
PCA-CGF	:	Neural network using principal component analysis and Conjugate gradient backpropagation with Fletcher-Reeves updates algorithm
PCA-GDX	:	Neural network using principal component analysis and Gradient descent with momentum algorithm
PCA-LM	:	Neural network using principal component analysis and Levenberg Marquardt algorithm
PCA-RP	:	Neural network using principal component analysis and Resilient backpropagation algorithm
PCA-SCG	:	Neural network using principal component analysis and Scaled conjugate gradient algorithm
PSO	:	Particle swarm optimisation

PSO-NN	:	Neural network using particle swarm optimisation
ROC	:	Receiver operating characteristic
RP	:	Resilient backpropagation
SCG	:	Scaled conjugate gradient
Sen	:	Sensitivity
Spec	:	Specificity
T1DM	:	Type 1 Diabetes mellitus
VIF	:	Variance inflation factor

List of Symbols

Q	:	the beginning of QRS complex
R	:	the peak of QRS complex
S	:	the end of QRS complex
P	:	the beginning of P-wave
T_P	:	the peak of T-wave
T_E	:	the end of T-wave
RR	:	the interval between two consecutive R points
PR	:	the interval from the beginning of P-wave to the beginning of QRS complex
QT	:	the interval from Q point to the end of T-wave (T_E)
QT_C	:	the corrected QT interval by heart rate (Bazett's formula)
RT	:	the interval from R point to the end of T-wave (T_E)
RT_C	:	the corrected RT interval by heart rate (Bazett's formula)
T_PT_E	:	the interval from the peak of T-wave (T_P) to the end of T-wave (T_E)
T_PT_{EC}	:	the corrected T_PT_E interval by heart rate (Bazett's formula)
$MeanRR$	:	mean RR interval
$SDNN$	:	the standard deviation of the RR interval index
$RMSSD$	:	the root mean square of successive RR interval differences
$pNN50$	:	the percentage of beats with a consecutive RR interval more than 50 ms
$HRVi$	:	HRV triangular index
$TINN$	:	baseline width of the RR interval histogram evaluated through triangular interpolation
VLF	:	Very low frequency
LF	:	Low frequency
HF	:	High frequency
$TotalPw$	:	Total spectral power
LF/HF	:	ratio between LF and components
p_{best}	:	best solution of the particle has achieved so far
g_{best}	:	the global best
$lbest$	:	the local best
w_i	:	position of the particle
v_i	:	velocity of the particle

c_1, c_2	:	acceleration constants
ω	:	inertia factor
ω_{max}	:	initial weight
ω_{min}	:	final weight
v_{max}	:	maximum allowable velocity
$v_{max\text{-}perc}$	:	velocity clamping percentage

Abstract

Hyperglycaemia is the medical term for a state caused by a high level of blood glucose, resulting from defects in insulin secretion, insulin action, or both. Hyperglycaemia is a common dangerous complication to glycaemic control in Type 1 diabetic patients. The chronic hyperglycaemia of diabetes is associated with long-term damage, dysfunction, and failure of different organs, especially the eyes, kidneys, nerves, heart, and blood vessels. Therefore, reliable detection of hyperglycaemic episodes is important in order to avoid major health conditions.

Conventionally, diabetic patients need to frequently monitor blood glucose levels to determine whether they have hyperglycaemia or not. A patient has to prick their finger (finger-stick) for a drop of blood several times a day, which can therefore significantly discourage many patients from periodically checking blood glucose levels. Another choice for hyperglycaemia detection might be continuous glucose monitoring systems (CGMS), which measure the glucose level in the interstitial fluid. For patients using CGMS, finger-sticks are still required to calibrate the sensor. The main shortcoming of CGMS is that glucose levels in interstitial fluid lag temporally behind blood glucose values, normally 10-15 minutes, which absolutely limits the accuracy of the detection. There is a strong demand to have a non-invasive technique to help patients to diagnose the disease easily and painlessly. Few methods have been reported to detect hyperglycaemia non-invasively or minimally invasively such as in exhaled methyl nitrates, and early detection of ongoing β cell death. However, the purpose of these studies was on real-time glucose control rather than disease diagnosis.

Electrocardiography (ECG) is a broadly used technique to obtain a quick, non-invasive clinical and research screen for diagnosing abnormal rhythms of the heart caused by diseases. In fact, observations of ECG changes have been found in hypoglycaemia and hyperglycaemia states in T1DM, such as increased heart rate and prolongation of *QT* interval in hypoglycaemia, whereas hyperglycaemia was related to reduced heart rate variability. By using these findings in hypoglycaemia,

researchers have developed an effective and sensitive system to detect hypoglycaemia non-invasively. These excellent performances of hypoglycaemia detection using ECG is the motivation of this thesis to study the effect of hyperglycaemia on ECG signals, and based on the findings to exploit the computational intelligence on the non-invasive detection of hyperglycaemia.

This research firstly explores the changes of ECG parameters associated with the hyperglycaemic state in T1DM. The ECG parameters consist of ECG intervals relating to repolarisation phase and heart rate variability (HRV) measures. A clinical study of ten T1DM patients and ECG feature extraction process are conducted to collect ECG features. Statistical analysis is then applied to every ECG feature to estimate the significant difference between hyperglycaemic and normoglycaemic states. The results show that the selected ECG parameters in hyperglycaemia differ significantly from those in normoglycaemia ($p< 0.05$). It implies that certain ECG parameters are correlated with high blood glucose levels and they possibly contribute to the performance of hyperglycaemia detection. Thus, the ECG parameters are used for input data of hyperglycaemia classifiers in this thesis.

Furthermore, the thesis introduces novel computational intelligent methods for hyperglycaemia detection using the ECG parameters. A neural network using Levenberg-Marquardt algorithm is the first method explored for hyperglycaemia detection in this thesis, known as LM-NN. The second algorithm is the integration of principal component analysis (PCA) with a neural network utilising the Levenberg-Marquardt algorithm, which is called a PCA-LM-NN network. PCA is a useful tool for dimensionality reduction to diminish the computational requirement and overcome the problem of multicollinearity. It is employed to filter the data so that only the significant independent ECG variables responsible for the high blood glucose levels can be used as input for the network training, in order that the neural network performs well for hyperglycaemia detection. The third method is for the improvement of the second method where particle swarm optimization is included. This algorithm is a combination of PCA, PSO and neural network, which is called PSO-NN. The PSO is utilised as an effective training algorithm to optimise the weights of the neural network. The proposed methods are compared with each other

and with other traditional classifiers. All the algorithms are investigated with the clinical electrocardiographic data extracted from ten T1DM patients.

The results show that the performance of PCA-LM model for hyperglycaemia detection is better than that of LM-NN (70.88% vs. 67.94%, in terms of geometric mean). In addition, the PSO-NN outperforms the PCA-LM-NN (77.58% vs. 70.88%, in terms of geometric mean). In short, the PSO-NN significantly improves the performances of both the LM-NN and PCA-LM-NN, with considerable sensitivity, specificity and geometric mean of 82.35%, 73.08% and 77.58%, respectively.

CHAPTER 1 . INTRODUCTION

Diabetes mellitus is a globally challenging problem, which is responsible for an enormous public health and social burden. The number of people living with diabetes is increasing in every country and half of these people do not know they have it. Meanwhile, hyperglycaemia, the most common complication of diabetes, has been recognised as detrimental and one of the risk factors for adverse outcomes. Achieving normoglycaemia can be difficult due to variable insulin protocols without consensus. There is still limited literature related to hyperglycaemia detection. Therefore, adequate detection of hyperglycaemia is absolutely essential for diabetic patients to avoid circulation problems and other diseases caused by diabetes such as kidney failure, heart disease, etc.

Chapter 1 begins with the background of this research to provide a brief introduction about the meaning of hyperglycaemia and a general understanding of the key topics discussed throughout the rest of the thesis. The second section explains the motivation for conducting this research and the necessity for a novel non-invasive detection of hyperglycaemia. Continuously, the third section introduces the aim of the thesis. The next two parts present the contribution and structure of the thesis. Finally, publications relating to this thesis during the PhD candidature are provided in the last section.

1.1 BACKGROUND

Diabetes mellitus is a collection of diseases characterised by insufficient insulin production and/or insulin resistance. Insulin is a hormone that mediates the uptake of glucose in liver, muscle and fat tissue. The lack of insulin or insulin resistance results in elevated blood glucose concentration (BG), hyperglycaemia, which is the cardinal symptom of diabetes. Over time, high blood glucose damages nerves and blood vessels, leading to complications such as heart disease, stroke, kidney disease, blindness, dental disease, and amputations. The alternative condition in diabetes is hypoglycaemia, an opposite state resulting from too low blood glucose levels. Severe hypoglycaemia can lead to seizures, coma, and even death.

The two major types of diabetes are type 1 (T1DM) and type 2 (T2DM) constituting approximately 10% and below 90% of the total diabetes population, respectively. A small proportion (3-5%) of pregnant women develop gestational diabetes (GDM) that resembles T2DM in manifestation and aetiology (Ben-Haroush *et al.*, 2004). GDM along with other types of the disease including pre-diabetes will not be discussed further in this report. In T1DM, a progressive destruction of the insulin producing beta cells in the pancreas causes an absolute insulin deficiency (Committee, 2003). Therefore, people with T1DM are dependent on daily insulin injections. Without exogenous insulin people with T1DM will die from ketoacidosis within a short time. Type 1 diabetes typically occurs in children and young adults, though it can appear at any age. In the past, type 1diabetes was called juvenile diabetes or insulin-dependent diabetes mellitus.

T2DM, the most common type of diabetes, is characterised by impaired insulin sensitivity and secretion. Unlike T1DM, persons with T2DM retain a certain production of insulin, although of an insufficient level to keep the BG within normal range. T2DM progresses slowly and as a result people can have the disease for years without knowing it. This form of diabetes can be controlled by a properly managed diet and exercise, by taking oral anti-diabetes drugs; or by insulin therapy (Joslin *et al.*, 2005). Hyperglycaemia is the technical term for high blood glucose (blood sugar). High blood glucose happens when the body has too little insulin or when the body

cannot use insulin properly. People who have type 1 and have not have given themselves enough insulin may experience hyperglycaemia symptoms. Other causes for this condition include eating too much food, infection or illness, injury, surgery and emotional stress. The common symptoms of this condition are frequent urination, thirst, drowsiness, fatigue, excessive hunger, dry mouth, blurry vision and unexplained weight loss. If people do not treat it properly, hyperglycaemia can also lead to more serious conditions, including ketoacidosis, mostly in people with type 1.

Hyperglycaemia witnesses worse effects in conjunction with severe illness. Hyperglycaemia was reported as being present in up to 68% of patients admitted to a medical ICU (Intensive Care Unit) (Freire *et al.*, 2005). It is an independent predictor of death in many acute settings, including acute myocardial infarction (Capes *et al.*, 2000), trauma, head injury, and stoke (Mowery *et al.*, 2009; Prakash *et al.*, 2008). Postulated mechanisms by which hyperglycaemia causes harm include decreased cerebral blood flow, intracellular acidosis, and low ATP levels; these may be similar to the actions of hyperglycaemia witnessed in diabetes mellitus (Brownlee). A raised blood glucose level is also a risk factor for infection and is associated with increased mortality in acute illness (Lin *et al.*, 2009). Furthermore, hyperglycaemia is associated with poor gut motility, a factor that may be important in bacterial overgrowth and translocation (Brealey *et al.*, 2009). In general, hyperglycaemia causes harm through a variety of mechanisms, and this damage is accentuated in the critically ill where there is concurrent activation of multiple inflammatory processes. There is broad consensus that hyperglycaemia should be avoided, and predicted as early as possible before optimal treatment end points can be clarified.

In contrast, hypoglycaemia, the most acute and common complication of type 1 diabetes, is a condition characterized by abnormally low blood glucose. Extreme hypoglycaemia is suggested as a factor of death in diabetic patients. Hypoglycaemia of a type 1 diabetic patient, which often happens during sleep, is associated with the cause of the “dead in bed” syndrome (Tattersall *et al.*, 1991). To prevent or minimize the consequences of hypoglycaemia, a lot of research has been done to obtain comprehensive understanding of hypoglycaemia mechanisms and to develop early detection systems of hypoglycaemia based on physiological signals. For example,

hypoglycaemia has been found to alter results in cardiac repolarisation such as significant prolonged *QT* interval (Eckert *et al.*, 1998; Christensen *et al.*, 2010; Robinson *et al.*, 2004). The patho-physiological mechanism of changes in *QT* interval during hypoglycaemia has been clarified by a physiologically longer cardiac repolarisation at night and thus an electrically instable myocardium worsened by the hypoglycaemia induced adrenergic stimulation (Cryer, 2004). Non-invasive and minimally invasive methods of hypoglycaemia detection have been covered by several patents (Hung Nguyen, 2008; Bharmi, 2009; Evans, 2006). Many researches have also been undertaken to improve the performance of hypoglycaemia detection (Alexakis *et al.*, 2006; Ling *et al.*, 2011; Nuryani *et al.*, 2012). Hence, there is no need for further explorations regarding perception in hypoglycaemia detection but hyperglycaemia.

In clinical settings, the best screening test for high blood glucose is the fasting plasma glucose (FPG), which is also a component of diagnostic testing. The FPG test and the 75-g oral glucose tolerant test (OGTT) are both suitable tests for diabetes; however, the FPG test is preferred because it is easier and acceptable to patients, and less expensive. If there is a high suspicion for hyperglycaemia, an OGTT should be performed to confirm on alternative days. Fasting is defined as no consumption of food or beverage other than water for at least 8h before testing. If necessary, plasma glucose testing may be required on individuals who have taken food or drink shortly before testing. A confirmatory FPG test or OGTT should be completed on a different day if the clinical condition of the patient permits. Normally, it takes quite a lot of time for the completion of clinical diagnosis of hyperglycaemia in diabetes.

Intermittent measurement of blood glucose for self-monitoring is currently one of the main obstacles in achieving glycaemic control in Type 1 diabetic patients. Conventionally, to determine whether they are hyperglycaemic or not, diabetic patients need to frequently monitor blood glucose levels. One conventional self-monitoring technique of glucose concentration, for example, requires that the patients draw blood, typically by pricking the finger. The drawn blood is then analysed by a portable device to determine blood glucose levels. The technique can

be painful and therefore can significantly discourage the patent from periodically checking blood glucose levels. Finger-stick measurements can only provide information regarding a few minutes of each day and cannot be used to easily identify patterns of abnormal glycaemia (Kaufman *et al.*, 2001). There is a strong demand to have non-invasive or minimally invasive techniques to help patients to diagnose the disease easily and painlessly.

Non-invasive up to date methods proposed include systems such as: infrared/near-infrared spectroscopy (Malin *et al.*, 1999), iontophoresis (Tamada *et al.*, 1999), skin conductance (Rebrin *et al.*, 1999), etc. However, none of these have proved sufficiently reliable or unobtrusive. The ideal solution to this problem would be the development of a subcutaneous sensor, which could monitor glucose concentrations continuously. Even though recent technological advances offer the development of such a sensor, there are still significant technological problems to overcome (Ireland *et al.*, 2000). Continuous blood glucose monitors measure the glucose level in the interstitial fluid. Shortcomings of continuous glucose monitoring systems (CGMS) due to this fact are that continuous systems must be calibrated with a traditional blood glucose measurement and therefore require both the CGM system and occasional finger-stick measurement. Another limitation is that glucose levels in interstitial fluid lag temporally behind blood glucose values. With glucose fluctuations, it has been shown that there is a 5-min lag time with the CGMS which may vary up to 26 minutes in some systems. This lag is due to the delay in equilibration between blood and interstitial glucose, which limits the accuracy of CGM for predicting blood glucose concentrations especially when these concentrations are changing rapidly (Burge *et al.*, 2008). In summary, it is still more desirable to have an alternative non-invasive method to detect hyperglycaemia efficiently.

There are a few methods to detect hyperglycaemia non-invasively or minimally invasively such as in exhaled methyl nitrates (Galassetti, 2007), early detection of ongoing β cell death (Akirav *et al.*, 2011). To predict hyperglycaemia, other studies have focused on predicting blood glucose levels to forecast glycaemic states using dynamic models (Sparacino *et al.*, 2008; Le Compte *et al.*, 2010).

Alternatively, neural network approaches have been developed for early detection of hypoglycaemia/hyperglycaemia events based on glucose data from continuous glucose monitoring (CGM) (Daskalaki *et al.*, 2012; Pappada *et al.*, 2010). Nevertheless, the purpose of these studies was on real-time glucose control rather than disease diagnosis.

According to a variety of medical and healthcare information, electrocardiography (ECG) is a broadly used technique to obtain information ranging from a simple pulse to a very detailed description of heart activity. It indicates the overall rhythm of the heart, so ECG is known to be the best way to measure and diagnose abnormal rhythms of the heart (Bonow *et al.*, 2011). In ECG, during every heartbeat, *P, Q, R, S, T* and *U* waves can be seen. These waveforms result from depolarization and repolarisation of different parts of heart muscle. By evaluation and prognosis, ECG parameters can give early predictions of many heart diseases with a high rate of accuracy for patients. In fact, observations of ECG changes have been found in hypoglycaemia and hyperglycaemia states in T1DM, such as increased heart rate and prolongation of *QT* interval in hypoglycaemia (Christensen *et al.*, 2010), whereas hyperglycaemia has been related to reduced heart rate variability (Singh *et al.*, 2000). By using these findings in hypoglycaemia, researchers at the Centre for Health Technologies in University of Technology Sydney have developed an effective and sensitive system to monitor hypoglycaemia non-invasively (Nguyen *et al.*, 2009; Nguyen *et al.*, 2006). This hypoglycaemia monitor called Hypomon enables the identification of approximately 80% of night-time hypoglycaemic events and has been commercialised by AiMedics Pty Ltd. Based on this excellent performance of hypoglycaemia detection, ECG becomes a very promising technique to obtain more ubiquitous and non-invasive clinical and research screening for detection of hyperglycaemia than other physiological signals. To date, there has been no research into hyperglycaemia detection in T1DM based on changes in ECG.

1.2 MOTIVATION OF THE THESIS

With the idea of employing ECG to detect hyperglycaemia, comprehension of the effects of hyperglycaemia on ECG signals should be firstly investigated. Several

studies have been reported regarding ECG abnormalities which can possibly be found in hyperglycaemia. For example, shorter mean *RR* intervals were found in hyperglycaemic patients (Bathula *et al.*, 2010). Increased *PR* in hyperglycaemia have been discussed in a research with acute hyperglycaemia in healthy males (Marfella *et al.*, 2000). Besides, while hypoglycaemia was widely proved to alter ventricular repolarisation by lengthening of QT_C (corrected *QT* interval corrected by Bazett's formula), the role of hyperglycaemia in causing abnormal cardiac repolarisation in people with diabetes is not yet clear. Marfella reported significant increments of QT_C interval and *PR* interval in cases of acute hyperglycaemia in normal subjects (Marfella *et al.*, 2000). Gordin suggested that a hyperglycaemic glucose clamp induces QT_C interval prolongation in diabetic patients and in healthy control subjects (Gordin *et al.*, 2008). Inversely, in a recent study (Suys *et al.*, 2006) where simultaneously recorded QT_C values and glucose levels were analysed in patients with Type 1 Diabetes, QT_C prolongation was associated only with hypoglycaemia, not with hyperglycaemia. Results from clinical studies are less consistent; therefore further studies are needed in order to establish the clinical relevance of the relationship between hyperglycaemia and *QT* interval.

As mentioned above, hyperglycaemia may alter results in a corrected *QT* interval which are normally used as an index of ventricular repolarisation duration. *QT* interval covers the distance from the beginning of *Q*-wave to the end of *T*-wave. Therefore, there is no doubt that alteration in a QT_C interval may influence other ventricular repolarisation parameters such as corrected *RT* interval (interval from *R* point to peak of *T*-wave corrected by Bazett's formula) and corrected T_PT_E interval (depicted as interval from the peak to the end of *T*-wave corrected by Bazett's formula). So far, to the best of my knowledge, there is no report about the effect of hyperglycaemia on RT_C and T_PT_{EC} intervals. In addition, research regarding the relation of ventricular repolarisation parameters on the outcome of hyperglycaemia detection has not been explored.

Recently, reduced heart rate variability (HRV) has been seen among patients with Type 1 diabetic mellitus (Singh *et al.*, 2000; Jaiswal *et al.*, 2013). Diabetes mellitus is widely known as one of the main causes of autonomic neuropathy, which

can lead to abnormalities in heart rate control (Kudat *et al.*, 2006) and the main driver to the abnormality in heart rate variability appears to be hyperglycaemia. However, changes in HRV parameters have not yet been investigated for detection of hyperglycaemia. Thus, a question arises as to the contribution of HRV parameters in the performance of hyperglycaemia detection.

Other problems being considered are the methodology for the hyperglycaemia detection system and which computational intelligences are chosen to detect hyperglycaemia accurately and efficiently. Hung Nguyen *et al.* have suggested a lot of algorithms to detect the onset of hypoglycaemia using ECG parameters (Ling *et al.*, 2011; Nuryani *et al.*, 2012; Nguyen *et al.*, 2009). These studies have been reported with good sensitivities and with acceptable specificities. However, they have not yet been employed to detect hyperglycaemia. Among advanced computational intelligent algorithms, artificial neural networks (ANN) have been proved to be a powerful tool for classification models. Neural networks with the ability to learn by example achieve flexibility and effectiveness in medical diagnosis. Because of its advantages, ANN is selected as a classification technique for this research. Further investigation is going to be undertaken in a combination system which employs ANN. There is motivation to examine the contribution of a combined ANN system to the effectiveness of hyperglycaemia detection.

1.3 AIMS OF THESIS

According to the motivation statements above, this PhD research aims to introduce a novel non-invasive hyperglycaemia detection system for diabetic patients utilising computational intelligent algorithms. The classification techniques use electrocardiographic parameters as input signals. Hence, two main purposes are investigated in this study.

The first aim is to explore the changes of ECG parameters associated with the hyperglycaemic state in T1DM. The ECG parameters consist of ECG intervals and heart rate variability (HRV) measures. The objective of this investigation is to identify hyperglycaemic and normoglycaemic states based on ECG changes in

diabetic patients. To do that, statistical analysis is applied to every ECG feature to estimate the significant difference between hyperglycaemic and normoglycaemic states. Moreover, correlation analysis is performed to evaluate the relationship between ECG parameters and blood glucose levels. Then, the identification is used to find the potential ECG features which possibly contribute to the performance of hyperglycaemia detection.

The second aim is to introduce several computational intelligent methods for hyperglycaemia detection. The algorithms are arranged from the simple to advanced ones in order to compare the improvement of results. At the heart of the methodology in this research is the development of an artificial neural network (ANN) for hyperglycaemia detection. A basic approach of neural network is presented as the second algorithm. Five training methods are examined to choose the appropriate training type for the later hyperglycaemia detection. Following the idea of neural network development, an advanced two-stage ANN classifier is established with a combination of principal component analysis (PCA) to enhance the performance of detection. PCA is employed as the first stage to transform the original group of ECG parameters into a new appropriately lower dimensional size of variables to diminish the computational requirement and overcome the problem of multicollinearity. The objective of PCA focused especially on identification of the most important ECG factors among cardiac intervals and HRV measures that are responsible for the variation of high blood glucose levels in diabetic patients. The last method is the improvement of the third one which is the hybrid of the two-stage ANN classifier using PCA and particle swarm optimisation (PSO). PSO is adopted as the training method for ANN classifier to optimise the neural network parameters. PSO technique could be an effective alternative training algorithm for ANN since it is found to be quite accurate when compared to the existing conventional algorithms. Besides this, since PSO is a heuristic optimisation technique, convergence to a local minimum is avoided and instead a global optimal solution can be obtained (Ch *et al.*, 2012) so that performance of ANN could be improved.

A clinical study in ten T1DM patients is carried out to investigate the effects of hyperglycaemia in cardiac intervals and variations. An ECG acquisition and

analysis system based on LabVIEW software is developed for collecting ECG signals and extracting features with abnormal changes. Observations of ECG alterations are categorised and estimated by means of statistical software SPSS (IBM) to find out the significant changes among ECG parameters between hyperglycaemic and normoglycaemic states. The proposed hyperglycaemia detection system is performed in MATLAB environment. Performances of different algorithms are compared to show the enhancement in the outcome and reveal the advantages of combined ANN classifier.

1.4 CONTRIBUTIONS

Changes in electrocardiogram during hyperglycaemia in T1DM are still far from being well understood. This PhD research has contributed important findings about the effects of hyperglycaemia on electrical activity of the heart. This approach has provided useful clinical perception. The study focuses on the observations of abnormal changes in ECG parameters under hyperglycaemia conditions including the degradations of repolarisation parameters and HRV measures. This provides important insight into the development of a non-invasive strategy for the classification of hyperglycaemia in patients with diabetes. Advanced intelligent algorithms and ECG parameters have been introduced for hyperglycaemia detection. These results have been presented in several journal and conference papers which are mentioned in the section 1.6. All papers are full-reviewed, organised by the IEEE associations and published in IEEE *Xplore* (ieeexplore.ieee.org).

The contributions of this doctoral research are:

(i) This study has investigated and discovered the abnormal changes of ECG parameters under hyperglycaemic conditions in T1DM. Ventricular repolarisation parameters and HRV measures are decreased significantly under a hyperglycaemia state compared to normoglycaemia. Both cardiac intervals and HRV measures are highly correlated with high blood glucose levels. These findings

contribute to the limited literature of alterations in electrocardiogram during hyperglycaemia.

(ii) This research has indicated the potential ECG parameters for the classification of hyperglycaemia. By means of PCA, the most important ECG factors among cardiac intervals and HRV measures that are responsible for the variation of high blood glucose levels in diabetic patients have been identified. This optimal group of ECG variables is selected as input in order that the intelligent algorithms perform well for hyperglycaemia detection.

(iii) An advanced neural network has been introduced for a hyperglycaemia detection model. It is a two-stage ANN classifier using PCA to reduce the dimensional size of input and then implemented by the most popular training technique Levenberg-Marquardt. This approach produces very good sensitivity and acceptable specificity in hyperglycaemia detection.

(iv) A combined neural network with PSO has been investigated to enhance the performance of hyperglycaemia detection. PSO trained feed forward neural network can avoid local minimum and converge to global minimum to achieve better results. This combination shows significant improvement in the performance of hyperglycaemia detection compared with the previous approach.

1.5 STRUCTURE OF THESIS

This thesis consists of seven chapters, appendix and reference. Each of the main chapters III, IV, and V is completed with a discussion and conclusion. Excluding the first chapter, the remaining chapters of the thesis are organised as follows:

- Chapter II presents the literature review related to hyperglycaemia in diabetes. It describes the glucose regulation in hyperglycaemia and the blood glucose threshold definition of hyperglycaemia. Alternatively, the chapter introduces the effects of hyperglycaemia on the electrical activity

of the heart. It provides basic information on ECG intervals including QT and QT_C, RT and RT_C, T_PT_E and T_PT_{EC}, and PR intervals. Besides, the section describes heart rate variability and an analysis tool for HRV. This chapter also reports on current techniques in detecting hyperglycaemia either minimally invasively or non-invasively. It shows the advantages and disadvantages in each method and reveals the gap in the literature of identifying hyperglycaemia to motivate a novel non-invasive strategy of hyperglycaemia detection.

- Chapter III provides a basic method applying artificial neural network and ECG parameters for hyperglycaemia detection. The chapter begins with clinical study to generate the data set which would be used for hyperglycaemia detection in this research. This part includes the description of the study, the process of ECG acquisition, the feature extraction and the ECG parameters obtained from the study. After that, a multilayer feed-forward artificial neural network (ANN) is introduced for classification of hyperglycaemia. Different training algorithms are implemented to choose the best one. A comparison of results between multiple regression and ANN is also displayed to show the outperforming of ANN.
- Chapter IV proposes a novel method for classification of hyperglycaemia by including PCA to an ANN classifier. PCA provides the features dimensionality reduction to find the optimal group of ECG variables as input. Variety of inputs and training algorithms are implemented to compare the performance of hyperglycaemia detection when including, and not including PCA.
- Chapter V indicates an optimisation of ANN classifier for hyperglycaemia detection. The optimal group of ECG variables which resulted from PCA is fed to the input of the classifier and then PSO is used to train the feed-forward neural network with back-propagation. The performance of this approach is also compared with the outcomes of other methods in chapter IV and V.

- Chapter VI provides discussion and the overall conclusion for this research. The discussion involves the clinical insight of the influences of hyperglycaemia on abnormal changes in ECG parameters. It also discusses the performances of different methods utilised in the study. Main findings of the research are discussed as significant contributions of this thesis to the knowledge. In addition, the chapter presents future research on hyperglycaemia mechanism and other potential methods for hyperglycaemia detection.

1.6 PUBLICATIONS RELATED TO THESIS

The following journal and conference papers have been presented during the doctoral research:

Journal paper:

L. L. Nguyen, Steven Su, Hung Nguyen, "Non-invasive Detection of Hyperglycaemia in Type 1 Diabetes Mellitus using Electrocardiographic Parameters", submitted to *IEEE Trans on Biomedical Engineering*, 2014.

Conference papers:

- L. L. Nguyen, Steven Su, Hung Nguyen, "Identification of Hypoglycaemia and Hyperglycaemia in Type 1 Diabetic patients using ECG parameters," in *Engineering in Medicine and Biology Society (EMBC), 34th Annual International Conference of the IEEE*, pp. 2716-2719, 2012.
- L. L. Nguyen, Steven Su, Hung Nguyen, "Effects of Hyperglycaemia on Variability of RR, *QT* and corrected *QT* intervals in Type Diabetic Patients," in *Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE*, pp. 1819-1822, 2013.
- L. L. Nguyen, Steven Su, Hung Nguyen, "Neural Network Approach for Non-invasive Detection of Hyperglycaemia using Electrocardiographic

Signals", accepted in *Engineering in Medicine and Biology Society (EMBC), 36th Annual International Conference of the IEEE,* August 26-30, 2014.

CHAPTER 2 . LITERATURE REVIEW

The literature review in chapter 2 presents a comprehensive insight into hyperglycaemia and current research on hyperglycaemia detection. To understand the physiology of hyperglycaemia and propose a novel non-invasive method for hyperglycaemia detection, a detailed background of blood glucose regulation related to hyperglycaemia and the effect of hyperglycaemia on the electrical activity of the heart are described. An outline and evaluation on the current research and technologies available to monitor the blood glucose levels and hyperglycaemia are provided in the later section. The gap in the work of hyperglycaemia detection based on physiological signals is revealed afterwards. The proposed strategy for hyperglycaemia detection is presented in the last part of this chapter.

2.1 HYPERGLYCAEMIA

2.1.1 Definition and causes of hyperglycaemia

Hyperglycaemia is the technical term for high blood glucose (sugar). High blood glucose happens when the body has too little insulin or when the body can't use insulin properly.

Glucose levels are measured in either:

- Milligrams per decilitre (mg/dl), or
- Millimoles per litre (mmol/l), which can be acquired by dividing (mg/dl) by factor of 18

Scientific journals are moving towards using mmol/l; some journals now use mmol/l as the primary unit but quote mg/dl in parentheses (Wikipedia, 2013).

Chronic hyperglycaemia that persists even in fasting states is most commonly caused by diabetes mellitus, and in fact chronic hyperglycaemia is the defining characteristic of the disease. Intermittent hyperglycaemia may be present in pre-diabetic states. Acute episodes of hyperglycaemia without an obvious cause may indicate developing diabetes or a predisposition to the disorder.

In diabetes mellitus, hyperglycaemia is usually caused by low insulin levels (type 1 diabetes mellitus) and/or by resistance to insulin at the cellular level (type 2 diabetes mellitus), depending on the type and state of the disease. Low insulin levels and/or insulin resistance prevent the body from converting glucose into glycogen (a starch-like source of energy stored mostly in the liver), which in turn makes it difficult or impossible to remove excess glucose from the blood. With normal glucose levels, the total amount of glucose in the blood at any given moment is only enough to provide energy to the body for 20-30 minutes, and so glucose levels must be precisely maintained by the body's internal control mechanisms. When the mechanisms fail in a way that allows glucose to rise to abnormal levels, hyperglycaemia is the result.

2.1.2 Blood glucose regulation

Glucose in the blood is the most important carbohydrate fuel in the body; it provides a source of energy for all tissues of the body. Glucose levels vary before and after meals, and at various times of day; the definition of "normal" varies among medical professionals. Glucose levels are usually lowest in the morning, before the first meal of the day (termed "the fasting level"). In the fasting level, the blood sugar concentration is between 80 and 90 mg. per 100 cc. of blood. Even after several days of fasting the blood sugar will be maintained around this level in a well-nourished individual. It is essential to normal health that the blood sugar be maintained at this

level, and that it should not fall below this level for periods longer than an hour. After a meal containing sugar the blood sugar rises at once, usually reaches a concentration of 120 to 140 mg after 11/2 to 2 hours, and then gradually falls during the third and fourth hours to the previous fasting level (Sandler, 1951).

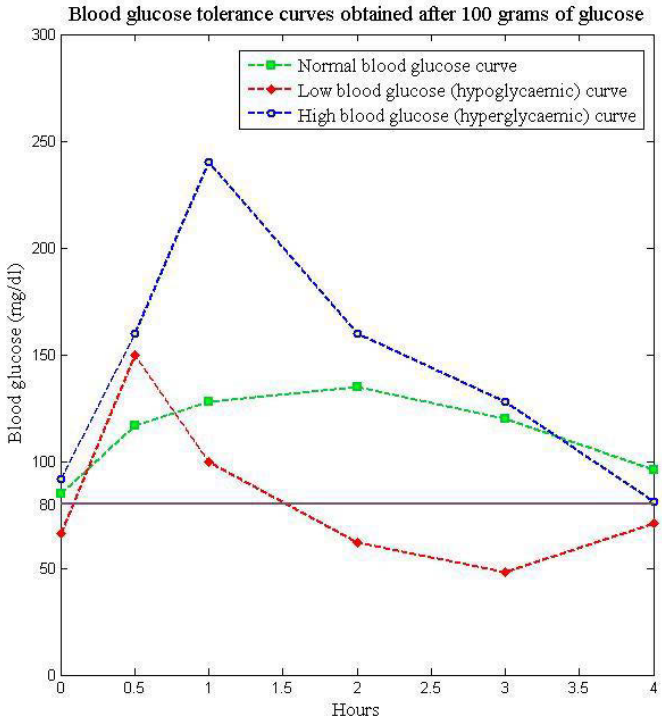


Figure 2.1: Blood glucose curves of three glycaemic states after a meal containing 100 grams of glucose

The body maintains blood glucose homeostasis mainly through the action of two hormones secreted by the pancreas. These hormones are insulin, which is released when glucose levels are high, and glucagon, which is released when glucose levels are low. Throughout the day, the release of insulin and glucagon by the pancreas maintains relatively stable levels of glucose in the blood. The maintenance of the blood sugar at normal levels is brought about by an efficient regulatory mechanism. The main organs in this mechanism are the liver, the autonomic nervous system, and certain glands of internal secretion called endocrine glands. The liver is at the centre of this mechanism and serves as a storehouse of the blood sugar supply. The foods are digested and broken down into simpler chemicals which are absorbed from the gastrointestinal tract, carried to the liver, and there built up (synthesised) by the liver cells into a complex compound called glycogen. This breakdown of liver

glycogen is controlled chiefly by nerves of the autonomic system which consists of two divisions, the sympathetic and the parasympathetic.

In the liver, insulin promotes conversion of glucose into glycogen and into fat. In muscle, insulin promotes the use of glucose as fuel and its storage as glycogen. In fat cells, insulin promotes the uptake of glucose and its conversion into fats. If glucose is being absorbed from the small intestine and there is no circulating insulin, most cells of the body cannot make use of the circulating glucose and its levels can rise to dangerous levels. This is what happens in a form of diabetes in which the pancreas does not secrete insulin. This condition is called hyperglycaemia, accompanying the high level of blood glucose; excess glucose may be found in the urine.

2.1.3 Blood glucose threshold definition

The blood glucose level threshold for defining hyperglycaemic events varies by study, depending on different patient populations, such as non-diabetic or diabetic patients, critically ill patients in intensive care units or outpatients. In this research, we focus on juvenile T1DM patients. Therefore, based on the recommendations of scientists (Hirshberg *et al.*, 2008) regarding levels of blood glucose among juvenile diabetic patients, glycaemic states can be classified as follows:

- Hypoglycaemia (Low range): blood glucose level (BGL) $\leq$ 60 mg/dl (or $\leq$ 3.33 mmol/l)
- Normoglycaemia (Normal range): 60 mg/dl <BGL < 150 mg/dl (or 3.33 mmol/l < BGL < 8.33 mmol/l)
- Hyperglycaemia (High range): BGL $\geq$ 150 mg/dl (or $\geq$ 8.33 mmol/l)

Cut-off value of 8.33 mmol/l is high enough for symptoms of hyperglycaemia to arise in adolescent diabetic patients during the night (Association, 2012). Levels that are significantly and persistently above this may necessitate treatment in hospitalised patients to avoid worsening outcomes.

2.2 EFFECT OF HYPERGLYCAEMIA ON THE ELECTRICAL ACTIVITIES OF THE HEART

2.2.1 Effect of hyperglycaemia on electrocardiogram intervals

The electrocardiogram (ECG) is a graphic recording of the electrical activity produced by the heart. It can be measured in twelve leads to give information about the regulation of heartbeats. In ECG signal, the important components are normally *P*-wave, QRS complex and *T*-wave. These components provide most of the useful information in the ECG in terms of intervals and characteristics defined by their features. The ECG features under investigation involve the parameters in depolarization and repolarisation stages of the heart. Medications or diseases that affect the cardiovascular system might reflect alterations in the characteristics of ECG signals.

In the past few years, several studies have reported a small number of ECG alterations in patients with hyperglycaemia. ECG alterations caused by the effects of hyperglycaemia have demonstrated changes of some ECG parameters, including *QT* interval, *PR* interval and mean *RR* interval (as shown in Figure 2.2). *QT* interval on the surface of ECG is measured from the beginning of the *Q*-wave to the end of *T*-wave. It is an indirect measure of the duration of depolarization, which is reflected by QRS complex, and ventricular repolarisation, which is indicated by the length from *J* point to the end of *T*-wave. *PR* interval is measured from the beginning of the *P*-wave to the beginning of the *Q*-wave. The *PR* interval reflects the time the electrical impulse takes to travel from the sinus node through the AV node where it enters the ventricles and is therefore a good estimate of AV node function. *RR* interval is the duration between the two consecutive *R* waves which occurs during ventricular depolarization. Heart rate in beats-per-minute can be calculated by 60 (number of seconds in a minute) divided by *RR* interval (measured in seconds).

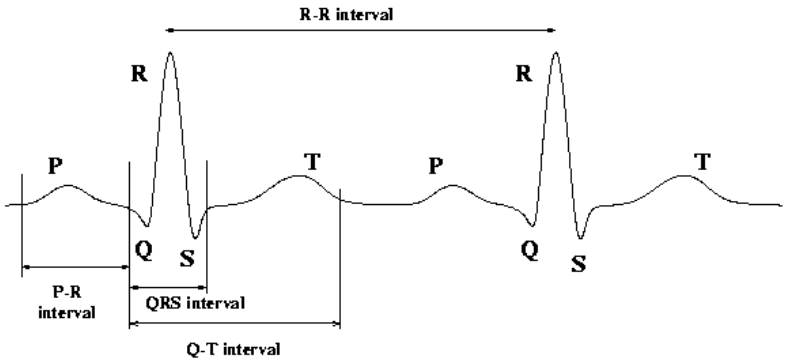


Figure 2.2: *QT*, *PR* and *RR* intervals in an electrocardiogram

The ability to define *QT* interval changes is important because of its association with ventricular abnormalities. The *QT* interval abnormalities are hampered by the fact that the *QT* interval is not constant and highly dependent on the preceding cardiac cycles and therefore on the heart rate. The *QT* interval varies not only with the heart rate, but also with many factors, including gender and because of diurnal variability, time of the day (Gowdar, 2010). To correct the inconstancy of dependence of the *QT* interval on preceding *R*-to-*R* changes, a rate-normalized or corrected *QT* interval (QT_C interval) is used. The object of QT_C is to normalize the *QT* interval to the value that it would have had if the heart rate was 60 beats/min, that is $QT_C = QT$ when *RR* interval is equal to one second. This is the meaning of *QT* correction formula developed by Bazett. In Bazett's formula, QT_C is defined as the *QT* interval divided by the square root of its preceding *R* to *R* time interval, which is given by:

$$QT_C = \frac{QT}{\sqrt{RR}} \tag{2.1}$$

Many other empirical formulas have been developed; however, there is no agreement as to which is the best method for *QT* correction. Bazett's formula therefore remains the most widely accepted method for rate normalisation of *QT* interval and is used in this research hereafter.

Changes in *QT* interval and QT_C under hyperglycaemic conditions have been observed, however, results from clinical studies are less consistent. Gordin *et al.*

(2008) recently reported that maintaining a plasma glucose concentration of 15 mmol/l by hyperglycaemia glucose clamp induces QT_C interval prolongation in type 1 diabetic patients and in healthy control subjects. The report stated that compared with normoglycaemia, acute hyperglycaemia increased the QT_C interval from 23 to 28 ms in type 1 diabetes and from 34 to 44 ms in healthy volunteers (as described in Figure 2.3) after 0, 60, 120 minutes of hyperglycaemia. These results agree with Marfella *et al.*'s finding that QT_C increased in a similar experiment of acute hyperglycaemia in healthy men (Marfella *et al.*, 2000).

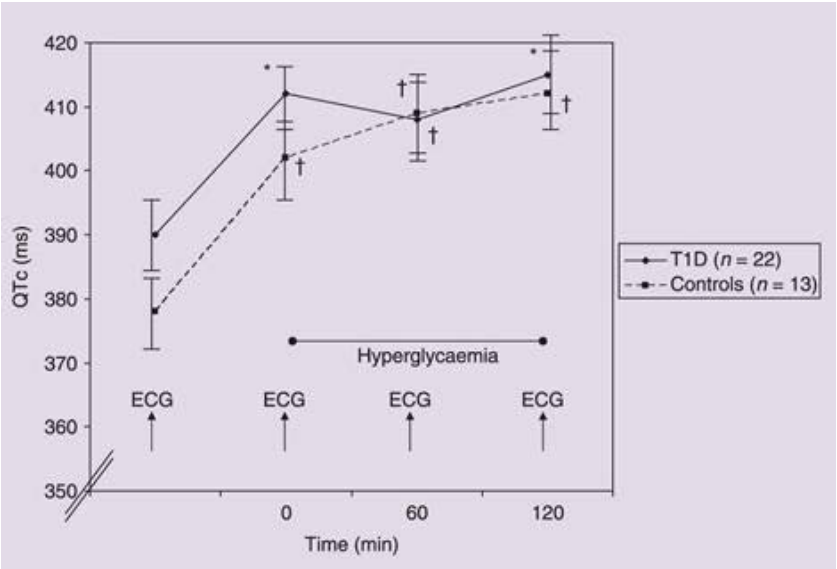


Figure 2.3: QT_C interval changes during 0, 60, 120 minutes of hyperglycaemia vs. normoglycaemia in study of Gordin et al. (Gordin *et al.*, 2008)

Inversely, in a recent study where simultaneously recorded QT_C values and glucose levels were analysed in patients with Type 1, QT_C prolongation was associated only with hypoglycaemia, not with hyperglycaemia (Suys *et al.*, 2006). They simultaneously recorded QT and QT_C values using 24-h Holter registration and glucose levels with a continuous glucose monitoring system in children and adolescents with type 1 diabetes mellitus. In this study, a significant correlation was found between QT, QT_C and QT_{Cmax} and glucose levels, with the higher QT_C values coinciding with lower glucose readings and vice versa (as shown in Figure 2.4). Apparently, the role of hyperglycaemia in causing QT abnormalities in people with diabetes is not clear. Pertinent to the reports of Gordin and Marfella *et al.* are the

results of a study showing that QT_C duration was significantly related to glucose concentrations 60 and 120 minutes after an oral glucose load in elderly men (Dekker *et al.*, 1996). Contrary findings have shown that on oral glucose test induced QT_C prolongation only in long *QT* syndrome patients; QT_C did not change in healthy control subjects (Nishizaki *et al.*, 2002). Further studies in type 1 diabetic patients are therefore needed to establish the clinical relevance of the findings of either Gordin and Marfella *et al.* or Suys.

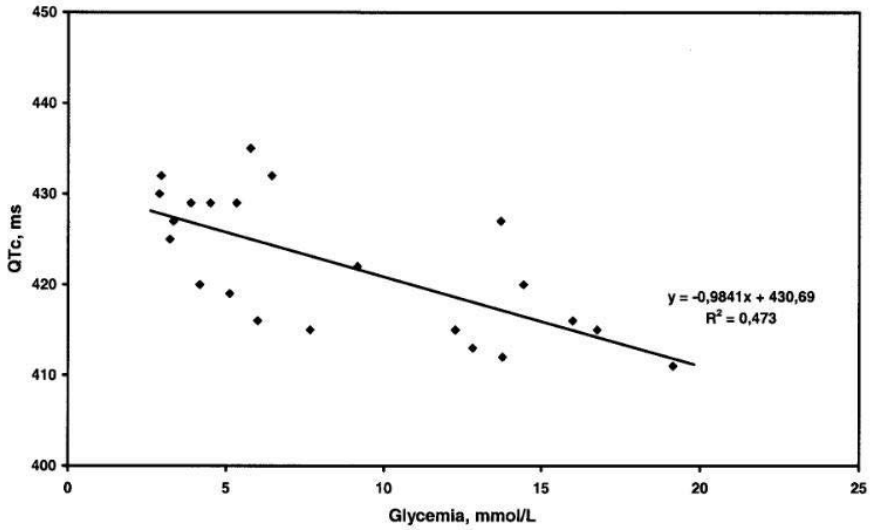


Figure 2.4: Individual correlation QT_C per hour versus glycaemia (r = 0.672, p = 0.003) in a study of Suys et al. (Suys *et al.*, 2006)

Increased *PR* interval in hyperglycaemia has been discussed in the research of Marfella *et al.* (Marfella *et al.*, 2000) with acute hyperglycaemia in healthy males. They reported that *PR* interval showed a significant increment at 120 min of the clamp (Figure 2.5). This effect is due to *PR* intervals commonly associated with atrial fibrillation (Lorsheyd *et al.*, 2005; Homoud, 2009), and chronic hyperglycaemia but not hypoglycaemia which may contribute to atrial fibrillation burden in several ways (Aksnes *et al.*, 2008; Gandhi *et al.*, 2005). The *PR* interval on the surface ECG serves as an index of atrio-ventricular conduction, in which the main component, *P*-wave duration, represents atrial depolarization that causes atrial contraction. Prolonged *P*-wave duration and *P*-wave dispersion have also been widely seen in pre-diabetic patients without coronary artery disease, hypertension or ischemia (Karabag *et al.*, 2011; Yazici *et al.*, 2007). This explains why diabetes

mellitus is known as an independent and strong risk factor for the development of atrial fibrillation.

A possible mechanism of the increments of any durations regarding *P*-wave under hyperglycaemic condition is that chronic hyperglycaemia causes structural and functional disorders by changing the chemical composition of the proteins present in cell membrane structure (Yazici *et al.*, 2007). An experimental study has reported that extracellular protein deposition and interstitial myocardial fibrosis might be other mechanisms responsible for the prolongation of *P*-wave dispersion, by influencing heterogeneity of atrial conduction time and atrial refractoriness in diabetic patients (Kato *et al.*, 2006). The reason for the prolongation of *P*-wave dispersion may be a structural defect in the atrium, which is likely to be caused by damage due to hyperglycaemia (Karabag *et al.*, 2011). At present, although there is no adequate explanation for the mechanism of increased *PR* interval in diabetic patients, the same mechanism contributing to the effect of hyperglycaemia on *P*-wave dispersion is possibly also operative.

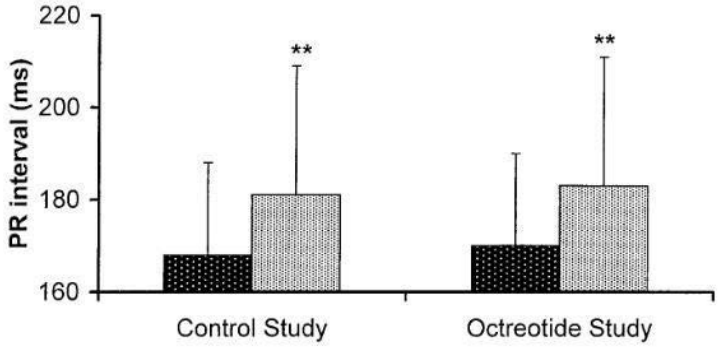


Figure 2.5: PR interval during hyperglycaemic clamps in Marfella *et al.*'s study (Marfella *et al.*, 2000). Left dark bar basal values, right light bar 120 min clamp values. ** $p < 0.01$ vs basal values.

2.2.2 Effect of hyperglycaemia on Heart rate variability

Variability of *RR* interval, known as heart rate variability (HRV) measures, has a close relationship to sympathetic and parasympathetic branches of the autonomic nervous system (ANS) (Electrophysiology, 1996). These branches are linked to the heart at the sinoatrial node (SA) where the beat is initiated for a normal heart. The sympathetic branch will raise heart rate, blood sugar and blood pressure

under stress while the parasympathetic branch will have an opposite effect once the body gets rest and relaxation (Shin *et al.*, 1989). The use of HRV has emerged as a simple, reliable and non-invasive method to evaluate the sympatho-vagal balance at the sinoatrial level. HRV signal, which is generated from ECG by calculating the inter beat intervals, is a non-stationary signal that represents the autonomic activity of the nervous system and the way it influences the cardiovascular system (Yaghouby, 2009). In addition, diabetes mellitus is one of the main causes of autonomic neuropathy, which can lead to abnormalities in heart rate control (Kudat *et al.*, 2006). Therefore, it is no doubt that there is association between HRV changes and diabetic patients. Changes of *RR* interval under hyperglycaemic conditions are then included in the analysis of HRV in this part. Since *RR* time intervals are less affected by the noise, analysis of HRV signal can be considered as a robust method in cardiology.

2.2.2.1 HRV measures

Measure of HRV features are basically relied on when analysing the variations of inter-beat *RR* time intervals. The variation analysis can be accomplished in the time or frequency domain. The time domain analysis, the simpler way, calculates a number of variables that describe beat to beat changes. The simplest time domain variable is mean *RR* interval. The most commonly used variable is the standard deviation of the *RR* intervals (*SDNN*), that is, the square root of variance. Since variance is mathematically equal to total power of spectral analysis, *SDNN* reflects all the cyclic components responsible for variability in the period of recording (Electrophysiology, 1996). Other popular measures derived from interval differences include *RMSSD*, the square root of the mean squared differences of successive *RR* intervals, and *pNN50*, the proportion derived by dividing the number of interval differences of successive *RR* intervals greater than 50 ms to the total number of *RR* intervals. These two measures of short-term variation estimate high-frequency variations in the heart rate and thus are highly correlated. In addition, the series of *RR* intervals also can be converted into a geometric pattern such as the sample density distribution of *RR* interval durations. The HRV triangular index

measurement is the integral of the density distribution (that is, the number of all *RR* intervals) divided by the maximum of the density distribution. The triangular interpolation of *RR* interval histogram (*TINN*) is the baseline width of the distribution measured as a base of a triangle approximating the *RR* interval distribution.

The frequency domain components of HRV signal, derived from its power spectrum analysis, has been commonly used to quantify the relative dominance of parasympathetic and sympathetic influence on the heart (Malliani *et al.*, 1991). Spectral analysis technique can distinguish among the intrinsic sources of HRV, as these rhythms occur at different frequencies. Typically, frequency domain has been divided into three major groups: high, low and very low frequency bands. Relative power in high frequency (*HF*) areas (0.15-0.4 Hz) has been used to infer parasympathetic nervous system activity. The low frequency (*LF*) component (0.04-0.15 Hz) reflects inputs from both branches of the autonomous nervous system. Very low frequency (0.0033-0.04 Hz) is influenced by many factors, including thermoregulation and the rennin-angiotensin system. The ratio of the *LF* to *HF* components represents a measure of sympatho-vagal balance (Barron *et al.*, 1996; Pagani *et al.*, 2000).

The variety of time and frequency measures of HRV is summarised in Table 2.1. It also includes the overview of the significations of components and their correlations between others. Although time domain measures can investigate the differences in heart rate or cardiac cycle length, the frequency measures are usually able to provide results that are more easily interpretable in terms of physiological regulations (Electrophysiology, 1996). These measures are sensitive indicators of autonomic function in diabetic patients and may provide important insight into the pathogenesis of autonomic neuropathy in hyperglycaemia (Noritake *et al.*, 1992).

	HRV measures	Units	Definition	Signification
Time domain	*Mean RR*	ms	mean of the *RR* intervals between two normal heartbeats	The same as average heart rate
	SDNN	ms	standard deviation of all normal *RR* intervals	Reflect both the sympathetic and parasympathetic influence on HR variability
	RMSSD	ms	Square root of mean squared differences of successive *RR* intervals	Reliable index of cardiac parasympathetic activity
	pNN50	ms	percentage of beats with a consecutive *RR* interval difference > 50 ms	Reliable index of cardiac parasympathetic activity
	HRV triangular index		the integral of the density distribution (that is, the number of all NN intervals) divided by the maximum of the density distribution	Both these measures express overall HRV measured over 24 hours and are more influenced by the lower than by the higher frequencies
	TINN	ms	The triangular interpolation of NN interval histogram (*TINN*) is the baseline width of the distribution measured as a base of a triangle approximating the NN interval distribution	
Frequency domain	*VLF*	ms^2	Very low frequency, from 0.0033 – 0.04 Hz	Represent an assortment of factors such as thermoregulation fluctuation
	LF	ms^2	Low frequency, from 0.04 – 0.15 Hz	Represent a mixture of parasympathetic and sympathetic activity of ANS
	HF	ms^2	High frequency, from 0.15 – 0.4 Hz	Represent the parasympathetic activity of ANS
	LF/HF		Low frequency/High frequency ratio	Reflect sympathetic/parasympathetic balance or sympathetic modulation
	Total Power	ms^2	≤ 0.4 Hz	Variance of all *RR* intervals

Table 2.1: Selected time and frequency domain measures of HRV

2.2.2.2 HRV in hyperglycaemia

Recently, reduced HRV has been seen among patients with Type 1 diabetic mellitus and the main driver to this abnormality appears to be hyperglycaemia. In the Framingham study, Singh *et al.* (2000) presented the association of hyperglycaemia with reduced HRV. The cohort study carried out with 1,919 subjects categorised by levels of fast blood glucose (FBG), including 1,779 who had normal FBG (< 110 mg/dl); 56 who had impaired FBG (≥ 110 and < 126 mg/dl) and 84 who were diabetic (FBG ≥ 126 mg/dl). They found that FBG levels were inversely and significantly correlated with *SDNN* and *LF* and *HF* powers, which mean high blood glucose levels caused lower HRV in subjects with diabetes. *SDNN* and *LF* and *HF* powers were reduced in the diabetic group and in subjects with impaired FBG compared with those with normal FBG. Their findings suggest that the presence of reduced HRV in subjects with diabetes imply the occurrence of heightened sympathetic activity and/or reduced vagal activity. Singh *et al.* agreed with the hypothesis that elevated FBG levels may be involved in the pathogenesis of autonomic neuropathy. An important strength of this population-based study was that subjects have been well characterized through many years of follow-up. The availability of reliable long-term data allowed researchers to select subjects who were free of clinical cardiovascular disease, which can alter autonomic function and HRV measurements.

Kudat *et al.* (2006) investigated heart rate variability in 31 diabetic patients (8 patients having Type 1 and 23 having Type 2) and in 30 control healthy individuals. None of the patients was on drugs that may affect HRV analysis, such as α- and β-blockers, and none of the patients had heart failure or uraemia. The study revealed that, all time and frequency domain parameters except mean *RR* interval and LF/HF were significantly lower in diabetic patients than in healthy controls, and among diabetes patients those with microvascular complications had the lowest HRV parameters. Their findings showed a relation between decreased HRV and the presence of micro vascular complications in diabetes.

The latest published study in the effect of hyperglycaemia on HRV, by Jaiswal *et al.* (2013), compared HRV measures in 354 youths with and 176 youth without type 1 diabetes mellitus and explored potential contributors of altered HRV. The results showed that HRV variables were altered among individuals with Type 1 diabetes compared with control subjects; such as *SDNN*, a marker of overall HRV, was significantly lower in Type 1 diabetic patients, and *RMSSD* and *HF* power, markers of parasympathetic function, were decreased in youth with Type 1 diabetes versus control subjects where *LF* power, a surrogate for sympathetic dysfunction, was higher among those with Type 1. The study found evidence of reduced overall HRV, including a pattern of parasympathetic loss with sympathetic overdrive, among youth and young adults with Type 1 diabetes mellitus, independent of traditional CVD risk factors. Their findings suggested an important role for hyperglycaemia in mediating these abnormalities in a contemporary cohort of diverse youth with an average duration of diabetes of approximately 10 years.

2.3 CURRENT RESEARCH OF HYPERGLYCAEMIA DETECTION

Early detection of hyperglycaemia is considered as a priority in preventing and controlling this dangerous condition. Many studies have been undertaken to establish effective methods and develop medical devices to detect hyperglycaemia. Basically, all methods are focused on monitoring the blood glucose to detect the high glucose levels. Optimal hyperglycaemia management relies on accurate glucose monitoring techniques. Figure 2.2 shows different classification methods of blood glucose monitoring: invasive, minimally invasive and non-invasive.

The following is the description of current methods used to identify hyperglycaemic episodes in diabetic patients. Existing techniques commonly used invasive point sample methods like finger stick and minimally invasive/non-invasive continuous glucose monitoring methods. Others utilised electrochemical and optical methods, such as reverse iontophoresis and near infrared spectroscopy, respectively. Alternative methods for hyperglycaemia detection employed effects of hyperglycaemia on physicochemical patterns on the body including β cell, methyl nitrates and tear glucose. Current techniques have their own outstanding features;

however every method still shows some drawbacks that need more investigation to overcome the limitations.

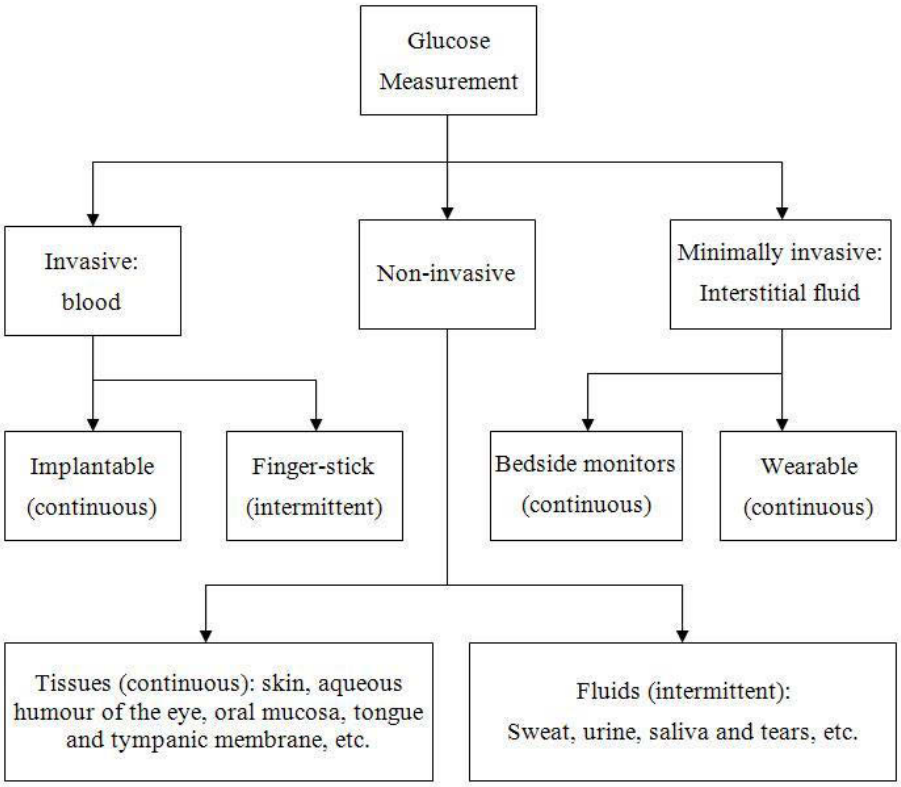


Figure 2.6: Overview of technologies for blood glucose monitoring (Ferrante do Amaral et al., 2008)

2.3.1 Test of blood glucose level - Finger stick method

Blood glucose level may be determined in whole blood, plasma or serum samples. If whole blood is used, the concentration of glucose will be lower than if the plasma or serum is used. This is due to the greater water content of the cellular fraction. Under usual circumstances, the concentration of glucose in whole blood is about 15% lower than in plasma or serum, but the difference will be less in patients with low hematocrits. The concentration of glucose is highest in the arterial circulation. The concentration of glucose in capillary samples is intermediate between venous and arterial. Glucose oxidase and reagents to measure the generation of hydrogen peroxide can be bonded to filter paper and the system used to measure glucose concentrations in a drop of capillary blood, is usually known as finger stick

method. This has resulted in the most important change in diabetes management since the introduction of insulin (McMillin, 1990).

Finger stick is the most common invasive method for blood glucose monitoring in patients (invasive means the penetration of body tissue with a medical instrument). This method is exactly as the name suggests. It's a stick for puncturing the finger to acquire a blood sample. Blood glucose monitoring can be done at home with a variety of blood glucose monitors (also called a glucose meter- more details available in http://diabetesinsight.ie/) to obtain the blood sample. Surprisingly, many patients consider the discomfort of the finger stick preferable to the inconvenience and aesthetic unpleasantness of obtaining a urine sample for testing.

Usually a drop of whole capillary blood obtained through a finger prick is sufficient to use on a test strip that is then measured in a monitor to detect how much glucose is present in the blood. A finger prick can be done with a small lancet (special needle) or with a spring-loaded lancet device that punctures the fingertip quickly. The drop of whole blood is then placed on the reagent bonded to a testing (paper) strip. Instead of using a known volume of blood, an excess of blood is exposed to a fixed quantity of glucose oxidase for a finite period of time to estimate concentration. After the specified time, usually 1 minute, the excess blood is removed by washing or wiping and the colour is allowed to develop. The concentration is then estimated by comparing to a colour chart, or by using a portable reflectance meter specific to the reagent strip, to measure the developed colour to read the blood sugar level. Reflectance meters or glucose meters for measuring blood glucose are becoming increasingly sophisticated, compact, and reliable. Shirt-pocket-size models are now available, and prototype models that store the time, date, result, and insulin doses for later graphic printing at the patient's home or physician's office have been developed.

In most hands, the glucose oxidase strip method or finger stick is accurate and reliable. Since whole blood is used, the results tend to be slightly lower than simultaneous venous samples, but this is balanced by the fact that capillary blood has a higher glucose concentration than venous blood. The major sources of error are in failing to put a large enough drop of blood on the strip and inaccurate timing. For

patients who use reflectance meters, another source of error is failure to keep the machine clean and calibrated.

A wide variety of diabetes test strips are available on the Australian market. The Free Style Lite® Blood Glucose Monitoring System, a product of Abbott, USA, is a typical example for blood glucose test device (https://www.abbottdiabetescare.com/products/patient/fs-lite-overview.html). This device is intended for use in the quantitative measurement of glucose in capillary whole blood from the finger, upper arm and palm and venous whole blood. It is used by healthcare professionals and people with diabetes mellitus at home as an aid in monitoring the effectiveness of a diabetes control program. Results of the test are available almost immediately. The meter displays results from 20 – 500 mg/dl. Repeating the test with a new test strip is recommended if low (hypoglycaemic) or high (hyperglycaemic) blood glucose results are found to confirm a potentially serious medical condition.

Undeniably, finger stick method is still chosen to be a popular choice among diabetic patients for self-monitoring in the management of diabetes mellitus due to its quick response, simplicity and high reliability. Checking blood sugar levels by finger prick several times a day is often a necessary task in the lives of many people with diabetes. However, pricking fingers multiple times a day is painful, time consuming, can cause calluses and sensitive fingers, and is difficult if you have visual or dexterity limitations. Therefore, despite being the most accurate method for blood glucose monitoring, finger-stick still discourages people to do it on a regular basis because of the pain and discomfort of the puncture. An alternative method for blood glucose monitoring that would not require pricking the finger for a blood sample to eliminate the pain is quite essential.

2.3.2 Continuous glucose monitoring system (CGMS)

Continuous glucose monitoring (CGM) systems use a tiny sensor inserted under the skin to check glucose levels in tissue fluid. The sensor stays in place for several days to a week and then must be replaced. A transmitter sends information about glucose levels via radio waves from the sensor to a pager-like wireless monitor.

Results of at least four finger stick blood sugar readings taken with a standard glucose meter at different times each day are entered into the monitor for calibration. Because currently approved CGM devices are not as accurate and reliable as standard blood glucose meters, users should confirm glucose levels with a meter before making a change in treatment (Johnson, 2012).

CGM systems are more expensive than conventional glucose monitoring, but they may enable better glucose control. CGM devices produced by Abbott, DexCom, and Medtronic (http://www.medtronicdiabetes.com/treatment-and-products/guardian-real-time-cgm-system) have been approved by the U.S. Food and Drug Administration (FDA) and are available by prescription. These devices provide real-time measurements of glucose levels, with glucose levels displayed at 5-minute or 1-minute intervals. Users can set alarms to alert them when glucose levels are too low or too high. Special software is available to download data from the devices to a computer for tracking and analysis of patterns and trends, and the systems can display trend graphs on the monitor screen.

The main advantage of continuous glucose monitoring is that with CGMS the patients can easily and discreetly view their current glucose values continuously throughout the day. Moreover, it can help identify fluctuations and trends that would otherwise go unnoticed with standard HbA1c tests and intermittent finger stick measurements. The monitors have trend arrows that show the patient if his level is rising or falling quickly. Most CGMS have alarm systems that tell the patient when he is getting close to a pre-set high or low limit.

CGMS, however, still exhibit many problems. Firstly, CGMS doesn't replace the finger prick method. Patients still have to check their blood glucose levels 2 to 4 times a day to keep the CGMS calibrated. Additionally, CGMS requires a sensor inserted under the skin, potentially making the patients uncomfortable from the pain of the insertion. The major disadvantage of CGMS is the time delay between blood and interstitial glucose that can reduce the accuracy of the monitoring system. It takes glucose around 10–15 minutes to move from blood into tissue fluid, or back, so the CGMS measures lag behind what's really happening in the blood if conditions change rapidly (Keenan *et al.*, 2009).

2.3.3 Reverse iontophoresis technique

Iontophoresis, a technique involving the application of a low electric current to the skin, has been used for transdermal delivery of drugs for over a century. This technique was firstly investigated to measure glucose concentration by glucose extraction through intact skin in a study by G.Rao (Rao *et al.*, 1993). Because glucose extraction proceeds in the opposite direction (from the skin outward), this process has been termed "reverse iontophoresis". Glucose collected through the skin via reverse iontophoresis has been shown to correlate well with blood glucose (Tamada *et al.*, 1995).

Success of measuring glucose concentration using reverse iontophoresis techniques has led to the development of the GlucoWatch® Biographer (Cygnus Inc.). This commercialised wrist-worn device monitors glucose continuously for up to 13h, recording six glucose readings per hour (Sieg *et al.*, 2004). It is a typical non-invasive blood glucose monitoring device that is designed to have an alarm system, detect trends and track patterns in glucose levels for diabetic patients. The GlucoWatch®, however, still has considerable drawbacks, which may limit widespread use. These include a long warming-up period of 3 hours to adjust for the variability in skin permeability, device malfunctioning particularly in cold weather, significant skin irritation, and sensitivity to excessive sweating. Each time before the GlucoWatch® is used it must be calibrated with the results of a finger-stick blood glucose test. There is also a lag-time of approximately 18-20 minutes corresponding to real-time glucose (Chan *et al.*, 2002). These potential problems may restrict its acceptability for hyperglycaemia detection.

2.3.4 Near-infrared light (NIR) spectroscopy technique

Non-invasive blood glucose sensing by near infrared spectroscopy requires glucose dependent absorption of near-infrared light as it passes through a selected region of the human body. The term "near infrared light" refers to the use of an external light source with wavelengths in the infrared spectrum near the wavelengths of visible light. An NIR source can pass through or be reflected by a body part.

Glucose and other body constituents absorb a small amount of the light at each wavelength. Spectroscopy, an established technology used to measure energy containing many wavelengths, detects the amounts of NIR absorbed at each wavelength. With spectroscopy, a data processing technique known as multivariate analysis simultaneously analyses the amount of light absorption at selected wavelengths for each blood glucose level. A polynomial formula is generated that converts the sum of relative contributions of absorption at the selected wavelengths to the blood glucose concentrations (Klonoff, 1997).

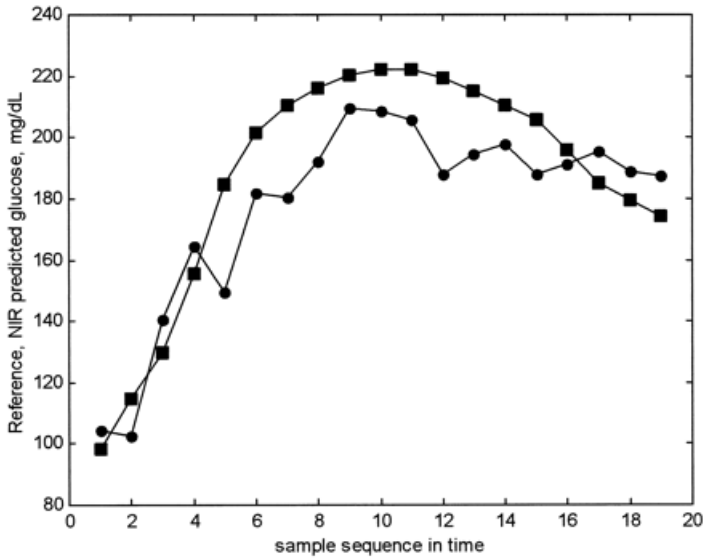


Figure 2.7: Reference glucose and NIR-predicted glucose profiles (Malin *et al.*, 1999)

The major problem with using NIR spectroscopy for blood glucose monitoring is the necessity for frequent recalibration. Glucose is responsible for < 0.1% of NIR absorbed by the body, whereas, water, fat, skin, muscle, and bone account for the vast majority of NIR absorption. Perturbations in the amounts of these substances can alter NIR absorption and thus invalidate the calibration formula for correlating light absorption with blood glucose concentrations that was generated during the calibration process (Aldhous, 1992). In addition, different probing locations and contact pressures between the probe and skin cause variations in the light propagation paths within tissue, making it very hard to obtain stable near-infrared spectra concerning body content. Selection of a measurement site must consider the physical and chemical characteristics regarding the accuracy of

measurement. Therefore, NIR spectroscopy is not a potential choice for an up-to-date hyperglycaemia detection system.

2.3.5 Early detection of β cell death

In diabetes mellitus, β cell destruction is largely silent and can be detected only after significant loss of insulin secretion capacity. Scientists at the University of Chicago proposed a suggestion of early detection of ongoing β cell death which would allow for earlier interventions at a time before the development of hyperglycaemia, when a more significant β cell mass is present (Akirav *et al.*, 2011). They reported a non-invasive method for the detection of circulating β cell-derived DNA in murine models of acute and chronic β cell destruction that provides a sensitive biomarker for β cell death in pre-diabetic mice in vivo and in human tissues and serum. Their method can identify β cell death before the onset of hyperglycaemia and soon after the onset of T1DM.

This finding is still under investigation in pre-diabetic mice in vivo and in human tissues and serum. Further human studies are needed to evaluate and validate this method's general applicability in human clinical settings. This strategy may prove useful for monitoring the development of hyperglycaemia with possible ongoing β cell destruction. However, reviewers raise concerns about accuracy of this method in real-life human experiments and unknown possible side effects.

2.3.6 Testing of methyl nitrates in exhaled air

UC Irvine scientists have been testing the levels of methyl nitrates in exhaled air as a marker for elevated blood sugar levels (Novak *et al.*, 2007). They hypothesized that different exhaled VOC (volatile organic compounds) profiles could be present in children with type 1 diabetes during spontaneous hyperglycaemia, due to the presence of low insulin and increased free fatty acids and ketones in individuals with diabetes. The children received intravenous insulin and glucose as needed, and the levels of plasma glucose and exhaled gases were monitored during constant normoglycaemia or initial hyperglycaemia with gradual correction. By

using a chemical analysis method developed for air-pollution testing, UC Irvine chemists and paediatricians have found that the kinetic profile of exhaled methyl nitrate was most strongly correlated with that of plasma glucose; children with type-1 diabetes exhale significantly higher concentrations of methyl nitrates when they are hyperglycaemic.

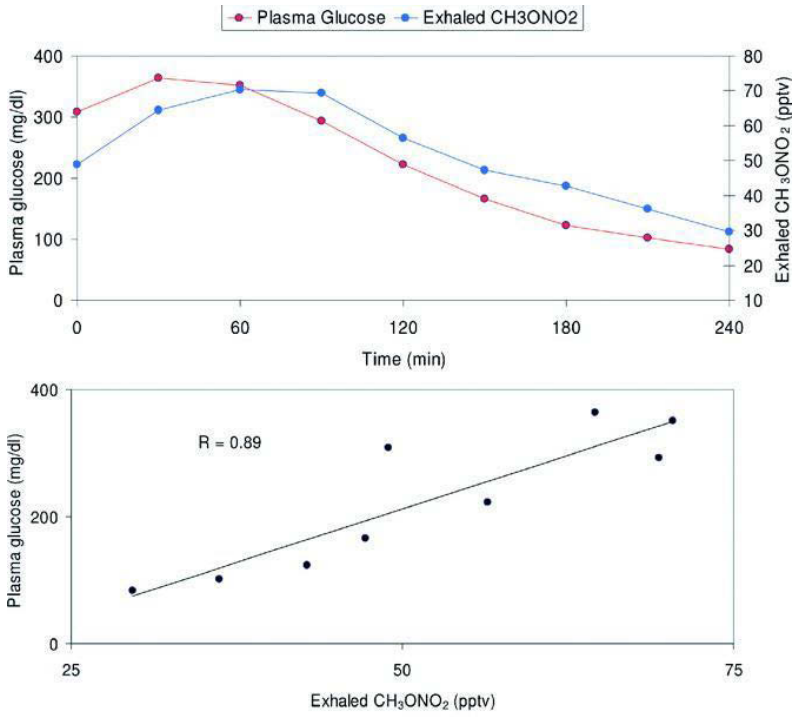


Figure 2.8: Plasma glucose and exhaled methyl nitrate profiles in one child with T1DM in study of Novak et al. (2007)

This study, however, showed estimation of glucose curves from exhaled gas only on an individual basis, not yet employed to a whole study population that requires identification and quantification of the relative contribution of all pertinent covariates. This technique, itself, has intrinsic difficulties in measurement and analysis, which have resulted in inconsistent outcome, limiting its practical applicability.

2.3.7 Measuring glucose concentration in tears

Tear glucose has been studied for several decades. Many reports have demonstrated that tear glucose is higher in diabetic subjects than in healthy ones. The correlation between tear glucose and blood glucose has been studied by different

methods in both humans and animals. Based on these promising results, scientists have attempted to continuously monitor the concentration of ocular tear glucose by using nano-structured and disposable contact lens-based sensors for diabetes (Zhang *et al.*, 2011; Bishop *et al.*, 2010). Sensors built onto lenses would let diabetic wearers keep tabs on blood-sugar levels without the need for finger pricks. Monitoring the tear glucose level may provide a feasible approach for non-invasive diagnosis and diabetes control.

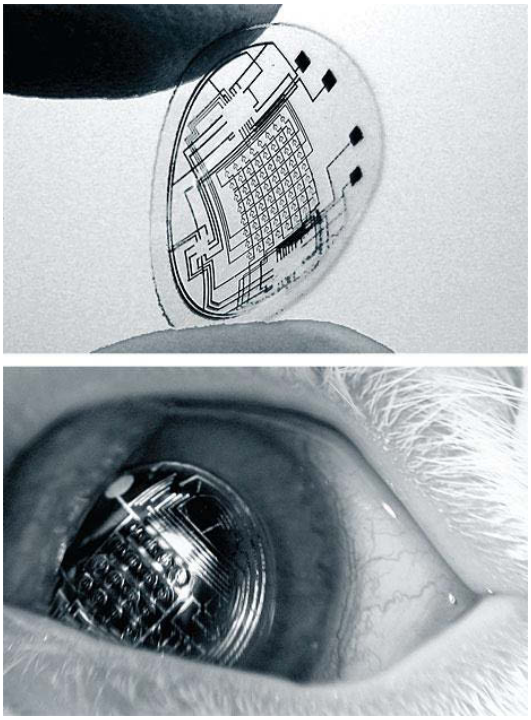


Figure 2.9: Contact based-lens sensors for tear glucose monitoring developed by researchers at University of Washington (Parviz, 2009)

A group of researchers at the University of Washington have fabricated prototype lenses with additional enhancements, including an antenna, which measure glucose concentration in tears and send the signals to a receiver (Yao *et al.*, 2011), as displayed in Figure 2.9. They have tested their first few prototypes successfully on animals and are undertaking developments for humans. However, accuracy and sensitivity of this technique are still unclear due to interfering factors when conducting tear glucose measurements. These include standards for normal tear concentration in users; time-lag between blood glucose and anterior aqueous humour

glucose concentrations; discrepancy in tear glucose levels with different methods of tear sampling. Further efforts are needed to develop a new design for contact based-lens sensors to overcome the shortcomings.

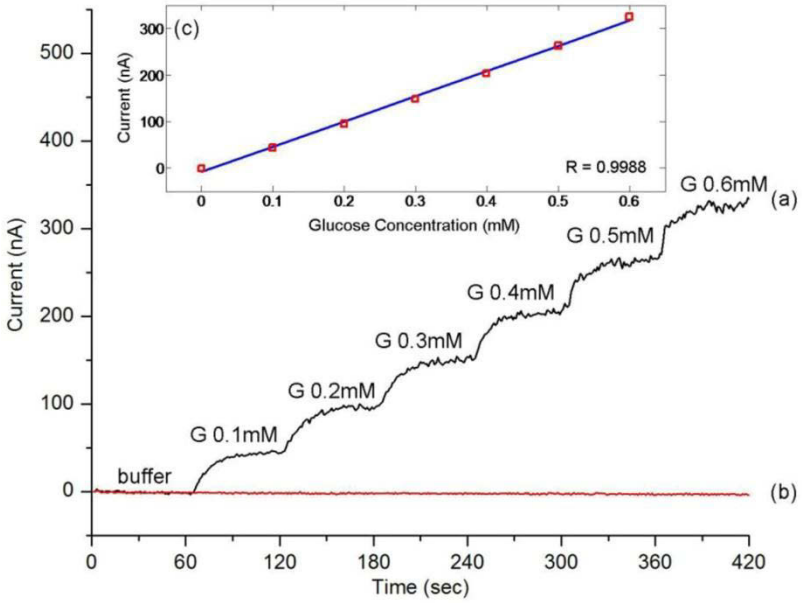


Figure 2.10: Measured glucose response and calibration curve of sensor in study of Yao et al. (2011)

2.4 THE GAP IN LITERATURE REVIEW

Despite the fact that almost all commercially successful blood glucose monitoring systems are invasive, there is an immense need to develop alternative techniques that will alleviate the pain and suffering from frequent pricking of fingers for blood samples in diabetic patients. Efforts have been made in order to reduce the level of invasiveness by decreasing the blood sample volume to a few micro-litres, and measuring areas of the body less sensitive to pain than fingertips, such as the forearm, upper arm, or thigh. Drawbacks of such systems are lack of control during sleep or manual activities, undiscovered episodes of hyperglycaemia, risks of infection, nerve damage and the discomfort of pricking the finger several times a day (Ferrante do Amaral *et al.*, 2008). Minimally invasive measurements, i.e. sampling the interstitial fluid (ISF) with subcutaneous sensors, remain uncomfortable for the patient's therapy, require continuous calibration, and are highly susceptible to bio-fouling (Gross *et al.*, 2000; Vashist, 2012). Hence, a non-invasive method is the goal

for the future of hyperglycaemia detection. Nevertheless, past attempts at non-invasive glucose monitoring failed to reliably separate glucose values from noise without an actual blood sample. By progressing different techniques, scientists are overcoming difficulties faced in the past. Some of the devices under development are truly non-invasive, while others require something to be injected or implanted under the skin to enable glucose detection (Ramchandani *et al.*, 2012).

The glucose concentrations can be determined by analysing the changes in measured body compartments, a substitution of blood, to minimize the invasiveness of blood glucose monitoring systems. For example, one option to non-invasive intermittent glucose control is using body fluids that could contain glucose, like saliva, urine, sweat or tears. Non-invasive continuous glucose monitoring could be accomplished through direct measurement of body tissues such as skin, aqueous humour of the eye, tongue or tympanic membrane (Ferrante do Amaral *et al.*, 2008). Still, there are always limitations regarding the accuracy of these measurements because of time-lag between real-time blood glucose levels and body fluid/tissue glucose. Moreover, the ratio of body fluids (intracellular, interstitial, plasma) are not only affected by factors such as activity level, diet or hormone fluctuations, but also by blood circulation, body temperature shift, metabolic activity and medication. Problems also occur due to changes in the tissue after the original calibration and the lack of transferability of calibration from one part of the body to another. Day-to-day changes in vasculature and tissue texture as well as the aging process may affect the long-term stability of glucose monitoring. Thus, there is a need to find an alternative potential physiological signal, which is highly associated with blood glucose level for effective non-invasive blood glucose monitoring strategy.

With hypoglycaemia, scientists have successfully detected this complication by utilising various physiological signals, especially non-invasive hypoglycaemia detection using electrocardiographic (ECG) signals. It indicates that ECG offers a quicker, more ubiquitous, non-invasive clinical and research screen for the early detection of hypoglycaemia, and potentially for hyperglycaemia, than other physiological signals. Indeed, few researches have revealed the association of hyperglycaemia and ECG parameters, so that further investigations are still required

to have a comprehensive insight in this issue. At least three main studies on hypoglycaemia detection, which employ the electrical activity of the heart, have achieved excellent performances with different computational intelligent algorithms. Meanwhile, there is no research on hyperglycaemia detection using ECG so far and no advanced computational algorithms have been employed for the classification of hyperglycaemia. Exploring this area may provide a great potential method for non-invasive hyperglycaemia detection to help patients manage the disease conveniently.

2.5 PROPOSED STRATEGY OF HYPERGLYCAEMIA DETECTION

Several ECG parameters have been investigated to correlate with the prevalence of hyperglycaemia. There still exists the controversy about the effect of hyperglycaemia on ventricular repolarisation parameters, like QT interval, RT interval and T_PT_E interval. Thus, this research carries out a new clinical study of a group of adolescent T1DM patients who only suffer from hyperglycaemic and normoglycaemic conditions to evaluate the influence of hyperglycaemia to the alterations of ventricular repolarisation intervals. Reduced heart rate variability has been seen among hyperglycaemic patients. In agreement with previous studies, this study demonstrates that high blood glucose levels cause lower HRV in subjects with diabetes. It extends with further observations of abnormalities and goes across the spectrum of hyperglycaemic and normoglycaemic events. Based on these explorations, certain ECG parameters could be determined as good markers for classification models to identify hyperglycaemic and normoglycaemic states in diabetics.

This research proposes a new non-invasive hyperglycaemia detection system based on electrocardiograms. The proposed hyperglycaemia detection system is presented in Figure 2.11. In general, input of the detection are ECG signals, which are then fed into the classification unit, including the computational intelligence. Computational intelligent algorithms play the most important role in the strategy of hyperglycaemia detection in order to achieve good performance of the classification. This classifier detects output of either hyperglycaemic, or normoglycaemic states. This thesis focuses on the developments of artificial neural network (ANN) for the

ECG-based hyperglycaemia classification. In biomedical areas, artificial neural networks (ANN) have been proved to be a powerful computational intelligence for classification models and pattern recognition. Neural networks with the ability to learn by example achieve flexibility and effectiveness in medical diagnosis. They have been widely used because of their high performance in complex situations.

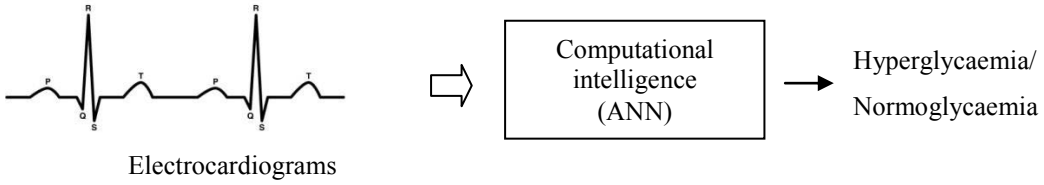


Figure 2.11: ECG-based hyperglycaemia detection system using ANN

At the beginning, capability of basic ANN was investigated using inputs of ECG parameters and compared with other classification techniques such as K-nearest neighbours and Linear discriminant analysis. Later advanced approaches attempted to improve the performance of hyperglycaemia detection. Principal component analysis (PCA) is proposed to reduce dimensionality of inputs and search for the appropriate ECG parameters. The integration of PCA and ANN is employed for classifying hyperglycaemia based on a subgroup of chosen datasets. Moreover, particle swarm optimisation (PSO), a heuristic optimisation technique, can potentially enhance classification performance by obtaining global optimal and disposing of local minima. PSO technique could be an effective alternate training algorithm for ANNs since it is found to be more accurate when compared to the existing conventional algorithms (Ch *et al.*, 2012). Also, impact of maximum velocity and acceleration constants of PSO on ANN convergence determines the best possible values of ANN's parameters to minimize classification error. The proposed PSO trained ANN is developed for hyperglycaemia classification to increase performance of the classification and the results are compared with the other existing gradient algorithms.

CHAPTER 3 . ELECTROCARDIOGRAPHIC BASED HYPERGLYCAEMIA DETECTION EMPLOYING ARTIFICIAL NEURAL NETWORK

3.1 INTRODUCTION

Artificial neural networks (ANN) have been widely applied in many scientific areas during the past few years. Already, they have been successfully used in various branches of biomedical systems, especially in medical diagnosis. Medical diagnosis using ANNs is currently a very active research area in medicine and it is believed that it will be more widely used in biomedical systems in the next few years. ANNs provide a powerful tool to help doctors to analyse, model and make sense of complex clinical data across a large number of medical applications. Most applications of ANNs to biomedical systems are classification problems; those are extensively utilised in pattern recognition to detect diseases or complex medical conditions. The task is on the basis of the measured features, which are extracted

from a background of irrelevant detail, to assign the patient to one of a small set of classes.

ANNs are ideal in recognizing diseases using scans since there is no need to provide a specific algorithm on how to identify the disease. Neural networks learn by example so the details of how to recognise the disease is not needed. In many instances, ANNs have proven more robust than other expert systems like rule-based or object-oriented systems in dealing with noise environments, and have been easier to train since they do not require intricate and potentially unreliable heuristics. Obviously, ANNs can achieve high performance and computation rates, which are vital in biomedical applications.

This chapter presents an ANN classification model for hyperglycaemia detection, which is based on ECG parameters. Multilayer feed-forward network is chosen to estimate the glycaemia states from T1DM data, because it is the most suitable network for generating nonlinear relationships for a given problem. The input of ANN classification consists of 16 original ECG variables which are extracted from ECG signals. The model provides output of binary glycaemic states (1 for hyperglycaemia or 0 for normoglycaemia) according to the data set fed into the ANN classification unit. The construction of the feed-forward ANN classification model is presented as follows. At first, an experimental environment is set up to generate the ECG data set which is used throughout this PhD research for hyperglycaemia detection models. It includes the subjects of clinical study, ECG acquisition, feature extraction, ECG parameters obtained and ECG statistical analysis. After that, methodology of feed-forward ANN classification model for hyperglycaemia detection is introduced. Five training algorithms are selected for ANN classification and their performances on hyperglycaemia detection are presented respectively. Furthermore, the influences of ANN training algorithms on the performance of hyperglycaemia detection are also discussed at the end of the chapter.

3.2 ARTIFICIAL NEURAL NETWORK FOR HYPERGLYCAEMIA DETECTION

The development of hyperglycaemia detection model based on ANN is presented in Figure 3.1. There are three main components in the strategy, including ECG acquisition, ECG feature extractions and ANN classifier. Firstly, the ECG signal is measured and acquired from the body of patients. In the second part, feature extraction component is developed to determine and obtain specific ECG parameters, such as *PR* and QT_C intervals, etc. A HRV analysis tool has been also utilised to generate certain necessary HRV measures. All ECG parameters are then fed into the ANN classifier. The output of ANN classifier is a binary glycaemic level (hyperglycaemia or normoglycaemia), which indicates a random episode to be under the state of hyperglycaemia or normoglycaemia.

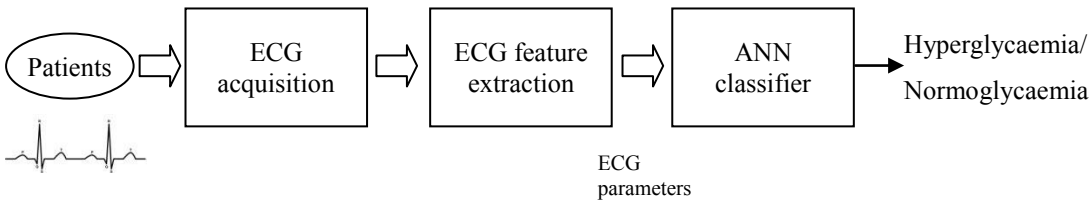


Figure 3.1: General structure of hyperglycaemia detection based on ANN and ECG parameters

3.2.1 ECG acquisition

3.2.1.1 Subjects

This thesis aims to explore the effects of natural occurrence of hyperglycaemic events on ECG parameters. Glycaemic levels were used to classify glycaemic states, normally considered as having low blood glucose levels or a hypoglycaemic state at the level of BGL less than or equal 3.33 mmol/l, as having normoglycaemic state when BGL varies from 3.33 mmol/l to 8.33 mmol/l and as having high blood glucose levels or a hyperglycaemic state at the level of BGL more than or equal 8.33 mmol/l. To prevent the interference of the hypoglycaemic state that may weaken the evaluation of the influence of hyperglycaemia in changes of

cardiac ventricular intervals, this thesis only focuses on subjects with normal and high blood glucose levels.

To perform, the database is based on the adolescent patients with T1DM (between ages of 12 to 18 years old) who volunteered for an overnight hypoglycaemia and hyperglycaemia study at the Princess Margaret Hospital for Children in Perth, Australia. In this thesis, the BGL threshold to define hyperglycaemia is set at 8.33 mmol/l (150 mg/dl), which is a high enough level for worsening outcomes (Hirshberg *et al.*, 2008). Ten Type 1 diabetic patients, who mostly had two glycaemic states, hyperglycaemia (BGL $\geq$ 8.33 mmol/l) and normoglycaemia (3.33 mmol/l < BGL < 8.33 mml/l), are chosen for collecting data in this thesis. All data were collected with approval of the Women's and Children's Health Service, Department of Health, Government of Western Australia, and with written informed consent. A comprehensive patient information and consent form was formulated and approved by the Ethics Committee. For the children participating in this study, the parent or guardian signed the relevant form.

3.2.1.2 Performance of ECG acquisition

A Compumedics system, which was the Siesta, was used to continuously measure and record electrocardiographic (ECG) signals every minute during the night with the sampling rate of 512 Hz. This study used single lead which was lead-II. The Siesta provides amplified channels for physiological signal collection. The software of Profusion PSG was installed in the PC to operate the Siesta for ECG acquisition. Data from this medical device are all clean signals to make sure artefacts and ectopic beats have been eliminated. The obtained ECG signals were exported into text files (*.txt). These files were then processed in the next stage of feature extraction for further required ECG parameters.

Actual blood glucose levels (BGLs) were routinely collected as reference using Yellow Spring Instruments (www.ysi.com). BGL was sampled at every 30-minute duration from 9 pm to 6 am corresponding to measured ECG parameters, approximately 17-20 data points were used for each patient. The results were then exported to text format files. None of the diabetic patients showed clinical evidence

of cardiac autonomic function. Their blood glucose levels increased during the hyperglycaemic phase, as shown in their actual BGL profiles, Figure 3.2.

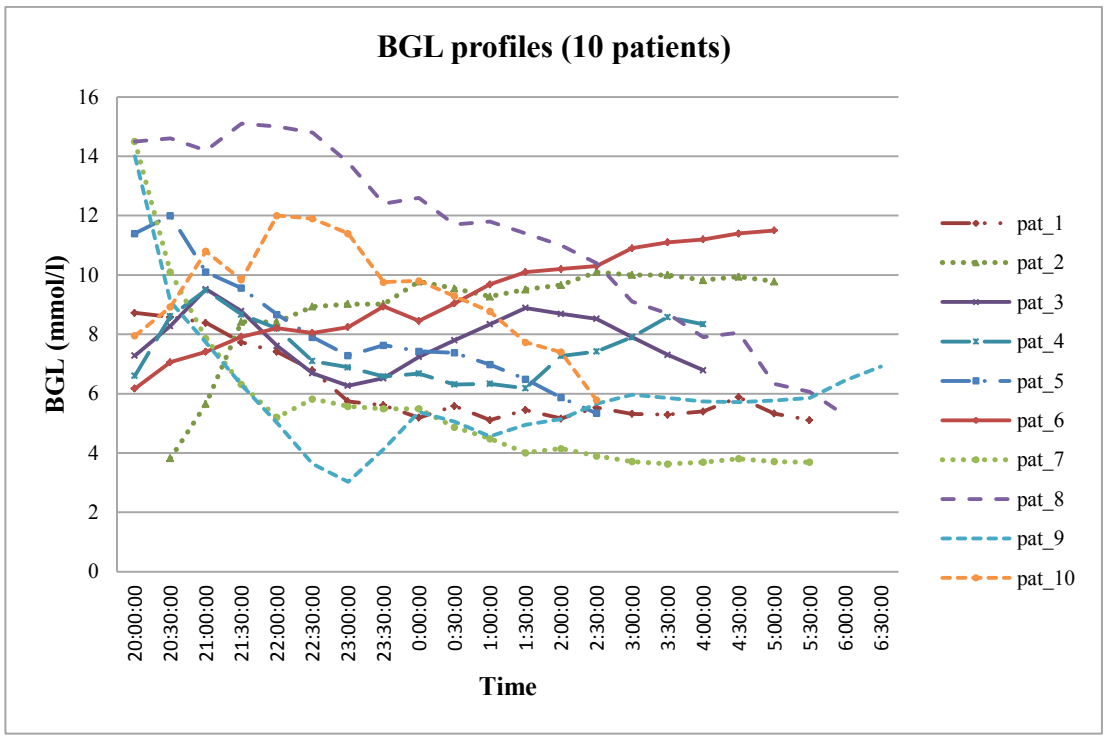


Figure 3.2: Actual blood glucose level profiles in 10 T1DM patients

3.2.2 ECG feature extraction

The ECG signals obtained from the ECG acquisition were processed in the feature extraction. This stage is aimed to extract an electrocardiogram to find the required ECG parameters. The ECG signals were processed using analysis software developed in LabVIEW Professional 2010. There were two main steps in the feature extraction, including ECG pre-processing and feature determination.

3.2.2.1 ECG pre-processing

After collecting data from the Compumedics system, to effectively remove noise effects on ECG signals (i.e., low frequency noise, baseline wander, power line interference, etc.), a pre-processing procedure was required and feature determination was then used to verify the ECG parameters. The algorithm that was

developed uses a digital Butterworth band-pass filter with high and low cut-off frequencies of 0.5 Hz and 40 Hz respectively. The baseline wandering was reduced in the filtered signal due to the removal of low frequency components, therefore improving signal-to-noise ratio. The electrocardiogram was then passed through a differentiator to enhance the signal (Figure 3.3).

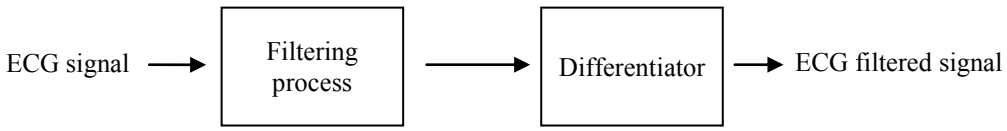


Figure 3.3: Block diagram of ECG pre-processing

3.2.2.2 Position of ECG characteristic points

ECG signal is characterized by several main points in a cycle length. Position of these characteristic points in the ECG signal was compulsory before performing feature determination. This description is to determine the positions of ECG characteristic points in time in an ECG recording. In this research, ECG parameters include ECG intervals and HRV measures. An ECG interval indicates time duration (in milliseconds) between two specific ECG points. HRV measures have been described in Table 2.1.

The position of ECG characteristic points is presented in Figure 3.4, which consists of:

- *P* : the beginning of *P*-wave
- *Q* : the beginning of QRS complex
- *R*-peak: the peak of QRS complex
- *S*: the end of QRS complex
- T_{peak} (T_P): the peak of *T*-wave
- T_{end} (T_E): the end of *T*-wave

3.2.2.3 ECG intervals

For the purpose of examining the effects of hyperglycaemia on ventricular repolarisation, the ECG intervals used in this work are as follows:

- *RR*: the interval between two continuous *R* peaks
- QT_C: the interval from the beginning of QRS complex (*Q*) to the end of *T*-wave (T_E)
- RT_C: the interval from the *R*-peak of QRS complex to the peak of *T*-wave (T_P)
- T_PT_{EC}: the interval from the peak of *T*-wave (T_P) to the end of *T*-wave (T_E)
- *PR*: the interval from beginning of *P*-wave (*P*) to the beginning of QRS complex (*Q*)

Index of *c* in the parameters indicates the correction by heart rate for the variables using the Bazett's formula 2.1, which is normalized using the square root of *RR* interval (the period between two continuous *R*-peaks).

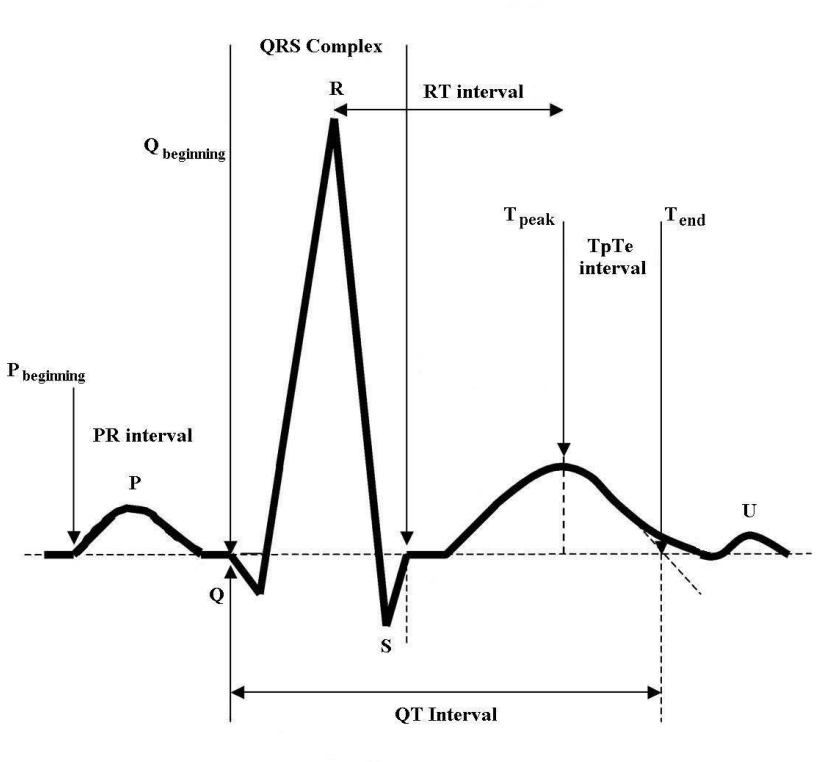


Figure 3.4: Schematic representation of normal ECG points and intervals

3.2.2.4 Feature determination

R peaks are found by setting a threshold in order to find a peak having higher amplitude than that value. Detection of peaks of *P*-waves, *Q beginnings* and peaks/end of *T*-waves are carried out by zero crossing method and window

thresholds. Once *R* peaks are found, these are taken as reference and the waves are scanned on both sides of the *R* peaks to get the zero crossing points for obtaining peaks of *P*-waves, *Q*, and peak/end of *T*-waves (T_P, T_E). To detect the beginning/end of the *P*-wave, an optimum window width of 100 ms is selected in which 50 ms are on either side of the peak of *P*-wave. Within the window, if the signal is not crossing an isoelectric line, then the minimum values are considered as beginnings and ends. *Q* points are one zero crossing behind *R* peaks. *T* peaks (T_P) are two zero crossings ahead of *R* peaks. *T* ends (T_E) are three zero crossings ahead of the *R* peaks. During the peaks/ends detections, signal annotation was also checked visually to eliminate the improper one. If a *U*-wave happened (which was very rarely seen in this study), followed by a *T*-wave before returning to baseline, the end of the *T*-wave was defined as the zero crossing point between *T*-wave and *U*-wave. The zero-crossing method has better advantages compared to other methods using ECG baseline (i.e., tangent method) that is not sensitive with ECG baseline drift. Figure 3.5 presented an example of two free-noise electrocardiogram signals with denoted significant ECG points, one is normoglycaemic ECG and one is hyperglycaemic ECG. The required ECG parameters have also been computed in LabVIEW program, which is shown in Figure 3.6. Compared to manual method, all features measurements have the error differences under 5%.

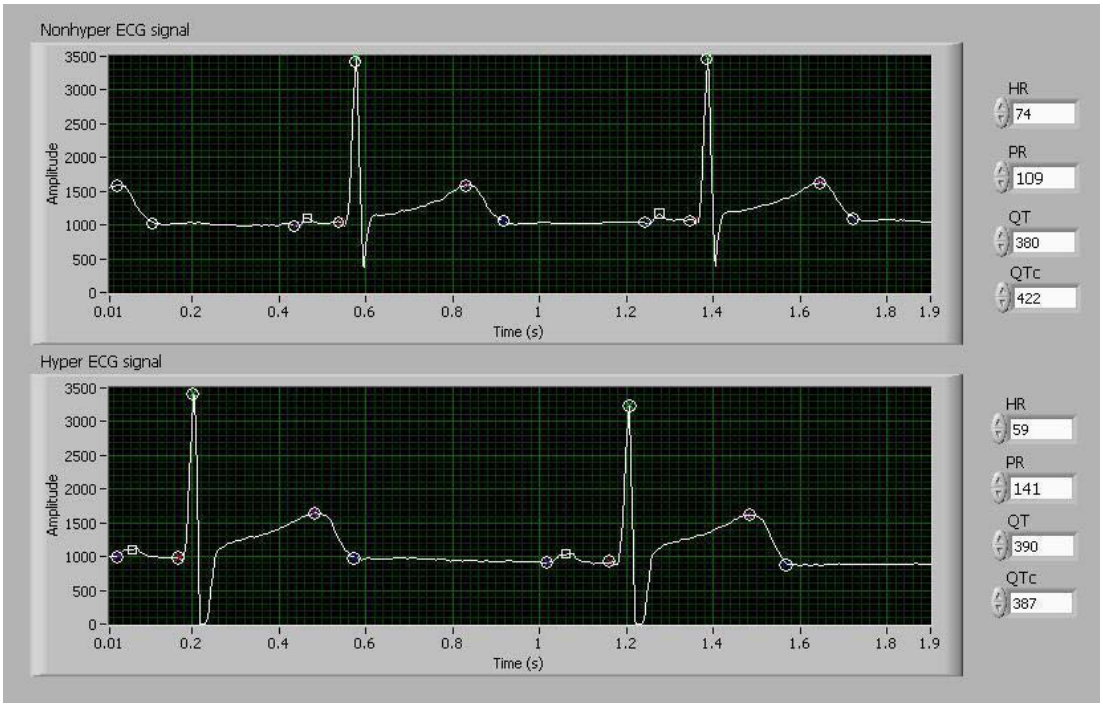


Figure 3.5: The ECG signals with denoted significant points, upper graph with non-hyper ECG and lower graph with hyper ECG

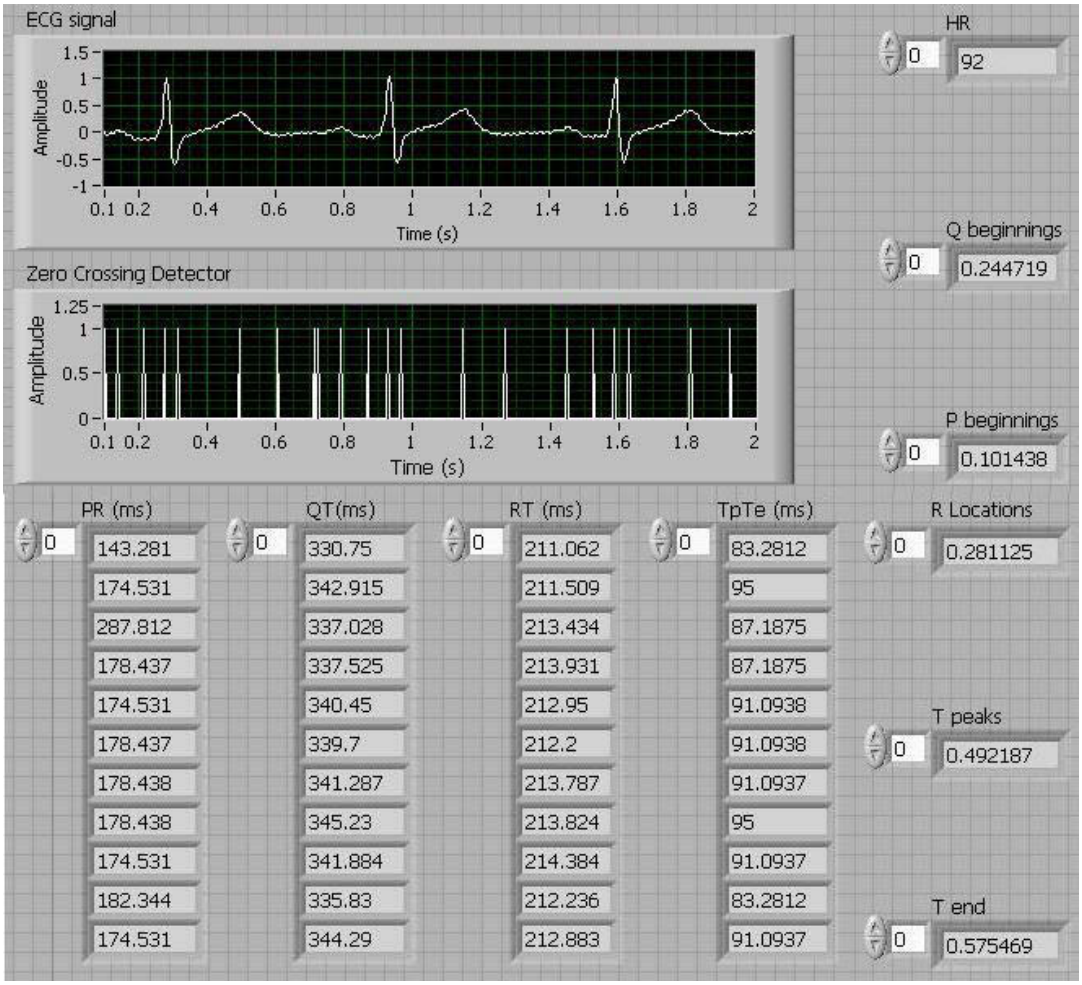


Figure 3.6: Processed ECG signals with zero crossing detector and calculated ECG parameters

3.2.2.5 HRV analysis tool

After extracting *R* peaks from raw ECG signals, RR interval time series were formed to analyse HRV. A selection of time domain and frequency domain measures of HRV were considered and performed following the Task Force guidelines (Electrophysiology, 1996). These HRV measures have been described in section 2.2.2.1 and are now again summarized as below:

Time domain measures:

- Mean RR interval (*MeanRR*)
- Standard deviation of the RR interval index (*SDNN*)
- Root mean square of successive RR interval differences (*RMSSD*)

- Percentage of beats with a consecutive RR interval more than 50 ms (*pNN50*)
- HRV triangular index (*HRVi*)
- Baseline width of the RR interval histogram evaluated through triangular interpolation (*TINN*)

Frequency domain measures:

- Very low frequency (*VLF*)
- Low frequency (*LF*)
- High frequency (*HF*)
- Total spectral power (*TotalPw*)
- Ratio between LF and HF components (*LF/HF*)

HRV analysis was conducted using the Kubios HRV analysis software package (http://kubios.uku.fi/). This software package was used for analysing the variability of heart beat intervals. It has an easy-to-use graphical interface that shows the HRV waveform and calculates the time-domain, frequency domain, and non-linear dynamics parameters from the raw RR signals. The software computed all required time and frequency domain features.

In the frequency domain analysis, the power spectrum is calculated either by using established Fast Fourier Transform (FFT) based method or parametric spectral estimation based on autoregressive time series modelling. Compared to autoregressive modelling, FFT based methods have been preferred for HRV analysis in diabetic patients (Chemla *et al.*, 2005). In this work, frequency domain analysis was done by FFT using Hanning window to derive spectral power. The intervals were linearly interpolated at the rate of 5 Hz to obtain evenly sampled values. The power components are usually expressed in absolute units (ms^2). The Kubios software generated a report sheet for each sample, as shown in Figure 3.7, including results of *RR* and heart rate distributions, time and frequency domain measures.

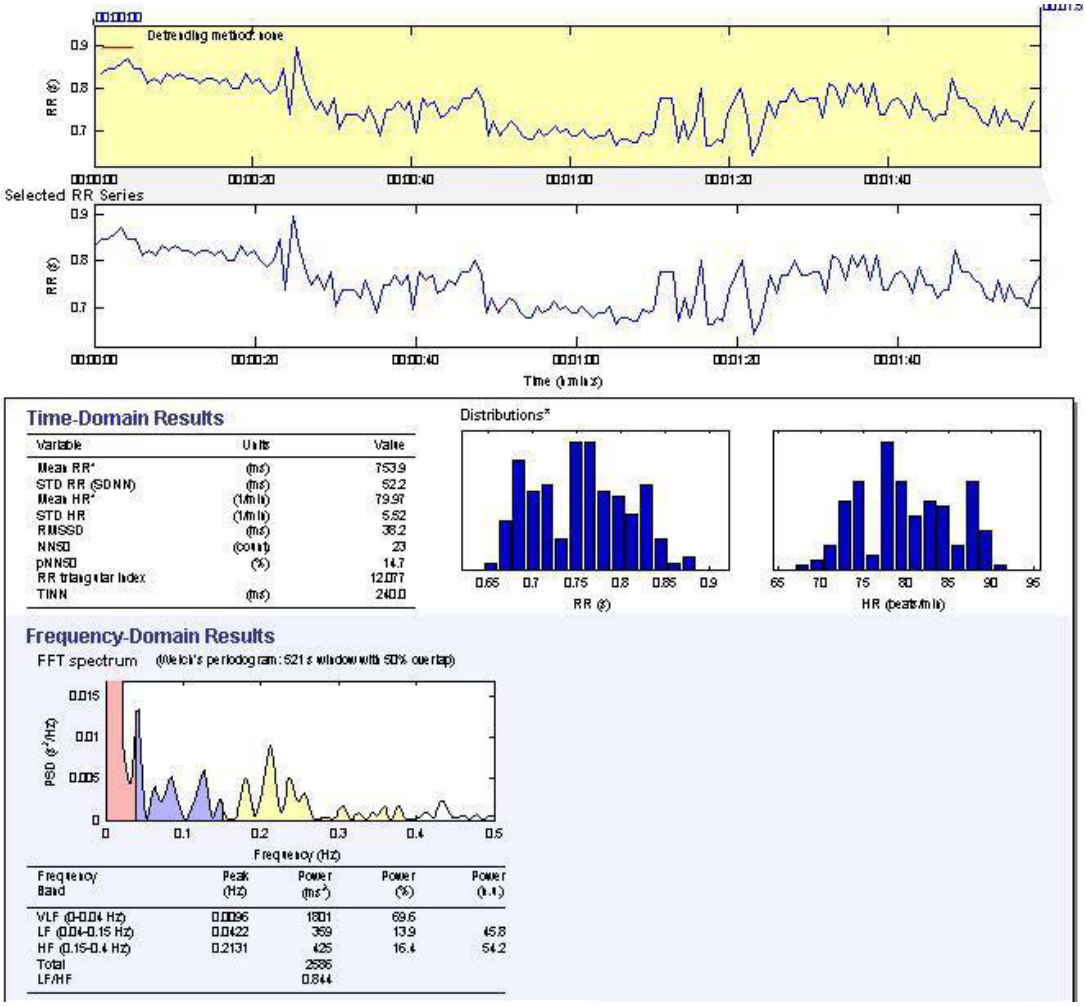


Figure 3.7: HRV analysis results in Kubios for 1 patient under hyperglycaemic condition

3.2.3 ANN classifier for hyperglycaemia detection

3.2.3.1 Neural network

An artificial neural network (ANN) is a computational model that attempts to account for the parallel nature of the human brain. An ANN is a network of highly interconnecting processing elements (neurons) operating in parallel. These elements are inspired by biological nervous systems. A neural network is a massively parallel-distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

(i) Knowledge is acquired by the network through a learning process,

(ii) Inter-neuron connection strengths known as synaptic weights are used to store the knowledge

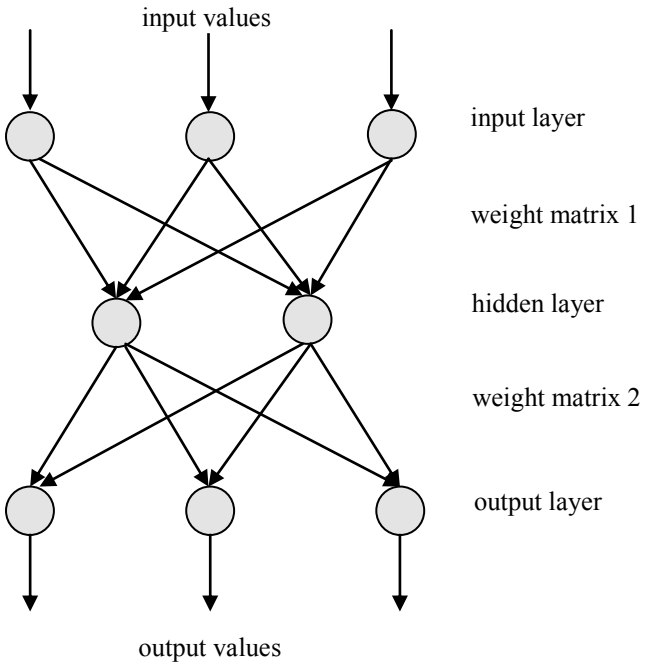


Figure 3.8: Neural network structure

In theory, neural networks can do anything a normal computer can do. We can train a neural network to perform a particular input which leads to a specific target output (Demuth H., 2001). The network is adjusted based on a comparison of the output and the target, until the network output matches the target, which is called supervised learning. Typically, many input/target pairs are used, in this supervised learning, to train a network.

In practice, neural networks have been trained to perform complex functions in various fields of application. As in nature, the connections between elements largely determine the network function. They are especially useful for signal classification. If there are enough training examples and enough computing resources it is possible to train a feed-forward neural network to perform almost any mapping to an arbitrary level of precision. A subgroup of processing element is called a layer in the network. The first layer is the input layer and the last layer is the output layer. Between the input and output layer, there may be additional layer(s) of units, called

hidden layer(s). Figure 3.8 represents the typical neural network. A neural network can be trained to perform a particular function by adjusting the values of the connections (weights) between elements.

The most common ANN architecture consists of many neurons organised in layers. Each neuron in a layer is connected to all the neurons of the next layer. It can always distinguish between input, hidden and output layers. There is only one input and one output layer whereas more than one hidden layer is allowed. The most common form of ANN is the multilayered neural network.

3.2.3.2 ANN architecture

In this study, the detection of hyperglycaemic state (BGL ≥ 8.33 mmol/l) using sixteen ECG variables is based on an advanced neural network algorithm developed from the obtained clinical data. This neural network has a multilayer feed-forward neural network structure with one input layer, one hidden layer and one output layer, which is shown in Figure 3.9.

In total, there are 16 ECG parameters employed as inputs in a multilayer neural network for hyperglycaemia detection in this thesis. They are:

(i) *HR*
(ii) QT_C
(iii) *PR*
(iv) RT_C
(v) T_PT_{EC}
(vi) *VLF*
(vii) *LF*
(viii) *HF*
(ix) *TotalPw*
(x) *LF/HF*
(xi) *MeanRR*
(xii) *SDNN*
(xiii) *RMSSD*
(xiv) *pNN50*

(xv) *HRVi*

(xvi) *TINN*

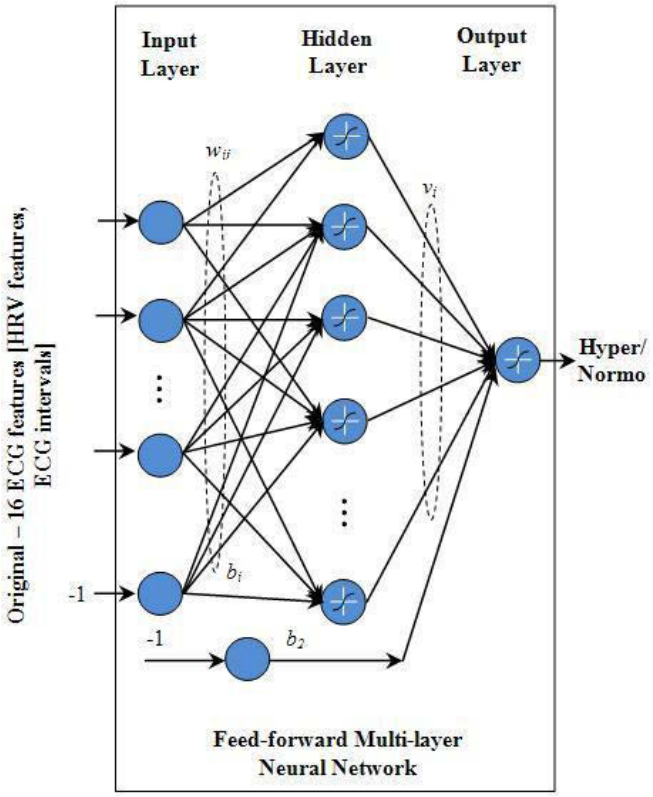


Figure 3.9: Hyperglycaemia detection feed-forward Multilayer neural network classifier and input of ECG parameters

Normalized values of ECG parameters, which have been shown previously in Table 3.2, are used for the input of ANN to reduce the patient-to-patient variability because of different scales from different ranges of features.

The structure of the neural network has 16 input nodes which are 16 mentioned ECG parameters above. The hidden nodes are varied from 3 to 20 in order to select the one that gives the best performance with the minimum mean square error (MSE). The training set, validation set and testing set are randomly selected with proportions of 35%, 35%, and 30% out of overall data, respectively. A validation set is utilised for determining early stopping of the network.

The output layer has one node which indicates the state of hyper or non-hyper. The input-output relationship of the proposed three-layer neural network for the output target y can be written as follows:

$$y = \sum_{i}^{H} w_i \tan sig\left(\sum_{j=1}^{N} v_{ij} - b_i \right) - b_2 \tag{3.2}$$

where H denotes the number of the hidden nodes; w_i, i= 1, 2, …, H is the weight of the link between i^{th} hidden nodes and the output; v_{ij}, i = 1,2,…, H; j = 1,2,…, N is the weight between the i^{th} hidden nodes and the j^{th} input; b_i, b_2 are the biases for i^{th} hidden nodes and output node, respectively; *tansig* denotes the hyperbolic tangent sigmoid transfer function used in both the hidden layer and the output layer:

$$\tan sig(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \tag{3.3}$$

3.2.3.3 Initialisation

To compare the performance of different networks, the same condition was kept in initialising the networks. The same training parameters and learning function were adopted during the training process.

The Nguyen-Widrow (Nguyen *et al.*, 1990) initialisation algorithm was used in this research. This algorithm chooses values in order to distribute the active region of each neuron in the layer evenly across the layer's input space. It generates initial weight and bias values for a layer, so that the active regions of the layer's neurons will be distributed roughly evenly over the input space. The Nguyen-Widrow initialisation algorithm has advantages over purely random weights and biases with few neurons wasted (all the neurons are in the input space) and the training works faster (each area of the input space has neurons).

3.2.3.4 Training algorithms

To accelerate the convergence of the error back propagation learning method, the ANN is trained by the Levenberg-Marquardt (LM) algorithm, which appears to be a fast and effective training algorithm. In general, The LM algorithm estimates the

second directional derivative of the error function, in order to direct the training process to a local minimum and find the optimised network parameters (i.e, w_i, v_{ij}, b_i, b_2).

Training parameters were chosen delicately. Maximum epoch was 1000. Minimum gradient limited was 0.0001 and goal was set to 0.001. Training networks were done to evaluate the performance of the algorithms after all details of the algorithms and parameters had been finalised.

The overall data set consists of a training set, validation set, and testing set. The validation set is used as an early stopping method to make sure the ANN does not over-train. The testing set is then used to test the generalization of the neural network. For these, the whole data set which includes both hyperglycaemic data part and normoglycaemic data part were used.

Moreover, five different training methods were also applied to the neural network model to compare their performances with the proposed one. Same input and structure were used with different algorithms: Levenberg-Marquardt, Gradient descent with momentum, Scaled conjugate gradient, Resilient back-propagation, Conjugate gradient back-propagation with Fletcher-Reeves updates, which are called LM, GDX, SCG, RP and CGF, respectively.

(a) Levenberg-Marquardt

Like the quasi-Newton methods, the Levenberg-Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares (as is typical in training feed-forward networks), then the Hessian matrix can be approximated as:

$$H = J^T J \tag{3.4}$$

and the gradient can be computed as

$$g = J^T e \tag{3.5}$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and e is a vector of network errors. The Jacobian matrix can be computed through a standard back-propagation technique that is much less complex than computing the Hessian matrix.

The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \tag{3.6}$$

When the scalar μ is zero, this is just Newton's method, using the approximate Hessian matrix. Newton's method is faster and more accurate near an error minimum, so the aim is to shift towards Newton's method as quickly as possible. In this way, the performance function will always be reduced at each iteration of the algorithm.

This algorithm appears to be the fastest method for training moderate-sized feed-forward neural networks (up to several hundred weights). It also has a very efficient MATLAB implementation, since the solution of the matrix equation is a built-in function, so its attributes become even more pronounced in a MATLAB setting.

(b) Conjugate gradient with Fletcher-Reeves update

Conjugate Gradient Algorithms are a basic back-propagation algorithm adjusting the weights in the steepest descent direction (negative of the gradient). This is the direction in which the performance function is decreasing most rapidly. It turns out that, although the function decreases most rapidly along the negative of the gradient, this does not necessarily produce the fastest convergence. In the conjugate gradient algorithms a search is performed along conjugate directions, which produces generally faster convergence than steepest descent directions (Demuth H., 2001).

Fletcher-Reeves Update

All of the conjugate gradient algorithms start out by searching in the steepest descent direction (negative of the gradient) on the first iteration.

$$p_0 = -g_0 \tag{3.7}$$

A line search is then performed to determine the optimal distance to move along the current search direction:

$$x_{k+1} = x_k + \alpha_k p_k \tag{3.8}$$

Then the next search direction is determined so that it is conjugate to previous search directions. The general procedure for determining the new search direction is to combine the new steepest descent direction with the previous search direction:

$$p_k = -g_k + \beta_k p_{k-1} \tag{3.9}$$

For the Fletcher-Reeves update the procedure is:

$$\beta_k = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \tag{3.10}$$

This is the ratio of the norm squared of the current gradient to the norm squared of the previous gradient.

The conjugate gradient algorithms are usually much faster than variable learning rate back-propagation, and are sometimes faster than resilient back-propagation, although the results will vary from one problem to another. The conjugate gradient algorithms require only a little more storage than the simpler algorithms, so they are often a good choice for networks with a large number of weights.

(c) Scaled Conjugate Gradient

The scaled conjugate gradient (SCG) algorithm is based on conjugate directions, but this algorithm does not perform a line search at each iteration. SCG algorithm is a variation of a standard conjugate gradient algorithm. The major idea of conjugate gradient is that they up to second order produce non-interfering directions of search. This means that minimization in one direction d_t followed by minimization in another direction d_{t+1} imply that the error has been minimized over the whole subspace spanned by d_t and d_{t+1}. The search directions are given by:

$$d_{t+1} = -E'(w_{t+1}) + \beta_t d_t \tag{3.11}$$

where w_t is a vector containing all weight values at time step t and β_t is

$$\beta_t = \frac{|E'(w_{t+1})|^2 - E'(w_{t+1})^T E'(w_t)}{|E'(w_t)|^2} \tag{3.12}$$

In the standard conjugate gradient algorithms the step size ϵ_t is found by a line search which can be very time consuming because this involves several calculations of the error and the first derivative. The step size is estimated by a scaling mechanism thus avoiding the time consuming line search. The step size is given by:

$$\epsilon_t = \frac{-d_t^T E'(w_t)}{d_t^T s_t + \lambda_t |d_t|^2} \tag{3.13}$$

where s_t is

$$s_t = \frac{E'(w_t + \sigma_t d_t) - E'(w_t)}{\sigma_t} \quad , \qquad 0 < \sigma_t \leq 1 \tag{3.14}$$

ϵ_t is the step size that minimizes the second order approximation to the error function. s_t is one sided difference approximation of $E''(w_t)d_t$. λ_t is a scaling parameter whose function is similar to the scaling parameter found in Levenberg Marquardt method. The calculation work per iteration for SCG can be shown to be in the order of two times the calculation work quick-prop algorithm. SCG contains no dependent parameters (Müller *et al.*, 1995).

3.2.3.5 Matlab Toolboxes

All ANN models in this thesis will be developed and run by the Neural Network Toolbox (Matlab). Neural network toolbox is a very powerful tool that is commonly used in applications where formal analysis would be difficult or impossible, such as pattern recognition and non-linear system identification. The Neural Network Toolbox provides a comprehensive support for many proven network paradigm as well as an inclusive set of training and learning functions. It supports modular network representation, allowing an unlimited number of inputs setting layers and network interconnections that enable the experiment to design and

manage the networks. The toolbox also supplies pre- and post-processing functions for improving network training and assessing network performance.

Moreover, Classification toolbox (version 3.1) in Matlab has also been used to perform alternative classification techniques in order to compare the performances of hyperglycaemia detection with ANN results. Matlab Classification Toolbox contains implementations of the various classifiers, including supervised and unsupervised classification algorithms. This toolbox allows users to compare classifiers across various data sets. Linear Discriminant Analysis and K-nearest neighbours have been chosen to test the classification of hyperglycaemia.

3.3 EXPERIMENTAL RESULTS

3.3.1 ECG parameters obtained from the study

There are 16 ECG parameters obtained from ten T1DM patients for hyperglycaemia detection in this thesis. These are a combination of ECG intervals and HRV measures, which are generally indicated following:

(i) *HR*
(ii) QT_C
(iii) *PR*
(iv) RT_C
(v) T_PT_{EC}
(vi) *VLF*
(vii) *LF*
(viii) *HF*
(ix) *TotalPw*
(x) *LF/HF*
(xi) *MeanRR*
(xii) *SDNN*
(xiii) *RMSSD*
(xiv) *pNN50*

(xv) HRVi

(xvi) TINN

The dataset from 10 T1DM patients consists of 143 data points with 2 classes, hyperglycaemia and normoglycaemia. Among those, 56 samples are hyperglycaemia and 87 samples are normoglycaemia. The dataset is stored in a text file in the form of 16x143 arrays, the rows being the ECG parameters and the columns being the samples.

An independent *t*-test was applied to every parameter to estimate the significant differences between hyperglycaemia and non-hyperglycaemia conditions. Moreover, Pearson's correlation analyses were used to evaluate and test the strengths of association between BGLs and variables in time and frequency domains. All statistical analyses were conducted with IBM SPSS version 19 (SPSS Inc, Chicago, IL, USA). Significance value (*p*-value) less than 0.05 is considered to be significant. Results are presented as mean ± standard errors. Table 3.1 shows the changes of 16 natural ECG parameters under hyperglycaemic conditions compared to normoglycemic conditions with corresponding *p* values.

To reduce patient-to-patient variability and dealing with parameters with different scales, normalisation was performed by using rescaling on the data set, which scales all numeric variables in the range [0,1]. Suppose that a *d*-dimensional ECG parameter vector *x* at time *i* is x_i. Its normalization x_i^* will be transformed to have normal distribution with its maximum value and minimum value as defined by the equation below:

$$x_i^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}};\ i = 1,2,...,d; \tag{3.1}$$

where x_{max} denotes the maxima and x_{min} represents the minima among all the data points. After normalisation, all parameters have the same scale for further comparison between them. Normalised data are presented in Table 3.2.

ECG parameters	Hyperglycaemia	Normoglycaemia	*p* value
HR	69.089 ± 7.209	70.379 ± 8.407	0.346
QT_C	394 ± 18.172	402.701 ± 16.112	**0.003**
PR	119.732 ± 11.887	112.299 ± 11.92	**< 0.0001**
RT_C	264.607 ± 16.891	273.299 ± 13.897	**0.001**
$TpTe_C$	90.143 ± 7.506	93.149 ± 8.344	**0.03**
VLF	208.879 ± 242.913	803.548 ± 1713.039	**0.001**
LF	531.570 ± 682.075	1496.638 ± 3359.613	**0.006**
HF	999.238 ± 1263.431	1929.816 ± 2430.304	**0.002**
TotalPw	1739.691 ± 1801.946	4250.792 ± 6246.882	**< 0.0001**
LF/HF	0.881 ± 0.932	1.047 ± 1.509	0.455
MeanRR	857.037 ± 81.288	879.670 ± 66.095	0.077
SDNN	54.043 ± 22.482	73.966 ± 38.544	**< 0.0001**
RMSSD	67.658 ± 30.550	87.034 ± 38.801	**0.002**
pNN50	30.622 ± 17.562	38.471 ± 15.776	**0.005**
HRVi	5 ± 1.783	5.886 ± 1.882	**0.005**
TINN	173.036 ± 77.252	222.598 ± 86.525	**< 0.0001**

Table 3.1: Changes of natural 16 ECG parameters under hyperglycaemic conditions compared to normoglycaemic conditions

ECG parameters	Hyperglycaemia	Normoglycaemia	*p* value
HR	0.313 ± 0.160	0.342 ± 0.187	0.346
QT_C	0.438 ± 0.204	0.536 ± 0.181	**0.003**
PR	0.572 ± 0.229	0.429 ± 0.229	**< 0.0001**
RT_C	0.433 ± 0.231	0.552 ± 0.190	**0.001**
$TpTe_C$	0.27 ± 0.167	0.337 ± 0.185	**0.03**
VLF	0.016 ± 0.018	0.06 ± 0.129	**0.001**
LF	0.016 ± 0.022	0.047 ± 0.108	**0.006**
HF	0.077 ± 0.101	0.151 ± 0.193	**0.002**
TotalPw	0.032 ± 0.035	0.080 ± 0.120	**< 0.0001**
LF/HF	0.092 ± 0.103	0.111 ± 0.167	0.455
MeanRR	0.475 ± 0.245	0.543 ± 0.2	0.077
SDNN	0.128 ± 0.085	0.203 ± 0.146	**< 0.0001**
RMSSD	0.220 ± 0.131	0.303 ± 0.167	**0.002**
pNN50	0.368 ± 0.211	0.462 ± 0.189	**0.005**
HRVi	0.469 ± 0.167	0.552 ± 0.176	**0.005**
TINN	0.393 ± 0.176	0.506 ± 0.197	**< 0.0001**

Table 3.2: Changes of normalised 16 ECG parameters under hyperglycaemic conditions compared to normoglycaemic conditions

The responses from 10 T1DM patients exhibited significant changes during the hyperglycaemic state against normoglycaemic state in both real data shown in Table 3.1 and normalised data in Table 3.2. This comparison resulted from a *t*-test, presenting ECG parameters in the form of (mean ± SD) with the associated *p*-values. It can be seen that thirteen out of sixteen parameters in hyperglycaemia significantly differed from those in normoglycaemia. There was no change in *HR*, while LF/HF

ratio and *MeanRR* were reduced but did not achieve statistical significance. Further details of these findings have been outlined in our previous studies (Nguyen *et al.*, 2012; Nguyen *et al.*, 2013). This result was expected since these altered ECG parameters are highly correlated with blood glucose, indicating that a rise in blood glucose levels is associated with a drop in QT_C, RT_C, $TpTe_C$ and HRV measures.

3.3.2 Performance of hyperglycaemia detection using ANN

3.3.2.1 Measuring the performance

To measure the performance of the classification results, sensitivity and specificity are use (Altman *et al.*, 1994). The definitions of sensitivity and the specificity are given as follows:

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.15}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3.16}$$

where *TP* (True Positive) is the number of hyperglycaemic events which are correctly classified as hyperglycaemia; FN (False Negative) is the number of hyperglycaemic events which are wrongly classified as normoglycaemia; TN (True Negative) is the number of normoglycaemic events which are correctly classified as non-hyperglycaemia; FP (False Positive) is the number of normoglycaemic events which are wrongly classified as hyperglycaemia. For the comparisons of the performances, a geometric mean *gm* was also used. The geometric mean equals to the square root of the multiplication of sensitivity and specificity.

$$gm = \sqrt{Sensitivity.Specificity} \tag{3.17}$$

Satisfactory results of sensitivity and specificity for disease diagnosis are required as follows:

- Sensitivity ≥ 70%
- Specificity ≥ 50%
- Sensitivity ≥ Specificity

3.3.2.2 ANN classification results

The neural networks applying five algorithms, namely LM, SCG, GDX, CGF, and RP used all 16 original ECG parameters as inputs. One performance of the error cycles is shown in Figure 3.10. MSE of training set decreased smoothly while the validation set dropped from the beginning up to 3 epochs then it started to increase continuously. Therefore, the training of the network was stopped at 3 epochs as the validation performance has started to increase continuously. This is how the validation set is used to prevent over-training of the ANN (Tetko *et al.*, 1995).

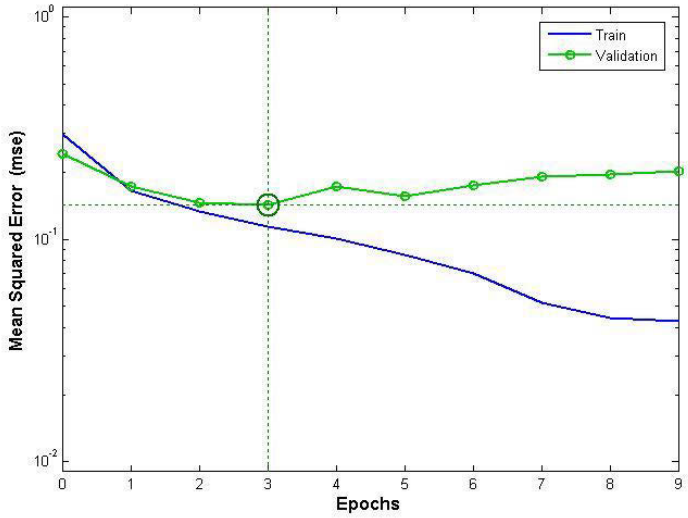


Figure 3.10: Training and validation error cycles of hyperglycaemia classification using ANN

The training performance of proposed detection has been analysed by means of ROC curves for five different algorithms in which the sensitivity (true positive rate) and the 1- specificity (false positive rate) are plotted in Figure 3.11. LM achieved the best corresponding *AuR* (area under ROC curve) of 0.8742 for the training set with input of all sixteen parameters, while the other comparison algorithms, GDX, SCG, RP, CGF only have 0.7726, 0.7980, 0.7740 and 0.7869, respectively. These ROC curves were used to find the optimal cut-off point to classify hyperglycaemic and normoglycaemic states. In order to analyse the optimum sensitivity and specificity, the cut-off point is selected at the point producing the sensitivity higher than 75%.

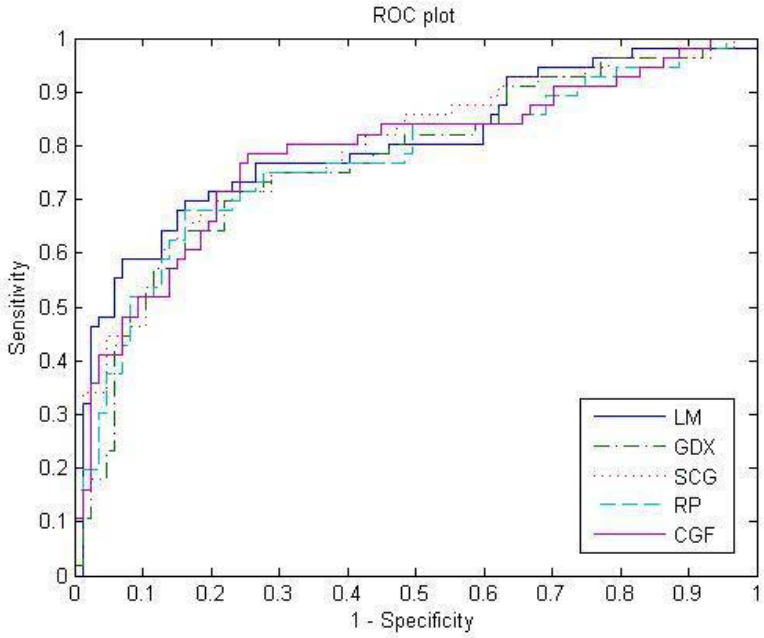


Figure 3.11: ROC curve: Sensitivity vs. 1-Specificity

At the defined cut-off point, the average (mean) training, validation and testing results in terms of sensitivity, specificity and *gm* are analysed and listed in Table 3.3 in which mean values are calculated by averaging over 20 runs. It can be seen in Table 3.3, the average (mean) testing result of ANN using LM algorithm with 16 inputs is satisfactorily found by giving the best sensitivity and specificity (70.37% and 62.58%) compared with other four algorithms such as GDX, SCG, RP and CGF whose mean sensitivity and specificity are (50.13% and 72.33%), (69.21% and 61.6%), (53.75% and 76.26%) and (59.95% and 64.65%), respectively.

Algo-rithms	Training			Validation			Testing		
	Sen (%)	Spec (%)	*gm* (%)	Sen (%)	Spec (%)	gm (%)	Sen (%)	Spec (%)	*gm* (%)
LM	88.43	75.42	80.86	70.33	68.35	67.71	70.37	62.58	65.05
GDX	63.93	80.27	70.07	50.13	75.77	60.17	50.13	72.33	58.57
SCG	83.65	67.94	72.98	70.34	66.31	64.41	69.21	61.6	62.24
RP	74.3	81.82	76.18	62.48	79.13	68.87	53.75	76.26	61.88
CGF	83.09	72.09	76.29	64.99	73.3	66.52	58.95	64.65	59.52

Table 3.3: Mean values of training, validation and testing results using five different algorithms

Algo-rithms	Hid. node	Train-ing *AuR*	Training			Validation			Testing		
			Sen (%)	Spec (%)	*gm* (%)	Sen (%)	Spec (%)	gm (%)	Sen (%)	Spec (%)	*gm* (%)
LM	9	0.8742	89.47	83.87	86.63	80	77.41	78.69	70.59	65.38	67.94
GDX	12	0.7726	78.57	81.82	80.18	75	61.76	68.06	66.67	54.84	60.47
SCG	10	0.7980	72.22	68.75	70.46	90.91	53.57	69.79	68.75	59.26	63.83
RP	11	0.7740	88.89	68.75	78.17	72.73	75	73.86	62.5	62.96	62.73
CGF	7	0.7869	82.35	81.82	82.09	71.42	86.21	78.47	55.56	68	61.47

Table 3.4: Best performance of hyperglycaemia detection using 16 ECG parameters and five different algorithms

The best performances of hyperglycaemia detection using five algorithms are presented in Table 3.4. Using LM algorithm, 9 hidden nodes produces the best classification in which the training results of 89.47% sensitivity and 83.87% specificity are gained and the testing set leads to 70.59% sensitivity and 65.38% specificity. These testing sensitivity and specificity are considered to be satisfactory result which meets the requirements of disease diagnosis (sensitivity $\geq$ 70% and

specificity ≥ 50%). The best testing geometric mean of 67.94% by using LM algorithm is also achieved with these sensitivity and specificity. Apart from LM, the sensitivities of the other four algorithms are less than 70%, which are not directed in biomedical application.

3.3.3 Comparison with other methods

Hyperglycaemia detection has been performed with other methods to compare with the performance of proposed ANN model. Linear Discriminant Analysis (LDA) and K-Nearest Neighbours (KNN), two of the most conventional and common classification techniques, have been chosen for the comparison. The methods are applied to the same ECG dataset from 10 diabetic patients. LDA and KNN classifiers have been run in training, validation and testing sets with proportions of 35%, 35%, and 30% out of overall data, respectively. The performances have been undertaken by using the Classification Toolbox (version 3.1) in Matlab.

LDA and KNN can easily handle the case where the within-class frequencies are unequal and their performance has been examined on randomly generated test data. LDA can provide class separability and draw a decision region between given classes to help to better understand the distribution of feature data (Balakrishnama S. *et al., 1998*). KNN is a nonparametric lazy learning algorithm, which means it does not make any assumptions on the underlying data distribution. It appears to be a useful technique in classification as well as regression, as most of the practical data does not obey the typical theoretical assumptions made (eg. Gaussian mixtures, linearly separable, etc.). To ensure the stability of the classifier, cross-validation has been performed in both classification techniques.

Different numbers of folds for cross-validation have been tested in order to discover the best results in hyperglycaemia detection. Finally, LDA using 3-fold and KNN using 4-fold cross validation with venetian blinds achieved the best outcomes, as shown in Table 3.5. Compared to the performance of ANN model using Levenberg-Marquardt for hyperglycaemia detection, LDA and KNN still got worse results with lower sensitivities and their sensitivities were much less than their

specificities in all three sets of training, validation and testing. The best testing sensitivities of LDA and KNN are 50% and 43.75%, which are poorer than the sensitivity of ANN (70.59%). Even Specificities of LDA and KNN are higher than ANN's (77.78% and 85.19% vs. 65.38%), LDA and KNN do not satisfy with the requirements for disease diagnosis. In general, ANN outperforms LDA and KNN with higher achievement in overall accuracy.

Algo-rithms	Training			Validation			Testing		
	Sen (%)	Spec (%)	gm (%)	Sen (%)	Spec (%)	gm (%)	Sen (%)	Spec (%)	gm (%)
ANN	89.47	83.87	86.63	80	77.41	78.69	70.59	65.38	67.94
LDA	73.68	84	78.61	57.14	79.31	67.32	50	77.78	62.36
KNN	78.94	93.55	85.94	57.14	93.1	72.94	43.75	85.19	61.04

Table 3.5: Comparison of best performances of three different classification techniques for hyperglycaemia detection

3.4 DISCUSSION

Based on the findings in the relation of hyperglycaemia with electrical activities of the heart, the proposed hyperglycaemia detection has been presented. It employs ANN algorithm and the input of ECG parameters. The system consists of ECG acquisition, feature extraction and ANN classification. Output of ANN classifier displays a binary value representing a hyperglycaemic or normoglycaemic event. The input was electrocardiographic clinical data. The ANN model is used to indicate whether the ECG parameters in the input belong to hyperglycaemia or normoglycaemia. Primarily, this ANN model could be used for the classification of hyperglycaemia.

The ECG parameters are extracted from ECG signals which are acquired through lead II. Lead II is frequently used for *QT* interval measurement (Salvi *et al.*, 2011), where *U*-wave is less prominent to minimize the error in the determination the end of *T*-wave (Garson, 1993). A zero-crossing detector was used to mark the points

of *P, Q, R,* T_P and T_E. The zero-crossing method has better advantages compared to other methods using ECG baseline (i.e., tangent method) that is not sensitive with ECG baseline drift. To avoid the error of finding those points, manually checking on the results of delineation was performed after the implementation of the zero-crossing detector. Compared to manual method, all feature measurements have the error differences under 5%.

It is generally accepted that hyperglycaemia among persons with diabetes causes degradation of the microvasculature that results in central and peripheral autonomic neuropathy. Yet, there are numerous pathways whereby autonomic dysfunction could in turn affect insulin function and glucose regulation. Major organs, including the pancreas, liver, and skeletal muscle, which are responsible for insulin secretion, glucose production, and glucose metabolism, respectively, are innervated by autonomic fibres (Carnethon *et al.*, 2003). In this study, as shown in Table 3.1, hyperglycaemia was related to thirteen out of sixteen original ECG parameters (i.e, QT_C, *PR,* RT_C, T_PT_{EC}, etc.) and had no relation with *HR, LF/HF* ratio and *MeanRR*. Interestingly, a higher blood glucose level was associated with decreased QT_C, RT_C, $TpTe_C$, increased *PR* and reduced HRV measures.

Hyperglycaemia has been shown highly correlated to ventricular repolarisation intervals (RT_C, QT_C, $TpTe_C$) in this study. Obviously, the shortened QT_C led to the reduction of RT_C, $TpTe_C$. While the prolonged repolarisation in hypoglycaemia could be explained by a physiologically longer cardiac repolarisation at night and thus an electrically instable myocardium worsened by the hypoglycaemia induced adrenergic stimulation, there is currently no clear explanation for the reduced repolarisation in hyperglycaemia. Meanwhile, these parameters should be utilised further for hyperglycaemia detection.

Increased *PR* in hyperglycaemia has been discussed in the research of Marfella *et al.* (2000) with acute hyperglycaemia in a healthy man. It can be explained that this effect is due to PR intervals commonly associated with atrial fibrillation (Lorsheyd *et al.*, 2005; Homoud, 2009) and chronic hyperglycaemia, but not hypoglycaemia which may contribute to atrial fibrillation burden in several ways (Aksnes *et al.*, 2008; Gandhi *et al.*, 2005). Therefore, the hyperglycaemic state in

type 1 diabetic patients with both hypoglycaemic and hyperglycaemic states can result in changes of *PR* interval. However, our ECG assessments of diabetic patients did not indicate atrial dysfunction.

Recent studies have emphasized the importance of HRV in evaluation of the autonomic nervous system on heart rate. Our study demonstrated that it is possible to use time domain measures and power spectral techniques to analyse autonomic influence on the cardiac cycle length and ventricular repolarisation. In agreement with previous studies, high blood glucose levels caused lower HRV in subjects with diabetes. Our findings extended the previous observations made in selected T1DM patients and went across the spectrum of hyperglycaemic and normoglycaemic events. *SDNN* was much lower in hyperglycaemic events than in normoglycaemic events, which means that both the sympathetic and parasympathetic influence HRV. *SDNN* is probably the best known HRV index (Bigger *et al.*, 1992), mostly used as a time domain measure of HRV, because the additional information from other time domain parameters of HRV is relatively scant. In frequency domain measurements, *LF* can be used to quantify the sympathetic component of autonomic function in addition to *HF*, the parasympathetic component. Table 3.1 showed that both these two features were significantly decreased under hyperglycaemic conditions. Therefore, *SDNN, LF* and *HF* are potentially good markers to identify hyperglycaemic and non-hyperglycaemic states using HRV in T1DM.

The Levenberg-Marquardt (LM) algorithm is often the fastest back-propagation algorithm in training a moderate-sized feed forward neural network (up to several hundred weights). It is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms. A three-layer feed forward neural network training with LM using the whole data set for hyperglycaemia detection received reasonable classification results (as tabulated in Table 3.2) though it might take more time and higher hidden nodes. The proposed ECG classifier achieved best results (70.59% sensitivity, 65.58% specificity and 67.94% geometric mean) with LM algorithm. In cases where other algorithms are applied, the proposed ECG classifier still resulted in worse performances as their

sensitivities were less than 70%, which does not meet requirements in medical diagnosis.

Two other conventional classification methods, Linear discriminant analysis (LDA) and K-nearest neighbours (KNN), have been tested to compare their performances with ANN's using LM algorithm. The results show that ANN model using LM detects hyperglycaemia events more accurately than LDA and KNN. Although ANN, compared to LDA and KNN, lacks a comprehensive visual representation, it cannot be denied that ANN is a more effective classification technique for hyperglycaemia detection than those usual methods.

As far as we are concerned, this study is the first to report a non-invasive hyperglycaemia detection system using ECG signals. To predict hyperglycaemia, many other studies have been worked on in different ways using various input parameters, but none of them examined the abnormal ECG parameters under hyperglycaemia conditions. Most of the alternative studies focus on predicting blood glucose levels to forecast glycaemic states. For example, Le Compte (Le Compte *et al.*, 2010) built a dynamic model for blood glucose prediction to provide the most probable future insulin sensitivity. Other approaches developed neural networks for early detection of hypoglycaemia/hyperglycaemia events based on glucose data from continuous glucose monitoring (CGM) (Daskalaki *et al.*, 2012; Pappada *et al.*, 2010). Nevertheless, the purpose of these studies was on real-time glucose control rather than the disease diagnosis. By the way of hyperglycaemia diagnosis, classification models concentrate on keeping sensitivity and specificity high enough in balance, although specificity should be lower than sensitivity as the true positive rate (proportion of hyperglycaemic events are correctly detected) is more important than false positive rate (proportion of normoglycaemic events are correctly detected). These findings contribute to the limited literature of problems related to hyperglycaemia diagnosis based on physiological signals among T1DM. Moreover, this study draws attention to the need for further investigations to understand the mechanisms responsible for abnormal changes in ECG signals during a state of hyperglycaemia. It should also be assessed whether subclinical changes in ECG are

predictive of the subsequent development of clinical hyperglycaemia conditions in these patients.

On the other hand, the results using the ANN model in this chapter is still not high enough for an accurate hyperglycaemia detection system. The proposed strategy needs to be enhanced to improve the overall accuracy of hyperglycaemia detection. Sixteen ECG parameters took extensive memory-cost and slowed down the computation rate. Moreover, a data set of a large number of variables may cause the multicollinearity which leads to incorrect classification where some of those variables are somehow dependent on each other. A strategy of reducing the dimension of input size such as Principal Component Analysis should be carried out to establish a set of interrelated parameters and find important ECG parameters, which significantly contribute to the performance of hyperglycaemia detection.

In this chapter, the changes of ECG parameters associated with hyperglycaemic and normoglycaemic states in Type 1 diabetic patients have been explored. An ECG acquisition and feature extraction based on LabVIEW has been developed to detect these alterations. Several statistical analyses have been done to evaluate the effect of ECG parameters on hyperglycaemia and normoglycaemia. The study also showed the different findings of decreased ventricular repolarisation and reduced HRV measures in Type 1 diabetic patients under hyperglycaemic states. A classification unit using the artificial neural network was developed to determine the presence of hyperglycaemia episodes based on T1DM patients' ECG parameters. The results of 70.59% sensitivity and 65.58% specificity by using LM algorithm are considered reasonable and meet the requirement for biomedical application. Various training algorithms also have been developed and compared. In general, the proposed ECG classifier using LM algorithm outperforms other algorithms. This result indicates that hyperglycaemic events in T1DM can be detected non-invasively and effectively by using ECG signals and ANN approach. The best performance of ANN models for hyperglycaemia detection in the experiment was 67.94% in terms of geometric mean with 9 hidden nodes using LM algorithm. This performance was slightly slow and thus a more advanced ANN is necessary to achieve higher results.

CHAPTER 4 . COMBINED PRINCIPAL COMPONENT ANALYSIS AND NEURAL NETWORK FOR HYPERGLYCAEMIA DETECTION

4.1 INTRODUCTION

As mentioned in chapter 3, an ANN has been applied using ECG signals for hyperglycaemia detection in Type 1 diabetic patients. Although that result of ANN is satisfactory, performance of classifier for hyperglycaemia detection still needs to be improved to adapt into the practical environment. Chapter 4 aims to develop another approach of ANN for classifying the events of hyperglycaemia and normoglycaemia with better results. Sixteen original ECG parameters seem to be a large input dimensionality that may create more redundancy due to correlation among features and increase in uncertainty within input datasets. These factors, might lead to a significant decrease of classification accuracy as well as an increase in processing costs. Features which do not make any or very little contribution to the performance of classification should be removed. Hence, the application of dimensionality reduction techniques, which reduce the data dimensionality while preserving informative features, is very important in order to enhance the performance of the classifier.

Dimensionality reduction is one of the key techniques in pattern recognition which focuses on revealing meaningful structures and unexpected relationships in the multivariate data. The main aim of the dimensionality reduction algorithms is to obtain a compact, accurate representation of the data that reduces or eliminates statistically redundant components. There are obviously three major advantages in using dimensionality reduction techniques. Firstly, the performance of a classification might be improved by selecting a subset of a smaller number of informative and less correlated variables. This issue is important for recovering the generalisation capability of a training dataset. Secondly, the processing time will be reduced with the use of a smaller number of input features. Thirdly, the smaller datasets would be more appropriate in cases where there exists multicollinearity in original datasets because of direct relationships among HRV measures.

Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are the two popular independent feature extraction algorithms. Both of them extract features by projecting the parameter vectors into a new feature space through a linear transformation matrix. But they optimise the transformation matrix with different intentions. PCA optimises the transformation matrix by finding the largest variations in the original feature space; whereas LDA pursues the largest ratio of between-class variation and within-class variation when projecting the original feature space to a subspace. For the purpose of reducing the size of original inputs to diminish the computational requirement while retaining the variations present in the data set, PCA has been selected to use as an ultimate choice.

The main contribution of chapter 4 is to propose and to evaluate a combination of PCA and ANN in a two-stage classifier for hyperglycaemia detection. This chapter investigates PCA for dimensionality reduction and proposes the application of ANN for classification of hyperglycaemia episodes. ECG variables, which contribute significantly to the performance of hyperglycaemia detection, have also been revealed by using PCA. Five training algorithms are selected for ANN classification and their performances on hyperglycaemia detection are presented respectively. Furthermore, the influences of PCA and not using PCA on the performance of hyperglycaemia detection are also discussed at the end of the chapter.

4.2 PRINCIPAL COMPONENT ANALYSIS (PCA)

4.2.1 Overview of PCA

Principal components analysis (PCA) is one of a family of techniques for taking high-dimensional data, and using the dependencies between the variables to represent it in a more tractable, lower-dimensional form, without losing too much information. PCA is one of the simplest and most robust ways of doing such dimensionality reduction. It is also one of the oldest, and has been rediscovered many times in many fields, so it is also known as the Karhunen-Loève transformation, the Hotelling transformation, the method of empirical orthogonal functions, and singular value decomposition.

Conceptually, the goal of PCA is to reduce the number of variables of interest into a smaller set of components. In other words, the goal is to extract as much variance with the fewest components. PCA analyses all the variance in the variables and reorganises it into a new set of components equal to the number of original variables. Regarding the new components, they are independent and decrease in the amount of variance in the originals they account for. First component captures most of the variance, second most and so on until all the variance is accounted for. Only some will be retained for further study (dimension reduction) since the first few capture most of the variance they are typically off focus.

In essence, PCA seeks to reduce the dimension of the data by finding a few orthogonal linear combinations (the PCs) of the original variables with the largest variance. The first PC, s_1, is the linear combination with the largest variance. We have $s_1 = x^T w_1$, where the p-dimensional coefficient vector $w_1 = (\omega_{1,1},...,\omega_{1,p})^T$ solves

$$w_1 = \arg\max_{\|w=1\|} Var\left\{x^T w\right\} \tag{4.1}$$

The second PC is the linear combination with the second largest variance and orthogonal to the first PC, and so on. There are as many PCs as the number of the original variables. For many datasets, the first several PCs explain most of the variance, so that the rest can be disregarded with minimal loss of information.

Since the variance depends on the scale of the variables, it is customary to first standardize each variable to have a mean zero and standard deviation one. After the standardization, the original variables with possibly different units of measurement are all in comparable units. Assuming a standardized data with the empirical covariance matrix

$$\sum_{pxp} = \frac{1}{n} XX^T \tag{4.2}$$

we can use the spectral decomposition theorem to write $\sum$ as

$$\sum = U \Lambda U^T \tag{4.3}$$

where Λ=diag($\lambda_1,\ldots,\lambda_p$) is the diagonal matrix of the ordered eigenvalues $\lambda_1 \leq \ldots \leq \lambda_p$, and U is a $p \times p$ orthogonal matrix containing the eigenvectors. It can be shown (Mardia *et al.*, 1995) that the PCs are given by the p rows of the $p \times p$ matrix S, where

$$S = U^T X \tag{4.4}$$

The subspace spanned by the first k eigenvector has the smallest mean square deviation from X among all subspaces of dimension k.

Another property of the eigenvalue decomposition is that the total variation is equal to the sum of the eigenvalues of the covariance matrix,

$$\sum_{i=1}^{p} Var(PC_i) = \sum_{i=1}^{p} \lambda_i = \sum_{i=1}^{p} trace(\Sigma) \tag{4.5}$$

and that the fraction $\sum_{i=1}^{k} \lambda_i / trace(\Sigma)$ gives the cumulative proportion of the variance explained by the first k PCs. By plotting the cumulative proportions in (4.5) as a function of k, one can select the appropriate number of PCs to keep in order to explain a given percentage of the overall variation. The number of PCs to keep can also be determined by first fixing a threshold λ_0, then only keeping the eigenvectors such that their corresponding eigenvalues are greater than λ_0. In this case, the problem is to Figure out how many components need to be considered, standard advice to keep only the eigenvalues larger than 1 is used.

An alternative way to reduce the dimension of a dataset using PCA is suggested in (Mardia *et al.*, 1995). Instead of using the PCs as the new variables, this method uses the information in the PCs to find important variables in the original dataset. As before, one first calculates the PCs, then studies the scree plot to determine the number k of important variables to keep. Next, one considers the eigenvector corresponding to the smallest eigenvalue (the least important PC), and discards the variable that has the largest (absolute value) coefficient in that vector. Then, one considers the eigenvector corresponding to the second smallest eigenvalue, and discards the variable contributing the largest (absolute value) coefficient to that eigenvector, among the variables not discarded earlier. The process is repeated until only k variables remain.

4.2.2 Rotation

After the number of components has been determined, and in order to facilitate the interpretation, the analysis often involves a rotation of the components that were retained. When the data follow a model (such as the psychometric model) stipulating 1) that each variable load on only one factor and 2) that there is a clear difference in intensity between the relevant factors (whose eigenvalues are clearly larger than one) and the noise (represented by factors with eigenvalues clearly smaller than one), then the rotation is likely to provide a solution that is more reliable than the original solution.

Two main types of rotation are used: orthogonal when the new axes are also orthogonal to each other and oblique when the new axes are not required to be orthogonal. Because the rotations are always performed in a subspace, the new axes will always explain less inertia than the original components (which are computed to be optimal). When performing a rotation, the term loadings almost always refer to the elements of loading matrix. An orthogonal rotation is specified by a rotation matrix, where the rows stand for the original factors and the columns for the new (rotated) factors. With oblique rotations, the new axes are free to take any position in the component space, but the degree of correlation allowed among factors is small because two highly correlated components are better interpreted as only one factor.

Oblique rotations, therefore, relax the orthogonal constraint in order to gain simplicity in the interpretation.

The rotation used in this thesis is an orthogonal rotation, called varimax rotation. It was developed by Kaiser (1958) and is the most popular rotation method. For varimax a simple solution means that each component has a small number of large loadings and a large number of zero (or small) loadings. This simplifies interpretation because, after a varimax rotation, each original variable tends to be associated with one (or a small number) of components, and each component represents only a small number of variables. In addition, the components can often be interpreted from the opposition of few variables with positive loadings to few variables with negative loadings. Formally, varimax searches for a linear combination of the original factors such that the variance of the squared loadings is maximized.

4.2.3 Analysis of principle component

In general, components are linear combinations of variables. These combinations are based on weights (eigenvectors) developed by the analysis. With multivariate research we come to eigenvalues and eigenvectors. Eigenvalues can be considered to measure the strength (relative length) of an axis in *N*-dimensional space and derived via eigen-analysis of the square symmetric matrix (the covariance or correlation matrix). Each eigenvalue has an associated eigenvector. While an eigenvalue is the length of an axis, the eigenvector determines its orientation in space. The values in an eigenvector are not unique because any coordinates that described the same orientation would be acceptable. Any factor whose eigenvalue is less than 1.0 is in most cases not going to be retained for interpretation of components.

PCA will give the loadings of variables. The loading for each variable is the correlation between it and the component. Loadings that are more than 0.5 are typically considered strong, between 0.3 and 0.5 are acceptable, and less than 0.3 are typically considered weak.

The correlation matrix of each data set was obtained to assess the measure of pair wise association among the various variables. Stepwise multiple regressions were then carried out, and the results were also checked for multicollinearity by examining the variance inflation factors (VIF) of the predictor variables. The principal components of the predictor variables were obtained using the correlation coefficients of the variables. The results of the principal component analysis were employed for two purposes. First, they were used for principal component regression analysis (PCR), applying the stepwise regression option in the choice of the principal components to enter the regression equation. Next, a varimax rotation of the principal components was used as a variable selection technique to choose the appropriate variables for inclusion in the ultimate regression model. The objectives of this approach were to minimize the effect of multicollinearity on the estimation of the regression coefficients and achieve parsimony. The analysis of the data was carried out using the statistical software, SPSS (Statistical Package for Social Science, version 10.0).

4.3 TWO-STAGE ANN CLASSIFIER USING PCA FOR HYPERGLYCAEMIA DETECTION

A two-stage ECG classifier based neural network is developed for hyperglycaemia episodes detection. This classifier uses inputs of ECG parameters to detect output of either hyperglycaemic, or normoglycaemic states. In this study, PCA performs the features dimensionality reduction to find the optimal group of ECG variables as input. An overview of the proposed hyperglycaemia detection system is described in Figure. 4.1.

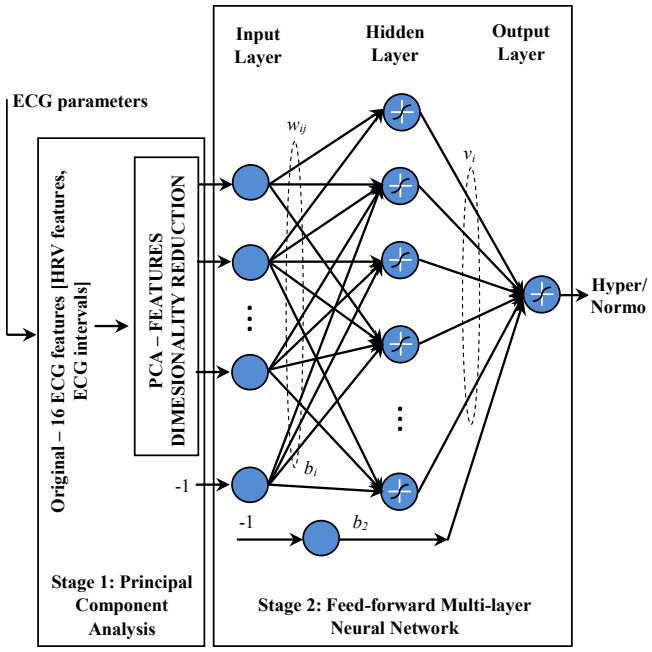


Figure 4.1: Hyperglycaemia detection using two-stage PCA-Multilayer neural network classifier and input of ECG parameters

4.3.1 First stage – PCA for feature reduction

PCA is an appropriate multivariate technique to reduce the dimension of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set (Sharma *et al.*, 1996). It is a special case of factor analysis that transforms the original set of inter-correlated variables into a new set of an equal number of independent uncorrelated variables or principal components that are linear combinations of the original variables. Therefore, high correlations (multicollinearity) between predictor variables can be avoided for correct identification of the most important variables. The principal components are ordered in a way that the first few principal components explain most of the variation in all of the original variables, and each subsequent one accounts for the largest proportion of variability that has not been retained by its predecessors (Jolliffe, 2002). The components are equally useful in regression analysis for mitigating the problem of multicollinearity and in exploring the relations among the independent variables, particularly if it is not obvious which of the variables should be the predictors. The new variables from the PCA become ideal to use as predictors in a regression equation since they optimise spatial patterns and remove possible complications

caused by multicollinearity. A multiple regression based on PCA is going to be utilised to components analysis.

The magnitudes of the coefficients give the contributions of each variable to that component. PCA can be formally stated in the following equations, assuming that there are p original variables x and producing p linear combinations:

$$\begin{aligned}
\xi_1 &= w_{11}x_1 + w_{12}x_2 + \ldots + w_{1p}x_p \\
\xi_2 &= w_{21}x_1 + w_{22}x_2 + \ldots + w_{2p}x_p \\
&\vdots \\
\xi_p &= w_{p1}x_1 + w_{p2}x_2 + \ldots + w_{pp}x_p
\end{aligned} \tag{4.6}$$

where ξ_p are the p principal components and w_{ij} is the weight of the j^{th} variable for the i^{th} principal component. Due to the differences in the units of variables, correlation matrix of variables was used to get eigenvalues and weight of variables.

After generalizing the principal component, varimax rotation was used to maximize the loading of a predictor variable on one component. Generally, implementation of the PCA procedures followed by an orthogonal rotation method (varimax rotation) produces a ranked series of factors. The higher the loading a variable, the more that variable contributes to the variation accounted for by the particular principal component. In practice, only loadings with absolute values greater than 50% are selected for the principal component interpretation (Jolliffe, 2002). A principal component with an eigenvalue greater than or equal to one is commonly considered for statistical significance (the Kaiser criterion). The statistical analyses of the data were carried out using the SPSS software (Statistical Package for Social Science, IBM, version 19.0).

In this study, firstly, Pearson correlation matrices of the ECG variables were obtained to access the measure of pair-wise association among the various variables. If multicollinearity exists with high correlations among variables, then PCA should be used to avoid this problem. Kaiser-Meyer-Olkin (KMO) and Bartlett's test was also mandatory for the verification of the applicability of PCA to our data set. We then employed PCA as the first stage to transform 16 original ECG parameters into a new appropriate lower dimensional size of variables, which will be used as input in

the second stage of ECG classifier, multilayer neural network. The objective of PCA focused especially on identification of the most important ECG factors among cardiac intervals and HRV measures that are responsible for the variation of high blood glucose levels in diabetic patients. The selection of the subset of ECG predictor variables followed two steps:

(1) Choose principal components that have eigenvalues equal or greater than one;

(2) Select variables with high loading associated with each principal component.

4.3.2 Second stage – ANN classifier

In the second stage of the ECG classifier, these features are employed as inputs in a multilayer neural network. The structure of neural network has nine input nodes which are the nine mentioned ECG parameters above. The hidden nodes are varied from 3 to 20 in order to select the one that gives the best performance. The overall data set consists of 107 episodes of normoglycaemia and 80 episodes of hyperglycaemia. The training set, validation set and testing set are randomly selected with proportions of 40%, 30%, and 30% out of overall data, respectively. To find the contribution of cardiac intervals and HRV measures to the classification of hyperglycaemia, different neural networks are trained with inputs from cardiac intervals. The training performance of proposed detection has been analysed by means of ROC curves for five groups of inputs.

The ANN is trained by the Levenberg-Marquardt (LM) algorithm, which appears to be a fast and effective training algorithm. For comparison of performances, five approaches using different training algorithms with the same structure and same input of all nine ECG features are presented as well. The algorithms are Gradient descent with momentum (PCA_GDX), Scaled conjugate gradient (PCA_SCG), Resilient back propagation (PCA_RP), Conjugate gradient back-propagation with Fletcher-Reeves updates (PCA_CGF).

4.4 EXPERIMENTAL RESULTS

4.4.1 PCA results

Multiple regression analysis based on 16 original ECG features has given the results of collinearity statistics which is been shown in Table 4.1. The result shows that there was a multicollinearity problem with the variance inflation factor (VIF) where some of VIFs are more than 5.0.

ECG parameters	Collinearity Statistics	
	Tolerance	VIF
HR	.843	1.186
QT_C	.175	5.714
PR	.965	1.036
RT_C	.217	4.601
$TpTe_C$	.510	1.959
VLF	.135	7.408
LF	.132	7.571
HF	.269	3.717
TotalPw	.548	1.824
LF/HF	.800	1.250
MeanRR	.079	12.709
SDNN	.091	11.037
RMSSD	.449	2.228
pNN50	.608	1.646
HRVi	.272	3.678
TINN	.096	10.372

Table 4.1: Collinearity Statistics of original ECG data

	HR	QT_C	*PR*	RT_C	$TpTe_C$	*VLF*	*LF*	*HF*	*LF/HF*	*Mean RR*	*SDNN*	*RMSSD*	*pNN50*	*HRVi*	*TINN*	*Total Pw*
HR	1	.356**	-.132	.210*	.129	-.046	-.053	-.049	-.074	-.537**	-.064	-.063	-.089	.187*	.136	-.052
QT_C	-	1	-.160*	.792**	.374**	.066	.091	.079	.045	.116	.109	.085	-.060	-.009	.038	.083
PR	-	-	1	-.149*	-.021	-.154*	-.166*	-.086	-.149*	-.022	-.198*	-.160*	.032	-.120	-.188*	-.160*
RT_C	-	-	-	1	-.061	.074	.083	.080	.051	.141*	.121	.112	-.031	.008	.107	.083
$TpTe_C$	-	-	-	-	1	.052	.050	-.006	-.048	.045	-.003	-.022	-.001	-.020	-.132	.047
VLF	-	-	-	-	-	1	.972**	.504**	.384**	.120	.774**	.432**	.008	-.018	.255*	.982**
LF	-	-	-	-	-	-	1	.589**	.420**	.156*	.859**	.539**	.049	.034	.283**	.992**
HF	-	-	-	-	-	-	-	1	.035	.233*	.792**	.802**	.490**	.292**	.378**	.637**
LF/HF	-	-	-	-	-	-	-	-	1	.021	.319**	.041	-.206*	.141*	.104	.378**
MeanRR	-	-	-	-	-	-	-	-	-	1	.234*	.241*	.169*	-.168*	-.050	.158*
SDNN	-	-	-	-	-	-	-	-	-	-	1	.852**	.360**	.256*	.475**	.861**
RMSSD	-	-	-	-	-	-	-	-	-	-	-	1	.597**	.318**	.546**	.554**
pNN50	-	-	-	-	-	-	-	-	-	-	-	-	1	.396**	.349**	.090
HRVi	-	-	-	-	-	-	-	-	-	-	-	-	-	1	.288**	.045
TINN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	.298**
TotalPw	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Statistical significant coefficients were denoted with * for $p< 0.05$ and with ** for $p< 0.0001$

Table 4.2: Pearson correlations matrix of original ECG parameters

Cross correlation (Pearson correlation) between 16 original ECG variables was also examined in this study. As shown in Table 4.2, statistically significant correlation coefficients *r* were denoted to confirm strong associations between ECG variables. There have existed high correlations between each component among cardiac intervals and HRV measures, i.e., QT_C with RT_C, *VLF* with *LF* and *TotalPw*, *LF* with *SDNN* and *TotalPw*, *HF* with *RMSSD* ($r > 0.7$ and $p < 0.0001$). It again proved that there was a significant amount of multicollinearity among ECG parameters, which may lead to high standard errors of the parameters' estimation in neural networks later. PCA is one of the approaches to avoid this problem. The main objective of PCA was to obtain a small number of components that would explain most (typically 60-90%) of the total variation in the original ECG parameters.

Kaiser-Meyer-Olkin (KMO) Measure of sampling adequacy and Bartlett's test for Sphericity were then applied to verify the applicability of PCA on an ECG data set, which has been shown in Table 4.3. The Bartlett's sphericity test is for all correlations, zero or for testing the null hypothesis where the correlation matrix is an identity matrix which was used to verifying the applicability of PCA. The value of Bartlett's sphericity test is 1536.187 which suggest that the PCA is applicable to our data sets ($p < 0.0001$). The value of KMO's Measure of Sampling was adequately found to be 0.783 which is higher than 0.5 implying that the sample size of ECG data set was enough for applying PCA.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		**0.783**
Bartlett's Test of Spericity	Approx. Chi-Square	1536.187
	df	55
	Sig.	.000

Table 4.3: KMO and Bartlett's test for ECG data

4.4.2 Subgroup data set

In PCA, original ECG parameters were transformed into an equal number of principal components. Sixteen original ECG parameters will give sixteen principal components (PC). After the transformation, varimax rotation was used to maximize the loading of a variable on one component. According to the results of PCA in Table 4.4, the first five principal components (PC1, PC2, PC3, PC4, PC5) which have eigenvalues greater than 1 are selected as reference for feature reduction. Because eigenvalues represent variances, a component with an eigenvalue of less than 1 is not significant. Thus, the first of five principal components provides an adequate summary of the data for most purposes. Only the first five principal components, explaining 76.635% of the total variation, should be sufficient for almost any applications. The first PC accounted for 33% of the total variation in the data. It is loaded heavily on *VLF, TotalPw, LF* and *SDNN*. The second PC, which accounted for 13% of the total variation, exhibited strong loadings on RT_C and QT_C. Principal components three, four and five loaded heavily on *pNN50,* $TpTe_C$ and *TINN*. Based on these PCA results, the final set with nine features are selected as inputs of classification, namely:

- *VLF*
- *TotalPw*
- *LF*
- *SDNN*
- RT_C
- QT_C
- *pNN50*
- $TpTe_C$
- *TINN*

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15	PC16
VLF	**.985**	.020	-.022	.022	.056	-.040	.011	-.043	-.013	.096	-.013	-.052	-.014	-.057	-.081	.000
TotalPw	**.982**	.029	.027	.018	.073	.001	.039	-.043	-.013	.098	.117	.026	-.002	-.024	-.001	.000
LF	**.978**	.031	-.005	.022	.063	-.004	.041	-.047	-.013	.140	.065	.063	.010	.006	.083	.000
SDNN	**.802**	.062	.218	-.016	.207	.145	.110	-.076	-.025	.104	.238	.346	.015	.188	.004	.000
RTc	.051	**.968**	-.006	-.129	.043	.007	.053	-.054	.041	.006	-.002	.009	-.185	-.003	-.001	.000
QTc	.044	**.891**	-.028	.303	.003	-.017	.061	-.063	.187	.026	.037	.022	.259	.006	.001	.000
pNN50	.047	-.038	**.937**	.007	.162	.206	.086	.032	-.046	-.125	.128	.088	-.001	.005	.000	.000
TpTec	.035	.067	.005	**.992**	-.068	-.006	.020	-.007	.057	-.027	-.008	-.006	.005	.000	.000	.000
TINN	.215	.041	.172	-.081	**.936**	.120	-.041	-.085	.067	.028	.072	.083	-.001	.003	.000	.000
HRVi	.014	-.007	.201	-.007	.116	.955	-.088	-.056	.085	.085	.081	.052	-.001	.003	.000	.000
MeanRR	.108	.106	.093	.029	-.035	-.089	.934	-.014	-.282	-.002	.062	.043	.002	.002	.000	.000
PR	-.106	-.088	.021	-.009	-.074	-.051	-.010	.983	-.053	-.054	-.008	-.023	-.001	-.001	.000	.000
HR	-.037	.201	-.047	.076	.067	.092	-.304	-.065	.916	-.048	-.006	-.012	.005	.000	.000	.000
LF_HF	.294	.024	-.125	-.031	.026	.087	-.001	-.061	-.044	.939	-.039	-.012	.001	.002	.000	.000
HF	.542	.041	.302	-.016	.136	.166	.115	-.008	-.010	-.081	.719	.171	.003	.004	.001	.000
RMSSD	.473	.062	.411	-.019	.300	.160	.122	-.072	-.036	-.061	.319	.601	.000	-.017	.001	.000
Eigenvalue	**5.349**	**2.14**	**2.05**	**1.6**	**1.122**	0.949	0.842	0.598	0.361	0.354	0.314	0.181	.095	0.032	0.012	0
% of Variance	33.43	13.376	12.811	10.003	7.015	5.931	5.261	3.738	2.259	2.212	1.962	1.129	0.594	0.203	0.076	0
Cumulatives %	33.43	46.806	59.617	69.62	76.635	82.566	87.827	91.565	93.825	96.037	97.998	99.127	99.721	99.924	100	100

Table 4.4: Total variance and rotated principal components loadings

4.4.3 Performance of two-stage ANN classifier using PCA for hyperglycaemia detection

In the second stage of the ECG classifier, these features are employed as inputs in a multilayer neural network. The structure of the neural network has nine input nodes which are the nine mentioned ECG parameters above. The hidden nodes are varied from 3 to 20 in order to select the one that gives the best performance. The overall data set consists of 107 episodes of normoglycaemia and 80 episodes of hyperglycaemia. The training set, validation set and testing set are randomly selected with proportions of 40%, 30%, and 30% out of overall data, respectively. To find the contribution of cardiac intervals and HRV measures to the classification of hyperglycaemia, different neural networks are trained with inputs from cardiac intervals (RT_C, QT_C, $TpTe_C$), 4 features selected from PC1 (*VLF, TotalPw, LF, SDNN*), HRV measures including time-domain (*SDNN, pNN50, TINN*) and frequency-domain (*VLF, TotalPw, LF*), separately.

The training performance of proposed detection has been analysed by means of ROC curves for five groups of inputs in which the sensitivity (true positive rate) and the 1- specificity (false positive rate) are plotted in Figure 4.2. The results of hyperglycaemia detection using LM algorithm are presented in Table 4.5. The corresponding *AuR* (area under ROC curve) for the training set with input of all nine parameters is 0.8455. This ROC curve is used to find the optimal cut-off point to classify hyperglycaemic and normoglycaemic states. In order to analyse the optimum sensitivity and specificity, the cut-off point is selected at the point producing the sensitivity higher than 80%. At that point, 6 hidden nodes produce the best classification in which the training results of 88.89% sensitivity and 71.88% specificity are gained and the testing set leads to 85.71% sensitivity and 58.62% specificity. These testing sensitivity and specificity are considered to be a satisfactory result which meets the requirements of disease diagnosis (sensitivity $\geq$ 70% and specificity $\geq$ 50%). The best testing geometric mean of 70.88% is also achieved with these sensitivity and specificity.

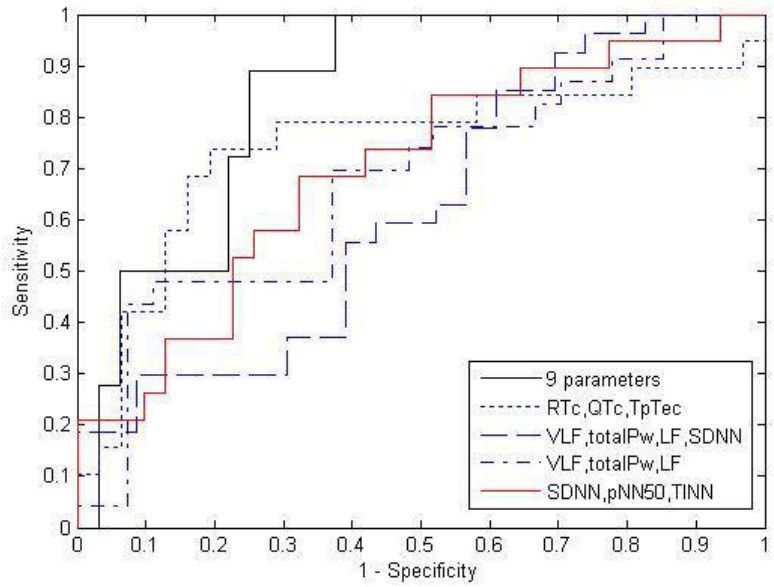


Figure 4.2: ROC plot

Inputs	Hid. node	AuR	Sen (%)	Spec (%)	*gm* (%)	Sen (%)	Spec (%)	*gm* (%)	Sen (%)	Spec (%)	*gm* (%)
9 parameters	6	0.8455	88.89	71.88	79.93	66.67	57.69	62.02	85.71	58.62	70.88
RT_C, QT_C, $TpTe_C$	4	0.7419	78.95	64.52	71.37	66.67	58.62	62.52	68.75	59.26	63.83
VLF, TotalPw, LF, SDNN	14	0.6216	66.67	43.48	53.84	68.75	52.94	60.33	69.23	60	64.45
VLF, TotalPw, LF	10	0.6747	73.19	51.85	61.91	50	64.29	56.69	72.73	62.5	67.42
SDNN, pNN50, TINN	18	0.6978	84.21	45.16	61.67	52.63	54.84	53.72	83.33	48	63.24

Table 4.5: Performances of ECG classifier using PCA and LM algorithm

The classification results from smaller groups of ECG parameters show comparable testing results such as input of (*VLF, TotalPw, LF*) and (RT_C, QT_C, $TpTe_C$) whose sensitivity and specificity are (72.73% and 62.5%) and (68.75% and 59.26%), respectively. These results indicate a potential non-invasive detection of hyperglycaemia using neural network combined PCA and ECG parameters.

For comparison of performances, five approaches using different training algorithms with the same structure and same input of all nine ECG features are presented as well. The algorithms are Gradient descent with momentum (PCA-GDX), Scaled conjugate gradient (PCA-SCG), Resilient back propagation (PCA-RP), Conjugate gradient back-propagation with Fletcher-Reeves updates (PCA-CGF). In the terms of sensitivity, specificity, and geometric mean, the performances of the detections using SCG, GDX, CGF, and RP algorithms, respectively, are described in Table 4.6. It can be clearly seen that the obtained testing LM result outperforms other algorithms such as SCG, GDX, RP and CGF in terms of sensitivity (85.71% against %. 69.23%, 70%, 71.43%, and 52.94%, respectively).

Algorithms	Hidden node	Sen (%)	Spec (%)	*gm* (%)
PCA-LM	6	85.71	58.62	70.88
PCA-GDX	13	70	63.64	66.74
PCA-SCG	9	69.23	66.67	67.94
PCA-RP	6	71.43	62.07	66.59
PCA-CGF	8	52.94	88.46	68.43

Table 4.6: Testing performances of neural network using PCA and five different algorithms

To evaluate the results of using PCA to reduce features dimensionality, hyperglycaemia detections utilising neural network which did not use PCA in the first stage, are also performed. The neural networks applying five algorithms, namely LM-NN, SCG-NN, GDX-NN, CGF-NN, and RP-NN used all 16 original ECG

parameters as inputs. Table 4.7 shows the best testing performances of five algorithms. As expected, multicollinearity problem in the data set led to a worsening performance, where the highest classification by using LM-NN was only 70.59%, 65.38%, and 67.94%, in terms of sensitivity, specificity and geometric mean, respectively. In terms of geometric mean, their performances are comparable to neural networks using PCA, however, apart from LM-NN, the sensitivities of the other four algorithms are less than 70%, which are not directed in biomedical application.

Algorithms	Hidden node	Sen (%)	Spec (%)	gm (%)
LM-NN	9	70.59	65.38	67.94
GDX-NN	12	66.67	54.84	60.47
SCG-NN	10	68.75	59.26	63.83
RP-NN	11	62.5	62.96	62.73
CGF-NN	7	55.56	68	61.47

Table 4.7: Testing performances of neural network without PCA using all 16 original parameters and five different algorithms

4.5 DISCUSSION

In this study, according to PCA results, HRV measures such as *VLF, TotalPw, SDNN, TINN,* which have the strongest significant decreases under hyperglycaemia conditions in statistical *t*-test, are selected to examine the presence of hyperglycaemia episodes. Although *HF* and *RMSSD* components are more statistically significant than *LF* and *TINN, LF* and *TINN* are still chosen as PCA considering the total variation of all the data associated with the contributions of each component. As a result, *VLF, TotalPw* and *LF* variables, which are ranked on the top of PC1 components, yielded better classification results compared to the group of time-domain measures including *SDNN, pNN50* and *TINN*. This outcome implies that there might be a continuous relation between glycaemia and all measures of

absolute spectral power of HRV (*VLF, LF, HF*). In our study, based on the PCA result, we found the most suitable group of parameters for hyperglycaemia detection. However, for clinical use, there is still no consensus about the most accurate HRV parameter.

The relation of hyperglycaemia and ventricular repolarisation intervals (RT_C, QT_C, $TpTe_C$) has been reported in previous studies (Suys *et al.*, 2006; Nguyen *et al.*, 2012). The geometric mean of 63.83% in the classification using these parameters indicates that they have an important contribution for hyperglycaemia detection. This important role also matches with results of statistical *t*-test showing that RT_C, QT_C, $TpTe_C$ are smaller in hyperglycaemia than in normoglycaemia with $p < 0.05$. Besides, based on Pearson correlation matrix (Table 4.2), RT_C, QT_C, $TpTe_C$ have strong correlations with each other which led to their high contributions to the total variance of the all data set. *PR* interval is significantly higher in hyperglycaemia than in normoglycaemia ($p < 0.0001$), yet its variance was not effectively reflected in the total variance of all variables (under 4% out of 100%) explaining why *PR* was rejected in PCA.

Levenberg-Marquardt (LM) algorithm is often the fastest back-propagation algorithm in training a moderate-sized feed forward neural network (up to several hundred weights). It is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms. A three-layer feed forward neural network training with LM using the whole data set for hyperglycaemia detection received reasonable classification results (as tabulated in Table 4.7) though it might take more time and higher hidden nodes. After being accompanied with principal component analysis to reduce the dimension of data set, the proposed ECG classifier achieved even better results (85.71% sensitivity, 58.62% specificity, 70.88 % geometric mean vs. 70.59% sensitivity, 65.38% specificity, 67.94% geometric mean) due to the multicollinearity between predictor variables which has been prevented to avoid incorrect identification of the most important predictors. In cases where other algorithms are applied, the proposed ECG classifier still results with acceptable performances, such as GDX (70% sensitivity, 63.64%

specificity, 66.74% geometric mean) and RP (71.43% sensitivity, 62.07% specificity, and 66.79% geometric mean).

As far as we are concerned, this study is the first to report a non-invasive hyperglycaemia detection system using ECG signals. To predict hyperglycaemia, many other studies have been worked on in different ways using various input parameters, but none of them examined the abnormal ECG parameters under hyperglycaemia conditions. Most of the alternative studies focus on predicting blood glucose levels to forecast glycaemic states. For example, Le Compte (2010) built a dynamic model for blood glucose prediction to provide the most probable future insulin sensitivity. Other approaches developed neural networks for early detection of hypoglycaemia/hyperglycaemia events based on glucose data from continuous glucose monitoring (CGM) (Pappada *et al.*, 2010; Daskalaki *et al.*, 2012). Nevertheless, the purpose of these studies was on real-time glucose control rather than disease diagnosis. By the way of hyperglycaemia diagnosis, classification models concentrate on keeping sensitivity and specificity high enough in balance, although specificity should be lower than sensitivity as the true positive rate (proportion of hyperglycaemic events are correctly detected) is more important than false positive rate (proportion of normoglycaemic events are correctly detected). Our findings contribute to the limited literature of problems related to hyperglycaemia diagnosis based on physiological signals among T1DM. Moreover, this study draws attention to the need for further investigations to understand the mechanisms responsible for abnormal changes in ECG signals during a state of hyperglycaemia. It should also be assessed whether subclinical changes in ECG are predictive of the subsequent development of clinical hyperglycaemia conditions in these patients.

In the next stage, we are pursuing the enhancement of the proposed strategy to improve overall accuracy of hyperglycaemia detection. In addition, other optimisation techniques, such as particle swarm optimisation (PSO), might need to be explored to optimise the neural network parameters in order to obtain better results. This hyperglycaemia detection strategy, after further validation, might have a prospect to be a non-invasive and painless method for diabetic patients who have a risk of high blood glucose.

In this chapter, a two-stage classification unit using artificial neural network combined PCA is developed to determine the presence of hyperglycaemia episodes based on T1DM patients' ECG parameters. The group of nine ECG parameters after being reduced by PCA shows significant contributions to the performance of hyperglycaemia detection. The results of 85.71% sensitivity and 58.62% specificity by using LM algorithm are considered reasonable and meet the requirement for biomedical application. Various training algorithms, which have been employed and not employed in the proposed classifier, also have been developed and compared. In general, the proposed ECG classifier using PCA and LM algorithms is a robust method and outperforms other ones. This result indicates that hyperglycaemic events in T1DM can be detected non-invasively and effectively by using ECG signals.

CHAPTER 5 . PARTICLE SWARM OPTIMISATION BASED NEURAL NETWORK FOR HYPERGLYCAEMIA DETECTION

5.1 INTRODUCTION

Artificial Neural Networks (ANN) have shown potential results in the classification of hyperglycaemic episodes in T1DM patients, which have been described in chapter 3 and 4. On the way to improve the performance of hyperglycaemia detection using ANN, particle swarm optimisation (PSO) is considered to associate with ANN to optimise the neural network's parameters.

Usually some gradient algorithms including the Back Propagation (BP) technique are used for training a network, because they can achieve faster convergent speedaround the global optimum. However they are prone to being trapped in the local minimum. On the contrary, the PSO algorithm converges rapidly during the initial stages of a global search, while it becomes extremely slow around the global optimum. Therefore, an adaptation with a transition from PSO search to train the weights and thresholds of neural network, which aims to exploit the advantage of

both algorithms, is proposed in this chapter. This PSO based neural network (PSO-NN) can achieve the strong global search ability of PSO, which could be an effective alternate training algorithm for ANN's since PSO is a heuristic optimisation technique and is found to be more accurate when compared to the existing conventional algorithms.

Regarding to optimisation techniques, the PSO algorithm and genetic algorithm (GA) are both popular methods. Similar to GA, PSO is an optimisation tool based on population, and the system is initialised with a population of random solutions and can search for optima by updating of generations. Conversely, unlike the GA, PSO has no complicated evolutionary operators such as crossover and mutation (Boeringer *et al.*, 2004). Particles update themselves with the internal velocity; they also have a memory important to the algorithm. Also, in PSO only the 'best' particle gives out the information to others. It is a one-way information sharing mechanism, the evolution only looks for the best solution. Compared to GA, the advantages of PSO are that PSO is easy to implement and there are few parameters to adjust. In general, the PSO algorithm has a strong ability to find the most optimistic result, although easily getting into a local minimum. After suitably modulating the parameters for the PSO, the rate of convergence can be speeded up and the ability to find the global optimistic result can be enhanced. That's why PSO has been chosen over GA as the optimization method to improve the hyperglycaemia performance.

Chapter 5 proposes a new hyperglycaemia detection model, which consists of two stages: feature dimensionality reduction and PSO based neural network classification unit. In the first stage, the dimension of features is reduced using Principal Component Analysis (PCA) to form more essential features. In the classification stage, a hybrid classification algorithm combines PSO and NN which is used to train the perceptrons of a three-layer neural network to optimise the weights of ANN. Further, a sensitivity analysis is later carried out in the study to evaluate the most suitable PSO-NN characteristics which include maximum velocity; acceleration constants to obtain minimize error. The proposed particle swarm optimisation trained neural network is employed in hyperglycaemia detection and the results are compared with the other existing algorithms.

5.2 PARTICLE SWARM OPTIMISATIONBASED NEURAL NETWORK (PSO-NN)

5.2.1 PSO algorithm

Particle Swarm Optimisation (PSO) algorithm is based on the theory of swarm intelligence. This algorithm can provide efficient solutions for optimisation problems through intelligence generated from complex activities such as cooperation and competition among individuals in the biologic colony. Compared with evolutionary computation, PSO still maintains the population based global search strategy, but its velocity displacement search model is simple and easy to implement.

Inspired by the social behaviour of animals such as fish schooling and birds flocking, Eberhart and Kennedy designed the Particle Swarm Optimisation (PSO) in 1995. This method is a kind of evolutionary computing technology based on swarm intelligence (Gao *et al.*, 2006).The basic idea of birds flocking can be depicted as follows: In a bird colony, each bird looks for its own food and in the meantime they cooperate with each other by sharing information between them. Therefore, each bird will explore a new area of potential by utilise its own experience and that of the others. Due to these attractive characteristics, i.e. memory and cooperation, PSO is widely applied in many research areas and real-world engineering fields as a powerful optimisation tool.

The basic PSO model consists of a swarm of particles moving in a d-dimensional search space. The direction and distance of each particle in the hyper-dimensional space is determined by its fitness and velocity. In general, the fitness is primarily related with the optimisation objective and the velocity is updated according to a sophisticated rule.

5.2.2 Adaptation of PSO to train neural networks

A population of particles is initialised for a random position w_i and velocity v_i in a sample problem space of dimension d, which can be expressed as:

$$d = (I+1)\times H + (H+1)\times O \tag{5.1}$$

where *I, H, O* are the numbers of input, hidden and output nodes, respectively. It then simulates the social behaviour of the particles in the swarm until it achieves the optimum solution by updating generations. The position and velocity of the i^{th} particle is denoted as $w_i = (w_{i1}, w_{i2}, ..., w_{id})$ and $v_i = (v_{i1}, v_{i2}, ..., v_{id})$ respectively. At each iteration step, the velocity is updated and the particle is moved to a new position by following two "best" values. The first one is the best solution (fitness) of the i^{th} particle it has achieved so far. This value is called *pbest*, which denotes as $p_{besti} = (p_{best1}, p_{best2}, ..., p_{bestd})$. Another "best" value that is tracked by the particle swarm optimiser is the best value, obtained so far by any particle in the whole swarm, also called the global best position *gbest*. It is denoted as $g_{besti} = (g_{best1}, g_{best2}, ..., g_{bestd})$. When a particle takes part of the population as its topological neighbor, the best value is a local best and is called *lbest* (Birge, 2003).

After finding the two best values, the new velocity and position of each particle are calculated as follows:

$$v_i = \omega \times v_i + c_1 \times rand_1 \times (p_{besti} - w_i) + c_2 \times rand_2 \times (g_{besti} - w_i) \quad (5.2)$$

$$w_i = w_i + v_i \quad (5.3)$$

where $rand_1$ and $rand_2$ are two independent random numbers uniformly distributed in the range of [0, 1]. c_1, c_2 are two constants called learning factors. ω is called the inertia factor to deliver a balance between global exploration and local exploitation.

The inertia weight ω keeps the movement inertial for the particle. It describes the influence of the previous velocity to the current velocity, makes it have the trend to extend the search space and have the ability to explore the new district. There is a function to adjust the rate of velocity of particle, as shown in equation 5.4. Usually the inertia weight ω is decreased linearly from 0.9 to 0.4.

$$\omega_i = \omega_{\max} - k \times (\omega_{\max} - \omega_{\min} / iter_{\max}) \quad (5.4)$$

where ω_{max} is the initial weight, usually ω_{max}= 0.9; ω_{min} is the final weight, usually ω_{min} = 0.4; $iter_{max}$ is the maximum number of iteration. k is the current number of iterations.

The available search range for the particles plays an important role in resolving the solution convergence. Each component of velocity is within the range

[$-v_{max}$, $+v_{max}$] so that the maximum allowable velocity v_{max} can control the exploration and exploitation of the search space. v_{max} determines the maximum change one particle can take during iteration and determines the precision between current position and the global best position. If v_{max} is too high, the particle may fly beyond the best solution; If v_{max} is a small value, the particle cannot perform enough searches and get trapped in some local minima (Lee *et al.*, 2008).

It is noticeable that three components typically contribute to the new velocity. The first part is proportional to the old velocity and is the tendency of the particle to continue in the same direction it has been travelling (Yuhui *et al.*, 1998). It can be thought of as a momentum term. The second component, which is associated with a local search, is considered as the cognitive part representing the private thinking. The third term, which is associated with a global search, is considered as the social part representing the social-psychological adaptation of knowledge. In addition, studies of inertia weight ω yield the conclusion that starting with a high value and lowering it throughout the iterations will help enhance the possibility of converging globally and save the computational expense in the local search procedure.

As mentioned earlier in chapter 3, a three-layer feed-forward neural network, which basically consists of input layer, hidden layer and output layer, is composed of a series of interconnected nodes and the corresponding weights between them. ANN is characterised by the ability of self-learning and error toleration. With the appropriate activation functions and trained weights, ANN can approximate any smooth, nonlinear function or relationship between the input and output. The training process is carried out on a set of data including input and output parameters. Usually, the data are split into two parts namely training samples and validation samples. The learning procedure is based on the training samples and the testing samples are used to verify the performance of the trained network. During the training, the weights in the network are adjusted iteratively until a desired error depicted as equation (5.5) is obtained.

$$E = \sum_{i=1}^{m} \sum_{k=1}^{n_o} (t_i^k - y_i^k)^2 / 2 \tag{5.5}$$

where t_i^k and y_i^k represent the actual and the predicted values respectively, m is the

number of training samples. The neural network is trained by minimising the above error function in a search space of weights.

The performance of each particle is evaluated according to a predefined fitness function, which is related to the problems to be solved. In this work, the fitness is determined by (5.5). The PSO algorithm has strong robustness for the initial weight of neural network. By the combination of the PSO into training neural network, it could improve the precision, speed and convergence rate of the network.

5.2.3 Hyperglycaemia detection using PSO-NN

A two-stage PSO-NN classifier has been developed for the classification of hyperglycaemia. It is an upgrade from the ANN model which has been explored in chapter 4. The PSO-NN still has stage 1 using PCA to reduce the feature dimensionality. After feature reduction, there are nine essential ECG variables needed for a hyperglycaemia classifier. A PSO based multi-layer ANN is implemented based on network in the second stage. PSO Research Toolbox (Evers *et al.*, 2009), Matlab Neural Network Toolbox (version R2012a) and PSO Research Toolbox – NN Matlab add-on (Rambharose, 2011) are used to train the neural network.

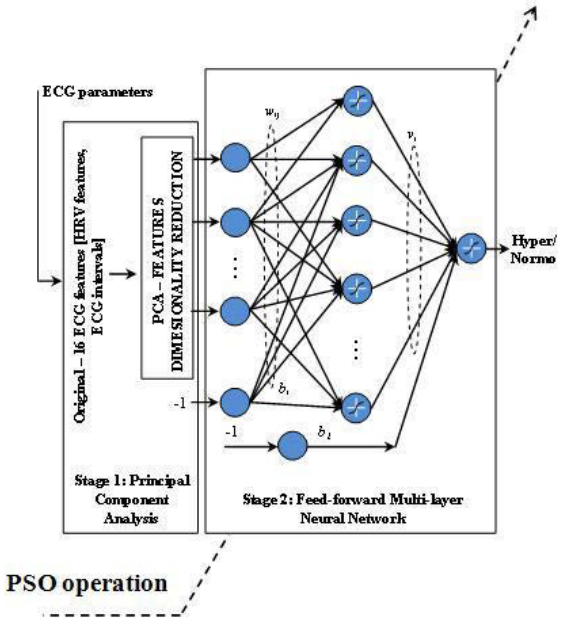


Figure 5.1: PSO based feed-forward Neural network for hyperglycaemia detection

The specific hyperglycaemia classifier based on the PSO-NN is illustrated in Figure 5.1. Stage 1 is for feature reduction using PCA, which was described clearly in chapter 4. In stage 2, the input after reduction is sent to the neural network model and then the PSO algorithm is applied to train the neural network until the goal fitness or MSE arrives (as shown in equation 5.5). The network was trained using PSO to arrive at the initial network weights.

Since a three-layer feed forward neural network is considered in this study, the weight matrix is presented as $w = \{w_{ih}, w_{ho}\}$ where w_{ih} and w_{io} are the set of weights connecting the input to the hidden layer and the hidden layer to the output layer respectively. The initial ranges of ANN parameters used in the PSO model are fixed according to Table 5.1. The values of MSE are evaluated by using the current position of each particle as an input and its value subsequently compared for each particle as its current position and the position at which a best fit is achieved. If the current fitness happens to be the best one so far, it is stored as p_{besti}. A random weighted difference between p_{besti} and the weight matrix w is then added to the particle velocity and the current best fitness value is stored as g_{besti}. Supposing that every weight and threshold in the network is initially set in the range of [-1,1], then every initial velocity and position of the particles is a set of weights and thresholds generated randomly in the range of [-1,1]. In addition, the maximum velocity v_{max} was regulated by velocity clamping percentage $v_{max\text{-}perc}$. It is the percentage of the range to which each dimension should be clamped, according to the following equation:

$$v_{\max} = v_{\max-perc}.range_i(\Omega); \tag{5.6}$$

where $v_{\max-perc} \in (0,1]$ and $range_i(\Omega) = w_i^U - w_i^L$; and Ω is the search space. Other primary PSO parameters are presented in Table 5.2.

The PSO algorithm uses local neighborhood best method, where the neighborhood was set to five particles. With this configuration, each neighborhood's best performing particle affects only velocities and directions of its five nearest neighbors. This causes the particles to converge on a solution more slowly, but

allows the algorithm to search more of the solution space. The flowchart 5.1 summarizes the local best PSO algorithm.

Parameters	Value
Maximum number of epochs	1000
Performance goal	0.001
Maximum validation failures	6
Epochs between display	10
Learning rate	0.01

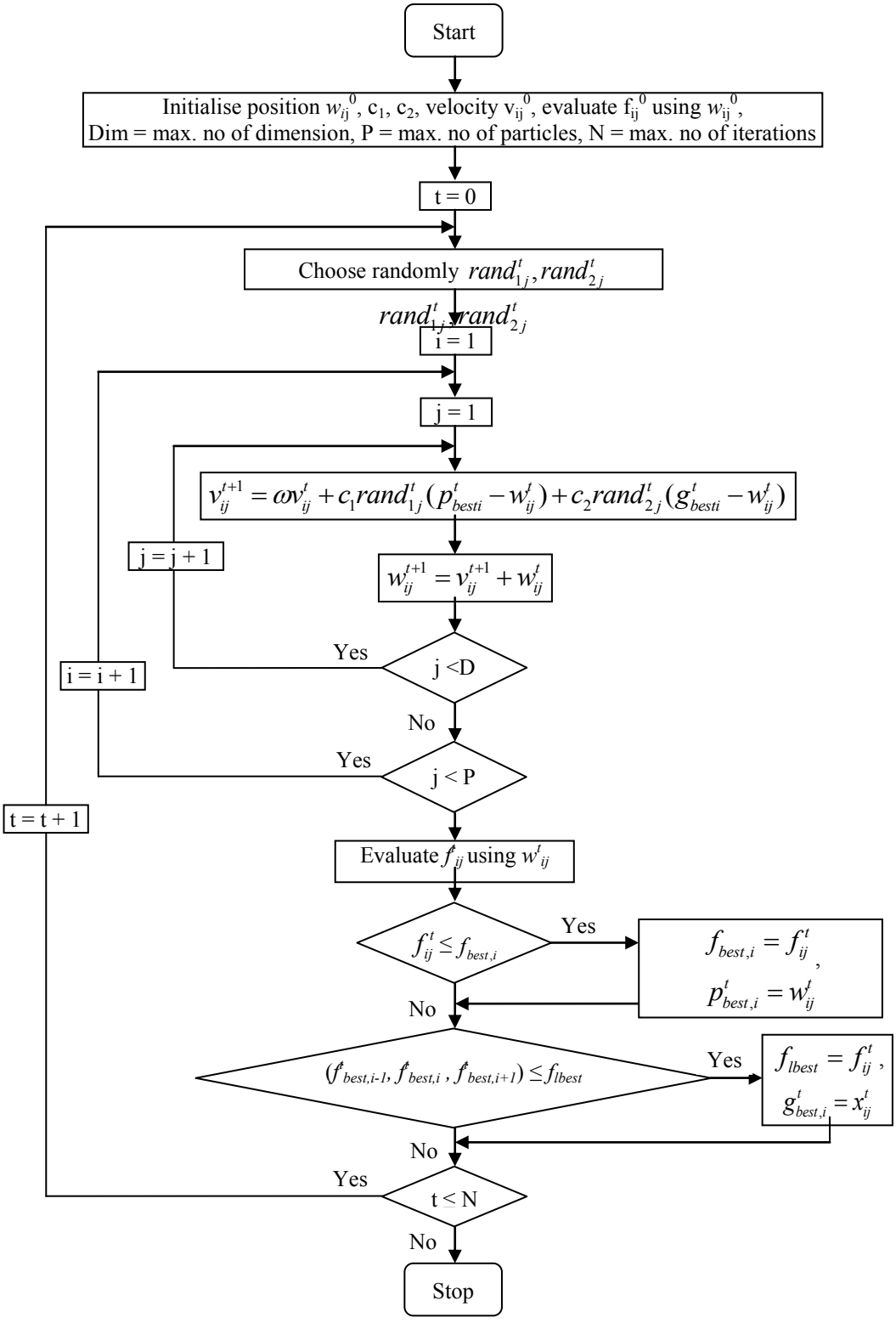
Table 5.1:ANN parameters used in the PSO-NN model

Parameters	Value
Population size	30
Velocity clamping percentage $v_{max\text{-}perc}$	0.2
Maximum number of iterations	1000
Acceleration constants (c_1, c_2)	$c_1 = c_2 = 2$
Inertia weight ω	0.7
Error goal	0.001
Minimum error gradient	10^{-9}
Dimension	Based on ANN architecture

Table 5.2: PSO parameters used in the PSO-NN model

TRAINPSO, a type of network training function, is used as the network training function that updates weight and bias values according to particle swarm optimisation. This is a modification of existing neural network training algorithms provided in Matlab's Neural Network toolbox, using ideas from Brian Birge's PSO toolbox (Birge, 2005). This PSO training function aims to be the interface between

the Matlab NN toolbox and the PSO Research Toolbox. This approach to training a NN by PSO treats each PSO particle as one possible solution of weight and bias combinations for the neural network. The PSO particles therefore move about in the search space aiming to minimise the output of the network performance function (Rambharose, 2011).



Flowchart 5.1: Local best PSO algorithm

The size of the feed-forward neural network is set as [9 *H* 1], which means there are nine neurons in the input layer according to nine input features and one neuron in the output layer of the network which indicates the state of hyper or normo. The hidden nodes "*H*" in the hidden layer are varied from 3 to 20 in order to find the optimal number of hidden nodes in the network. The Linear transfer function is the type of transfer function used for the last layer and hyperbolic tangent sigmoid transfer function is used for the rest of the layer of the designed neural network. Gradient descent with momentum weight and bias learning function is used as Back propagation weight/bias learning function.

Data input is the diabetic dataset from 10 T1DM patients, which was mentioned in chapter 3 and divided into training, validation and testing exemplars (with proportions of 35%, 35% and 30% out of overall data, respectively).

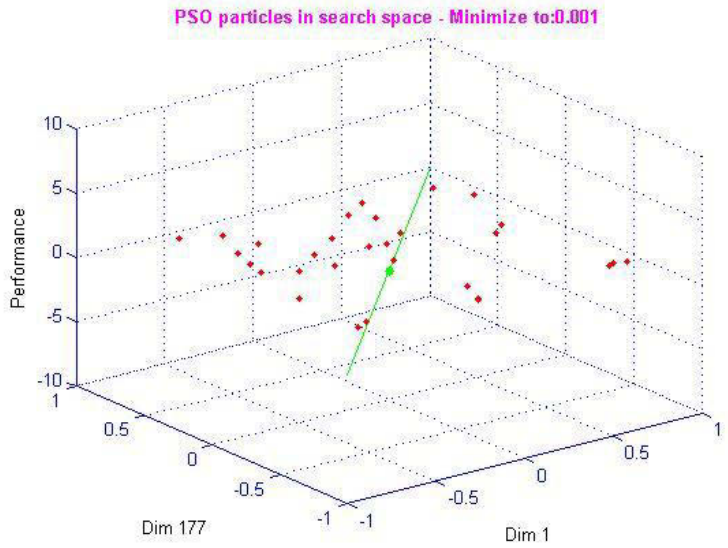
5.3 EXPERIMENTAL RESULTS

5.3.1 Performance of PSO-NN

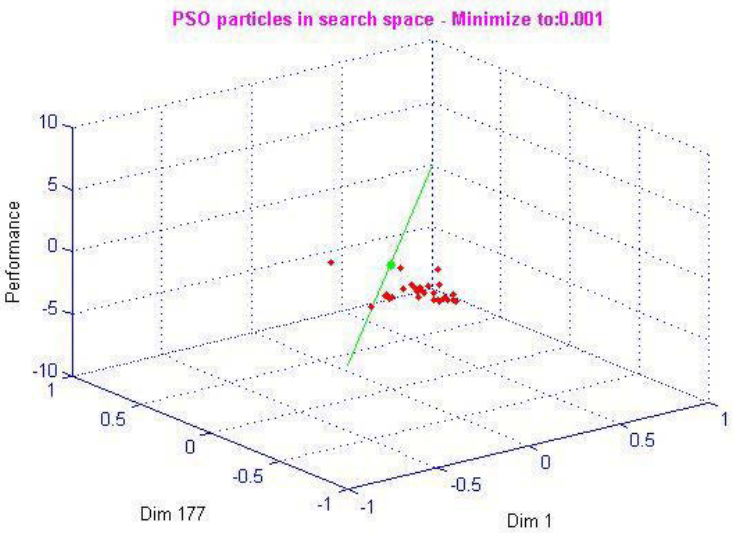
Diabetic data sets are studied using neural network trained based on PSO algorithm. The architecture of neural network is obtained by varying the number of neurons in the hidden layer from 3 to 20. After initialising training variables (i.e., training data parameters, training record parameters, training state parameter, to keep track of as the network is trained), PSO Research Toolbox is called to adjust network weights and biases, which is aimed to determine new weight such that error performance is minimized. Swarm motions in the search space are graphed on the 3-D coloured contour maps of Figure 5.2 (a, b, c, d), where particles can be seen flying from random initialization to eventual stagnation at the local minimiser.

The true global minimiser at [0,0] (which is spotted in the centred green point) is not discovered. A particle finds the relative quality region near the just found local minimiser and communicates this new global best to the rest of the swarm. As other particles fly in its direction, none finds a better global best, so all converge around

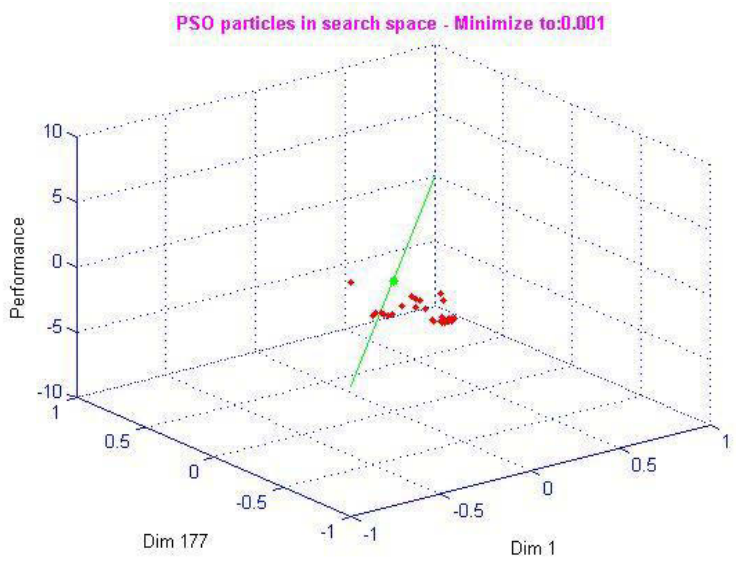
the latest local minimiser as momenta wane. According to the trajectory of PSO particles searching for global best, Figure 5.3 locates the positions of global best in one trial of PSO training, which are mainly moving around vertical range [-1,1].



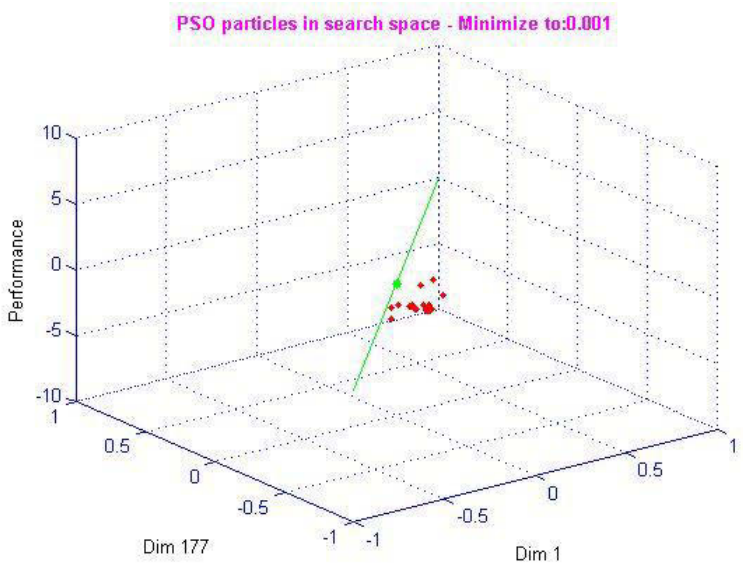
(a)



(b)



(c)



(d)

Figure 5.2: Swarm motions on the 3-D coloured contour maps

(a) Swarm initialisation – Particles are randomly initialised within the search space

(b) Particles are converging to local minimum via their attraction to the global best in the vicinity

(c) Cognitive accelerations toward personal best and momentum keep particles searching prior to settling down

(d) As momentum wane and no better global best is found, the local minimum is being honed in on when the progress reaches to the maximum number iterations.

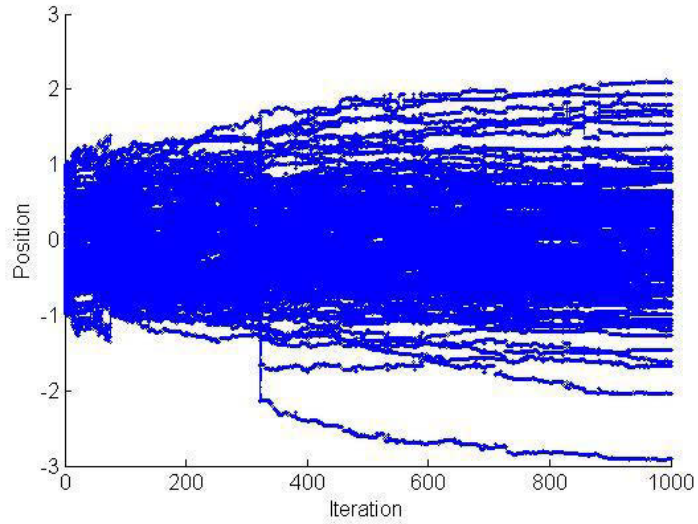


Figure 5.3: History graph of global best positions

The fitness function value of the global best at each iteration number is described in Figure 5.4 and Figure 5.5, with the iteration numbers of 1000 and 1200, respectively. The figures show that the convergence rates of the training by PSO algorithm at two different iteration numbers are nearly the same. The optimal fitness function value with 1000 iterations is 0.0626, which is nearly similar to the function value of 0.0658 with 1200 iterations. It shows that 1000 iterations are enough for the training with PSO to get the optimal fitness function.

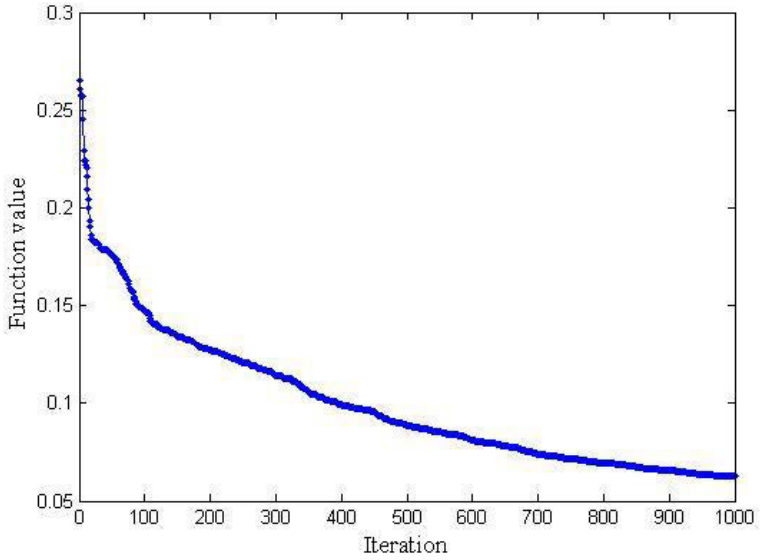


Figure 5.4: The function value of the global best obtained with 1000 iterations

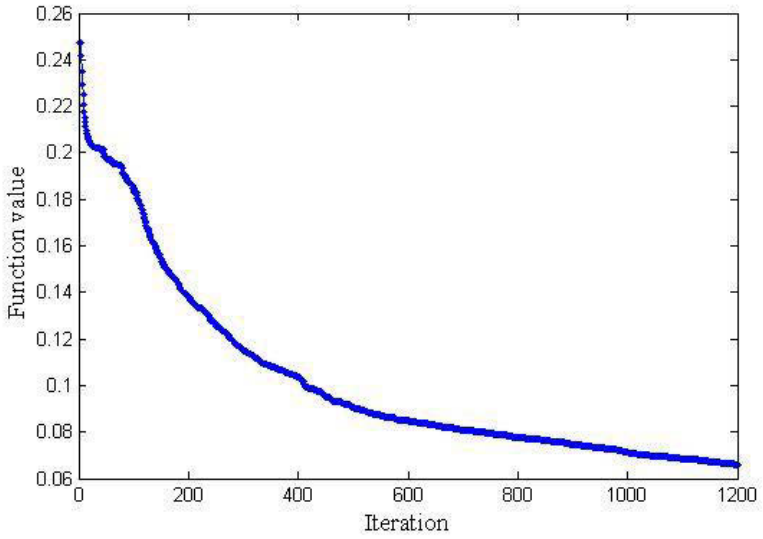


Figure 5.5: The function value of the global best obtained with 1200 iterations

From equation (5.3) it is clear that the particle swarm algorithm proceeds forward by updating the weight of each particle. If the velocity is not constrained, the equation could yield undesirable results since the search space expands and tends to infinity in the hyper space. The parameter v_{max}, which is manipulated by v_{max_perc}, is thus used to dampen the oscillations. The effects of v_{max} on the training and the performance of PSO-NN are thoroughly examined by assuming v_{max_per} values vary from 0.001 to 1. From Figure 5.6, it can be seen that when v_{max_per} is set to 0.001, the performance curve assumes nearly a horizontal line indicating no change in the weights. As the v_{max_per} increases to 0.1, the mean square error reduces to the value of 0.199 during the first 35 iterations, and at v_{max_per} value of 0.2, the MSE falls down to 0.143, giving properly the best performance value.

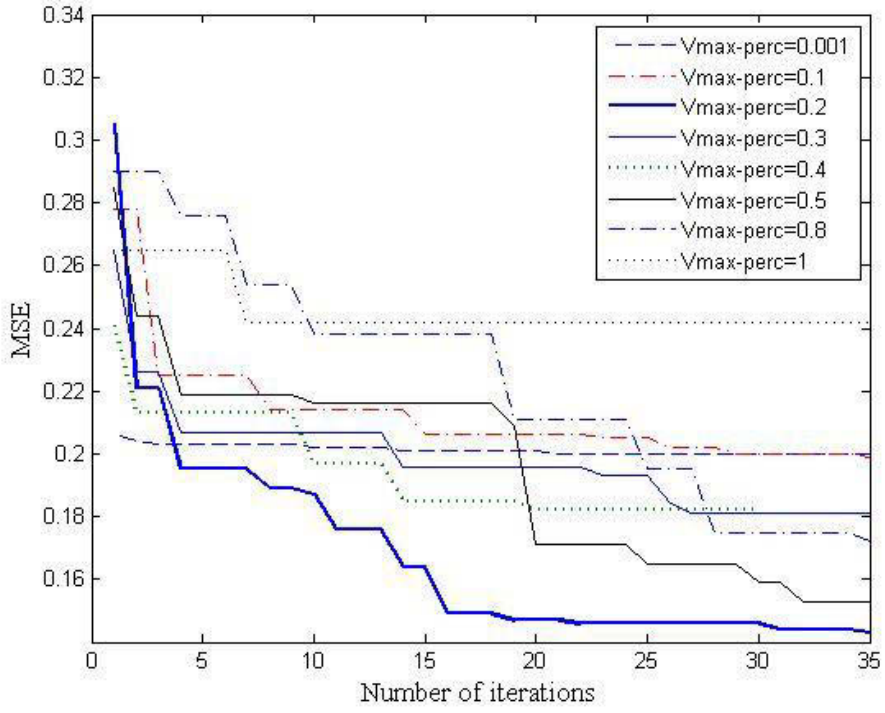


Figure 5.6: Performance curves of PSO-NN at various velocities

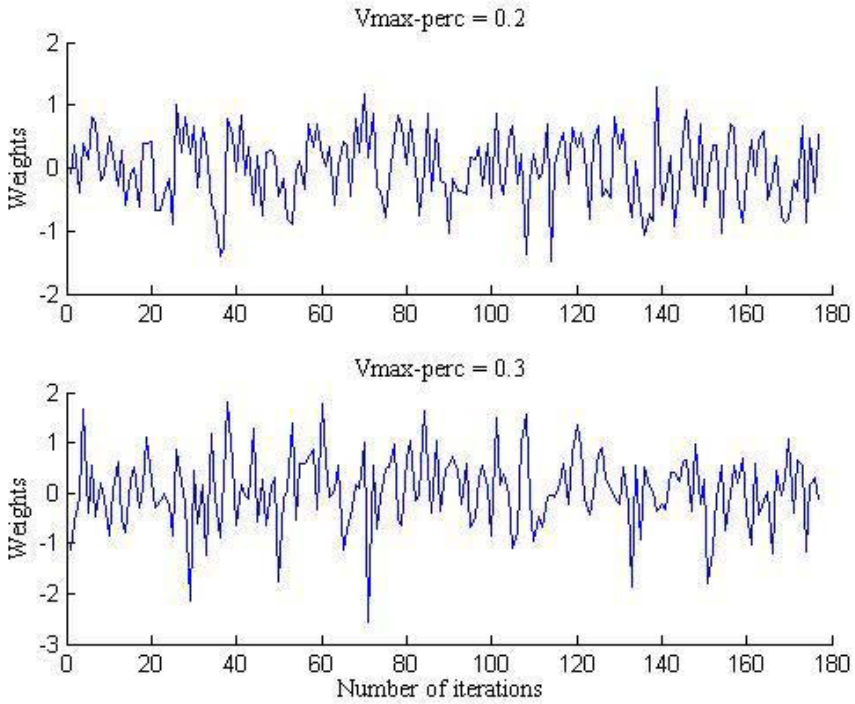


Figure 5.7: Variations of the weights of PSO-NN at various velocities: v_{mac_perc} = 0.2 and v_{mac_perc} = 0.3

Figure 5.7 depicts the variation in the weights of neural network training by PSO algorithm with an increase of the number of iterations. The variations of the weights are observed while fixing the v_{max_per} values at 0.2 and 0.3. Comparing two graphs in Figure 5.7, it can be seen that as v_{max_per} increase to value further than 0.2, the particles converge and thereby, the weights of the neural network increases considerably in a cyclic manner resulting into larger errors. It shows that v_{max_per} value of 0.2 is the optimal choice for the PSO-NN.

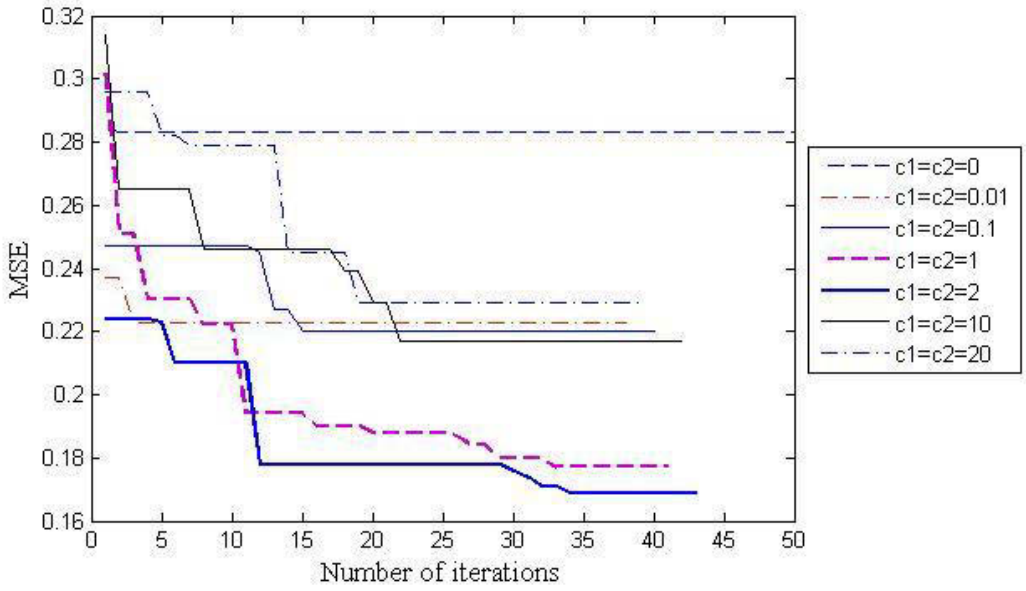


Figure 5.8: PSO-NN Performance curves at various acceleration constants

Acceleration constants have also affected the performance of the PSO algorithm. In this case, it has been assumed that $c_1 = c_2$. Figure 5.8 depicts the convergence curves of PSO algorithm for various values of c_1, c_2. Initially, at a value of acceleration constants equal to zero the MSE curve is more or less than a straight line because of equation 5.2. However, when $c_1 = c_2 = 0.01$, the network shows a very shallow convergence and the MSE value reaches 0.223 and remains constant thereafter. Similarly, when c_1, c_2 increase regularly until they approach to 2, a sharp drop in the value of MSE is observed within the first few iterations and the learning process is almost complete within 35 iterations. Furthermore, as c_1, c_2 increase beyond 2, the MSE value drops very rapidly in the initial iterations, yet higher than

MSE value at c_1, c_2 equal to 2. This shows that at high acceleration constants values, the ability of the network to generalise reduces.

The effect of change in weights of the network due to different values of c_1, c_2 at 1 and 2 are observed to evaluate the chosen value of c_1, c_2 of 2. Figure 5.9 depicts the fluctuations of the particles (value of the weights in the network) with an increase in the acceleration constants from 1 to 2. As c_1, c_2 are set at 2, the curve is seen to oscillate at slightly higher frequency of fluctuation than c_1, c_2 at 1. However, c_1, c_2 at 2 has smaller amplitude compared to the other. The choice of acceleration constants ($c_1 = c_2 = 2$) in PSO is based from the decrease of amplitude rather than the increase of frequency of fluctuation.

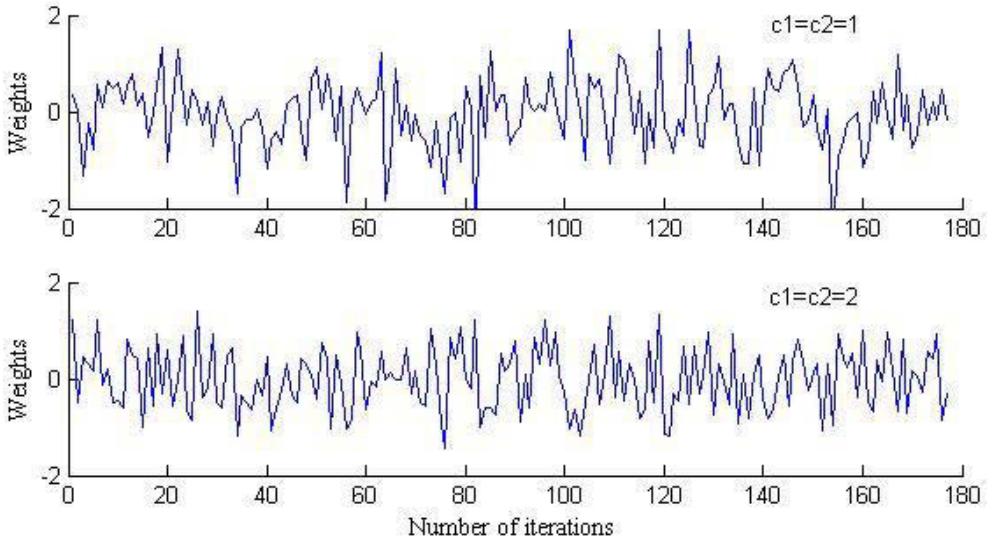


Figure 5.9: Values of weights in PSO-NN at $c_1 = c_2 = 1$ and $c_1 = c_2 = 2$

With the chosen PSO parameters, table 5.3 shows the best results of hyperglycaemia detection at different hidden nodes, varied from 3 to 20. Within 20 runs, the best performance for hyperglycaemia detection is found with 16 hidden nodes, in which the training results of 85% sensitivity and 70% specificity are gained and the testing set leads to 82.35% sensitivity and 73.08% specificity. These testing results not only meet the requirements of disease diagnosis (sensitivity $\geq$ 70% and specificity $\geq$ 50%), but also earn an excellent geometric mean of 77.58%.

Hidden node	Training			Validation			Testing		
	Sen (%)	Spec (%)	*gm* (%)	Sen (%)	Spec (%)	*gm* (%)	Sen (%)	Spec (%)	*gm* (%)
3	65.22	77.78	71.22	60.87	62.96	61.91	60	57.58	58.78
4	77.27	71.43	74.29	44.44	84.38	61.24	62.5	55.56	58.93
5	84.62	66.67	75.11	41.18	57.58	48.69	69.23	63.33	66.22
6	70.37	52.17	60.59	88.24	60.61	73.13	58.33	48.39	53.13
7	71.43	79.31	75.27	62.5	65.38	63.93	63.64	62.5	63.07
8	68.42	80.64	74.28	47.37	77.42	60.59	77.78	72	74.83
9	78.26	66.67	72.23	50	53.13	51.54	73.33	71.43	72.37
10	81.82	64.29	72.52	57.14	58.62	57.88	76.92	60	67.94
11	81.82	42.86	59.22	88.24	30.3	51.71	76.47	42.31	56.88
12	87.5	61.54	73.38	70	56.67	62.98	83.33	51.61	65.58
13	83.33	53.85	66.99	81.25	38.24	55.74	81.25	59.26	69.39
14	90.48	55.17	70.65	71.43	55.17	62.78	78.57	51.72	63.75
15	70.59	60.61	65.41	50	76.67	61.91	68.42	66.67	67.54
16	**85**	**70**	**77.13**	**77.37**	**61.29**	**68.86**	**82.35**	**73.08**	**77.58**
17	75	76.67	75.83	55	66.67	60.55	62.5	59.26	60.86
18	87.5	50	66.14	73.68	41.94	55.59	76.92	40	55.47
19	95.45	71.43	82.57	78.95	54.84	65.8	60	39.29	48.55
20	72.73	71.43	72.07	75	76.47	75.73	66.67	64	65.32

Table 5.3: Performances of hyperglycaemia detection at different hidden nodes

The corresponding ROC plot of PSO-NN for training set is also presented in Figure 5.10. AUC of 0.8279 is found by the PSO-NN, which express a high training accuracy in the performance of hyperglycaemia detection.

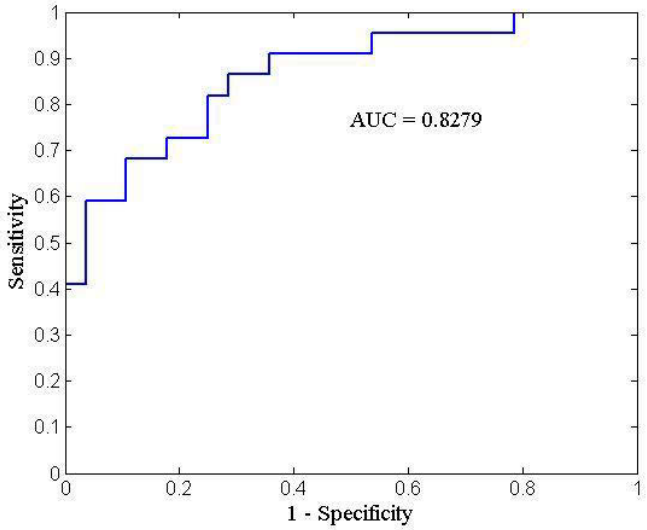


Figure 5.10: The ROC curve of best performance of PSO-NN

5.3.2 Comparison of PSO-NN with other NN-based algorithms

The performance of PSO-NN has been compared to the outcome of Levenberg-Marquardt NN (PCA-LM-NN) using PCA (chapter 4) for hyperglycaemia detection. Over 20 runs, mean values of training, validation and testing results of PSO-NN are higher than those of PCA-LM-NN, in terms of sensitivity, specificity and *gm* (as shown in Table 5.4). Table 5.5 shows the training execution errors of two methods. PSO-NN has lower error rates, which means it has more powerful classification capability and ability in terms of biomedical applications.

Method	Hid. Node	Training			Validation			Testing		
		Sen (%)	Spec (%)	*gm* (%)	Sen (%)	Spec (%)	*gm* (%)	Sen (%)	Spec (%)	*gm* (%)
PCA-LM-NN	6	78.45	74.86	76.63	71.18	65.04	68.04	68.18	59.55	63.72
PSO-NN	16	79.14	70.89	74.9	66.33	58.82	62.46	71.84	62.7	67.12

Table 5.4: Mean values of training, validation and testing results over 20 runs

Method	Mean Absolute Error (MAE)	Mean Square Error (MSE)	Root Mean Square Error (RMSE)
PCA-LM-NN	0.4076	0.1979	0.4449
PSO-NN	0.3755	0.143	0.429

Table 5.5: Training execution errors of PCA-LM-NN and PSO-NN models

For the assessment of agreement and reliability of two classification methods, Cohen's Kappa statistic is also performed with PCA-LM-NN and PSO-NN. The Kappa statistic is used to access the accuracy of any particular classification cases and to distinguish between the reliability of the predicted data and their validity. Once the Kappa statistic is performed, the results usually come along with Z-score and *p* value of the assessment, which is used to take into account that classifiers might sometimes agree or disagree simply by chance or not. Table 5.6 presents the comparison of Kappa statistic between two classification methods based on neural network, PCA-LM-NN and PSO-NN. The results show that PSO-NN has a better Kappa coefficient than PCA-LM-NN and both Kappa scores are around 0.4-0.6, which means the accuracy of the classification purpose is moderately reliable. PSO-NN has a significantly better Z-score than LM-NN (as $p < 0.05$), which means PSO-NN has significant classification results better than random choice compared to the other.

Method	Kappa statistic	Z-score	*p* value
PCA-LM-NN	0.4444	2.4898	< 0.05
PSO-NN	0.5326	3.5552	< 0.001

Table 5.6: Cohen's Kappa statistic of PCA-LM-NN and PSO-NN models

The performance comparison of PSO-NN to previous NN-based algorithms, which are LM-NN (Chapter 3) and PCA-LM-NN (Chapter 4), is presented in Table 5.7. PSO-NN achieved the best geometric mean of 77.58%, whereas those of PCA-LM-NN and LM-NN are much lower (70.88% and 67.94%, respectively). Although the sensitivity of PSO-NN is not as high as PCA-LM-NN, its specificity is definitely

better than the others. Obviously, PSO-NN outperforms the other two algorithms, which prove to be an excellent method for hyperglycaemia detection.

Method	Input	sen (%)	spec (%)	*gm* (%)
LM-NN (Chapter 3)	16 original parameters	70.59	65.38	67.94
PCA-LM-NN (Chapter 4)	9 parameters	85.71	58.62	70.88
PSO-NN (Chapter 5)	9 parameters	82.35	73.08	77.58

Table 5.7: Best testing results of three methods

5.4 DISCUSSION

A neural network model, known as PSO-NN, is developed to detect hyperglycaemia based on the ECG parameters extracted from ECG signals of 10 T1DM patients. The model is the improvement of the two-stage classifier ANN using PCA for hyperglycaemia detection, which has been explored in Chapter 4. The new model keeps the same stage 1, which uses PCA for feature reduction, and stage 2 is an advanced feed-forward neural network employing PSO algorithm to optimise the weights of the network. By means of PSO, the weights of the network are updated every iteration until the optimal weights are obtained. It takes advantage of the PSO algorithm's global searching ability characteristic.

For observing the movement of particles, PSO algorithm was run again by giving the same position and velocity of each particle initialised randomly. The dimension of search space is 177, according to equation 5.6 with 9 neurons in the input layer, 16 neurons in the hidden layer and 1 neuron in the output layer. From Figure 5.2 (a, b, c, d), it is easy to find out that the particles have searched more domains around the searching space then jumped from local optimal solution to converge to the global optimal solution when no better global best was found.

In the training process, the PSO algorithm overcomes the weakness of both algorithms and achieves the goal MSE. The PSO-NN, whose weights were initialised with PSO showed radically better performance than the networks without PSO pre-training, which have been described in Chapter 3 and 4. The parameters, which were estimated previously to use in the PSO algorithm, are found to collaborate well with the hybrid neural network model. The error index values calculated in the comparison between PSO algorithms with other NN-based algorithms indicate that the detection executed by PSO-NN involves minimum possible errors. Thus, PSO could be an alternative training algorithm for neural networks.

Sensitivity analyses for v_{max_per} and acceleration constants c_1, c_2 were also carried out to evaluate their impact on the network convergence and find out the optimal maximum velocity value and acceleration constants for the PSO algorithm with a minimum possible error. It was observed that as the maximum velocity percentage per clamp increases, the particles do not converge to the optimal weights and the weight of the network increase considerably in a cyclic manner thereby resulting in large errors while classifying hyperglycaemia. The evaluations of function values of global best at 1000 and 1200 iterations in Figure 5.4 and Figure 5.5 have shown that 1000 iterations are enough for PSO algorithm to get the optimal fitness function value. Obviously, the 1000 iterations one is far more efficient in terms of time for the training process than the 1200 iterations one. Likewise, a large value of acceleration constants gives rise to a sharp decrease in MSE that contributes to over-fitting. Similarly, a very small value of c_1, c_2 has to be ruled out because it involves high error performance and considerable computation time to predict the desired results. Thus the appropriate values of v_{max_per} and c_1, c_2 are identified by varying them at different values and observing the fluctuation of values of the weights in the network.

The number of hidden nodes of PSO-NN is varied from 3 to 20 to find the appropriate hidden node in order that the network performs well. From Table 5.3, one can observe that for the various hidden nodes, the geometric mean increases to the maximum value at 16 hidden nodes and then decreases gradually. Thus, the optimal neural network architecture using PSO algorithm is found to have a structure

9-16-1. That means there are 9 neurons in the input layers indicating 9 ECG input parameters, 16 neurons in the hidden layer and 1 neuron in the output layer. With this structure, the PSO-NN has achieved the best performance for hyperglycaemia detection.

Cohen's Kappa statistical analysis was also carried out with PSO-NN and PCA-LM-NN, the method has been used in the previous chapter. This analysis aims to assess the agreement of two classifiers with decisions made by chance and the reliability of their predicted date and their validity. As the disagreement increases, the point of kappa will decline more quickly than a proportion of agreement will, because of the penalty for chance agreement. With Kappa values over 0.4, both PSO-NN and PCA-LM-NN are in the moderate range between our classifiers and random choice. However, the PSO-NN model gave higher point of Kappa than PCA-LM-NN which means PSO-NN is the more reliable one.

From the figures and tables, it is clear that the PSO algorithm is more accurate when compared to the other algorithms. The inclusion of PSO algorithm in ANN contributes to a high improved outcome and is thus a most suitable algorithm for hyperglycaemia detection. This can be seen by comparing the results of PSO-NN with LM-NN (chapter 3) and PCA-LM-NN (chapter 4). PSO algorithm provides good sensitivity (82.35%) and a significantly higher specificity than the other two methods (73.08% vs. 58.62% and 65.38%). As a result, PSO-NN obtains much higher geometric mean than LM-NN and PCA-LM-NN (77.58% vs. 70.88% and 67.94%), which means PSO-NN performs far better than the others.

In this study, the parameters of the PSO-NN are decided by manually varying the values of maximum velocity and acceleration constants. However, an effort can be made in the future to optimise PSO parameters for neural network training by using PSO. Furthermore, the PSO can also be incorporated into a hybrid PSO and Levenberg-Marquardt algorithm to improve the global convergence rate. In this study, PSO-NN is employed to detect hyperglycaemic episodes of 10 T1DM patients. In the future, it could be applied to more complex and large data sets and can be extended for multi-stated classification as well.

PSO algorithm employed in a multilayer feed-forward neural network (PSO-NN) has been developed and tested for hyperglycaemia detection. The adaptation of PSO in learning neural network aims to further enhance the performance of hyperglycaemia detection. PSO is a more advanced algorithm compared to the pure LM and combination of PCA and LM which developed in Chapters 3 and 4. In PSO-NN, PCA has also been used as stage 1 to find the optimal ECG parameters before feeding them into the input of neural network. PSO algorithm is then used in stage 2 to train the network and update the weights until optimal weights are obtained. The optimisation of the weights in the network resulted in the improvement of classification performance. The PSO approach outperformed the classical LM and PCA-LM algorithms. It gave significantly higher geometric means than the two others in terms of mean values over 20 runs and best testing results. Cohen's Kappa statistical analysis has also been done to compare three methods which prove PSO-NN as a very effective model. Results of the investigation in this study reveal that the proposed method is prominent and useful for the classification of hyperglycaemia.

CHAPTER 6 .
DISCUSSION, CONCLUSIONS AND FUTURE DIRECTIONS

6.1 DISCUSSION

6.1.1 Hyperglycaemia and abnormalities in ECG signals

In this thesis, the association between hyperglycaemia and ECG signals has been investigated. It was revealed that hyperglycaemia can result in the shortening of ventricular repolarisation intervals such as *QT*, *RT*, and T_PT_E; increased *PR* intervals and reduced heart rate variability. The aim of this thesis does not involve clarifying the mechanisms of abnormalities of ECG features found in hyperglycaemic episodes, although some suggestions are still proposed for the clinical manifestation of hyperglycaemia in regards to the effect of this complication on the electrical activity of the heart.

Interestingly, earlier studies have shown that hypoglycaemia also affects cardiac repolarisation but in the opposite way, particularly on the prolongation of *QT* interval (Robinson *et al.*, 2004). Basically, the mechanism of *QT* changes under hypoglycaemia is now clear. Two main mechanisms can account for the prolongation of the *QT* interval in hypoglycaemia. One is the sympatho-adrenergic system activated by hypoglycaemia that releases adrenaline and non-adrenaline into the circulation. The other is the decreasing potassium level caused by the increased

insulin and adrenalin levels (Lee *et al.*, 2003). In hypoglycaemia, repolarisation is prolonged by the inhibition of a rapid component of the delayed rectifier potassium current which plays an important role in repolarisation (Zhang *et al.*, 2003), thus *QT* interval will be prolonged due to the increased repolarisation time.

In contrast, the mechanism of the abnormal alterations of *QT* interval in hyperglycaemia is poorly understood. There is still controversy in reaching a final confirmation as to whether *QT* comprehensively appears to shorten or lengthen under hyperglycaemia conditions. A report had suggested that hyperglycaemic glucose clamp induces QT_C interval prolongation in diabetic patients and in healthy control subjects (Gordin *et al.*, 2008). Marfella proposed that hyperglycaemia may produce ventricular instability by increased sympathetic activity, increased cytosolic calcium content in myocytes or both (Marfella *et al.,* 2001). Inversely, in a recent study (Suys *et al.*, 2006) where simultaneously recorded QT_C values and glucose levels were analysed in patients with Type 1 Diabetes, QT_C prolongation was associated only with hypoglycaemia, not with hyperglycaemia. In this thesis, the subjects for clinical study were selected to encounter only hyperglycaemic and normoglycemic events, without the interference of hypoglycaemic events which may weaken the evaluation of the influence of hyperglycaemia in changes of cardiac ventricular intervals. A significant decrease in QT_C interval has been seen among all ten patients in the study, which is in agreement with Suys's finding. Results from clinical studies are less consistent because of the different background of studies, and then in order to establish the clinical relevance of any findings, further studies are needed.

Clearly, the diagnosis of short cardiac ventricular repolarisation intervals can be challenging to establish. Shortened *QT* interval may be the result of a number of reasons, such as genetic channelopathies, hypercalcaemia, hyperkalaemia, acidosis, increased vagal tone, after ventricular fibrillation, digitalis use, and androgen use (Maluli *et al.*, 2013). Within this study context, one of the causes of short *QT* interval is suggested to relate to the elevated serum potassium level in hyperglycaemia of diabetes mellitus with long-term dialysis. Serum potassium concentration may be elevated because of an extracellular shift of potassium caused by insulin deficiency (ADA, 2004). That insulin deficiency is a major determinant of hyperkalemia in

hyperglycaemia has been concluded from recent studies of different hyperglycaemic states (Tzamaloukas *et al.*, 1987; Tzamaloukas *et al.*, 2011). Severe hyperkalemia will produce changes on the ECG signal, such as a prolonged *PR* interval and a shortened *QT* interval (Astle, 2005), which are seen in this research.

Recent studies have identified a possible link between reduced heart rate variability (HRV) and diabetes, particularly among patients with Type 1 diabetic mellitus (Singh *et al.*, 2000; Jaiswal *et al.*, 2013). The main driver for this abnormality appears to be high blood glucose concentration. The autonomic nervous system (ANS) regulates heart rate (HR) through sympathetic and parasympathetic (vagal) branches. HRV is reduced in diabetes mellitus (DM) patients, suggesting dysfunction of cardiac autonomic regulation which has been associated with increased risk for cardiac events (Tarvainen *et al.*, 2013). The mechanism of this phenomenon is due to the association of the autonomic nervous system and glucose tolerance. In healthy individuals, when blood glucose levels fall the sympathetic branch is activated, stimulating glucose production in the liver and kidney and reducing muscle use of glucose by way of adrenaline release. Meanwhile, parasympathetic (vagal) stimulation has the opposite effect: insulin release from the pancreas, which stimulates glucose uptake by cells; reduction in glucose release from tissues and increased glycogen formation by the liver. Hypoglycaemia, or low blood sugar, leads to sympathetic activation increasing blood glucose levels, while hyperglycaemia, high blood sugar, results in parasympathetic activation to reduce blood glucose levels. These two systems work in unison to maintain healthy blood glucose levels. Therefore, impaired parasympathetic regulation (decreased HRV) increases risk for chronic hyperglycaemia and hyperinsulinaemia (raised insulin levels, also known as insulin resistance), which is a pre-cursor for diabetes mellitus. This finding is consistent with known increased risk among diabetics for autonomic neuropathies, which can lead to loss of sensation in the peripheries and in extreme cases, infected wounds resulting in limb loss (often known as diabetic foot) (Wensley, 2013).

While there is a vast amount of literature on disorders of glucose tolerance and HRV, there are still many gaps in our understanding of this relationship. It

remains unclear, for example, whether elevated levels of insulin or glucose reduce HRV or whether impaired vagal function is a causal risk factor for the development of diabetes. Studies of HRV in healthy individuals followed up for many years or genetic studies may help to clarify this correlation.

6.1.2 Comparison of experimental results

Experimental results of the applications on hyperglycaemia detection are presented in Table 6.1. Three proposed methods, LM-NN, PCA-LM-NN, and PSO-NN are compared and analysed.

Method		Sensitivity (%)	Specificity (%)	*gm* (%)
LM-NN	Training	89.47	83.87	86.63
	Validation	80	77.41	78.69
	Testing	**70.59**	**65.38**	**67.94**
PCA-LM-NN	Training	88.89	71.88	79.63
	Validation	66.67	57.69	62.02
	Testing	**85.71**	**58.62**	**70.88**
PSO-NN	Training	85	70	77.13
	Validation	77.37	61.29	68.86
	Testing	**82.35**	**73.08**	**77.58**

Table 6.1: Best performances of Training, Validation and Testing sets of three proposed methods

The performance of PCA-LM-NN is better than LM-NN (70.88% vs. 67.94%) in terms of geometric mean. Although PCA-LM-NN and LM-NN employ the same learning algorithm (Levenberg-Marquardt), PCA-LM-NN applies feature reduction using principal component analysis (PCA) to get a smaller-size group of ECG variables as input of the network, whereas the LM-NN uses a total of sixteen original ones. The new subgroup of ECG variables found by PCA has significant ECG

parameters which are mostly correlated with high blood glucose levels. The multicollinearity among original ECG variables lowers the results of LM-NN. Thus, better performance of PCA-LM-NN is the consequence of efficiently choosing optimal inputs before feeding them into the network. It can be seen that the sensitivity of PCA-LM-NN is much higher than that of LM-NN (85.71% vs. 70.59%). It is because the PCA-LM-NN handles only optimal inputs which are significantly responsible for the variation of high blood glucose. This may lead to an increased likelihood of correct identification of hyperglycaemic events. Therefore, optimal ECG variables are still utilised in the next proposed method (the PSO-NN) to secure high classification performance.

Comparing the PSO-NN and PCA-LM-NN, the performance of PSO-NN outperforms the PCA-LM-NN (77.58% vs. 70.88%, in terms of geometric mean). The PSO-NN employs ANN structure and uses a particle swarm optimisation (PSO) to optimise the training of the weights of network, whereas PCA-LM-NN does not. The better performance of PSO-NN could be as a result of the swarm optimisation effectively optimising the ANN weights by particles. Local best PSO algorithm is used to train the network instead of global best PSO. The global version is faster but might converge to local optimum for some problems, whereas the local version is a little bit slower but not easily trapped into local optimum.

Tuning PSO parameters in order to minimise MSE value in the implementation of the PSO-NN also affects its better results. The concerned PSO parameters are velocity clamping percentage ($v_{max\text{-}perc}$) and acceleration constants ($c_1 = c_2$). In a PSO, velocity clamping controls the global exploration of the particle. It reduces the size of the step velocity, so that the particles remain in the search area, but does not change the search direction of the particle. Considering the term, the $v_{max\text{-}perc}$ value is a trade-off between the margin space and error. If the value is too high, the particles will search far beyond the best solution. If the value is too small, the particles cannot perform enough searches and soon get trapped in the local minimum. The acceleration constants, c_1 and c_2, represent the weighting of the stochastic acceleration terms, that pull each particle toward positions p_{besti} (best previous position) and g_{besti} (global best position). Thus, adjustment of this factor

changes the amount of 'tension' in the system. Low values allow particles from far target regions to explore the search space before being tugged back, while high values result in abrupt movement toward the target (Eberhart *et al.*, 1995). Therefore, it is important to finalize the appropriate values of velocity clamping and acceleration constants, so that the particles will soon effectively converge in the local area and give the best fitness value.

	PSO-NN	**PCA-LM-NN**	**LM-NN**
ANN structure	[9 16 1]	[9 6 1]	[16 9 1]
Learning algorithm	PSO	LM	LM
Stopping epochs	35	12	9
Sensitivity and specificity requirements	Satisfactory	Satisfactory	Satisfactory
Overall performance	**Best**	**Good**	**Acceptable**

Table 6.2: Comparisons of three proposed methodologies

Comparisons among three proposed methodologies in terms of ANN structure, learning algorithm, number of epochs at stopping, satisfactory sensitivity and specificity requirements, and overall performance are rated in Table 6.2. It can be seen that the structure of LM-NN is heavier than the other two due to the input features which have not been reduced yet. Three models seek for the proper number of hidden nodes, which is varied from 3 to 20. It does not matter which one has the smaller number until they achieve the best results. However, the more hidden nodes the longer the training process may take, which has been proved in higher epochs at stopping the PSO-NN. That means the PSO-NN may need a longer convergence time as it takes the local best PSO algorithm to avoid trapping in local minima. LM algorithm has the major advantage of faster convergence speed, but without any optimisation in network parameters in PCA-LM-NN and LM-NN, it could not help to improve their performances. In terms of sensitivity and specificity requirements (sensitivity $\geq$ 70%; specificity $\geq$ 50%), all three methods are satisfactory, which meets the requirements for biomedical applications.

In conclusion, the overall performance of three proposed methodologies surpasses the conventional methods such as LDA, KNN and classifiers employing other back propagation algorithms in terms of sensitivity, specificity and geometric mean. Handling with smaller-size and more effective data helps PSO-NN and PCA-LM-NN get better results than LM-NN. The optimisation of ANN weights in the PSO-NN assists to enhance its performance and is the best among three proposed methods. PSO-NN might be a potential solution for hyperglycaemia detection and has good prospects for further improvement.

6.2 CONCLUSIONS

6.2.1 Summary and conclusions

The objectives of this thesis were to develop an appropriate methodology for the classification of hyperglycaemic and normoglycaemic states using electrocardiographic signals.

To achieve this goal, firstly, an extensive literature review was conducted in chapter 2, including the major background of hyperglycaemia, the reported effect of hyperglycaemia on the electrical activities of the heart and related work in the detection of hyperglycaemia. The blood glucose threshold of identifying the hyperglycaemic state was set at 8.33 mmol/l (or 150 mg/dl). It was revealed that hyperglycaemia results on ECG signals differ from hypoglycaemia. For example, hyperglycaemia may cause significantly shorter ventricular repolarisation intervals such as *QT*, and reduced heart rate variability. Hence, these abnormal changes in ECG signals due only to hyperglycaemia, not hypoglycaemia or normoglycemia, can become crucial markers for hyperglycaemia detection in diabetic patients. However, it was a surprise that as far as we are concerned, there has been no research using ECG parameters for the detection of hyperglycaemia. The hyperglycaemia classifiers were reported in a number of studies but they were not adequately investigated.

Chapter 3 was dedicated to explain the experimental environment involving a clinical study, which has been utilised as data samples generated for hyperglycaemia detection throughout the thesis. A number of hyperglycaemic and normoglycaemic datasets were generated from ten selected type 1 diabetic patients, who had participated in an overnight hypoglycaemia and hyperglycaemia study at the Princess

Margaret Hospital for Children in Perth, Australia. Sixteen ECG parameters were chosen to identify hyperglycaemia and normoglycaemia regarding the effect of hyperglycaemia on ECG intervals and heart rate variability parameters. A LabView program was developed to acquire ECG signals and extract certain ECG intervals. The Kubios HRV analysis tool has also been used to produce essential HRV parameters. The total of sixteen ECG parameters obtained from the study has gone through statistical analysis to confirm the significant changes in ECG parameters under hyperglycaemic conditions compared to normoglycaemic conditions and clearly emphasised. The significant abnormalities observed are the considerable shortening of RT_C, QT_C, $TpTe_C$ intervals, and a major increase of PR intervals and reduced certain HRV parameters. These are very important findings as they are considered to be the unique input factor for the strategy of non-invasive hyperglycaemia detection, and totally different from the conventional methods of blood glucose monitoring for determination of hyperglycaemia.

Chapter 3 also involved the approach of the artificial neural network classifier for detecting hyperglycaemia using those ECG parameters as input for the network, while two commonly used k-nearest neighbours (KNN) and linear discriminant analysis (LDA) classifiers were also implemented for comparison purposes. The results show that hyperglycaemic events can be detected non-invasively and effectively by sixteen chosen ECG parameters and ANN approach. Compared to the performance of the ANN model using Levenberg-Marquardt for hyperglycaemia detection, LDA and KNN still experienced worse results with lower sensitivities and their sensitivity values were much less than their specificities in all three sets of training, validation and testing. This confirmed the advantage of ANN over other traditional classifiers for the classification of hyperglycaemia.

The integration of a feature reduction technique based on principal component analysis (PCA) with ANN classifier was proposed as a means to select optimal ECG parameters and improve the performance of hyperglycaemia detection. This new two-stage classifier was discussed in chapter 4 for accurate and robust ECG classification based on a subgroup of nine variables reduced by using PCA. PCA is usually useful for dimensionality reduction to diminish the computational requirement and to overcome the problem of multicollinearity. Exploration of the relationships among the independent variables can be used to filter the data so that only the significant independent ECG variables responsible for the high blood glucose levels can be determined. Rather, rearrangement of the input ranked by principal component scores of the data is also utilised to improve the performance of

feed forward multi-layer neural network employed in the second stage. PCA has been used in cardiac pattern classification studies to support various machine learning algorithms to achieve better results. Five different training methods were also applied to the neural network model to compare their performances and find the appropriate one. Cross correlation (Pearson correlation) between 16 original ECG variables was also examined in this study and confirmed strong associations between ECG variables. There have existed high correlations between each component among cardiac intervals and HRV measures. It proved that there was a significant amount of multicollinearity among ECG parameters, which may lead to high standard errors of the parameters' estimation in neural networks later. PCA is one of the approaches to avoid this problem. Based on these PCA results, the final set with nine features (namely, *VLF*, *TotalPw*, *LF*, *SDNN*, RT_C, QT_C, *pNN50*, $TpTe_C$ and *TINN*) are selected as inputs of classification. The results of hyperglycaemia detection using LM algorithm achieved the highest performance with the testing set leading to 85.71% sensitivity and 58.62% specificity. The best testing geometric mean of 70.88% was also gained with these sensitivity and specificity results. It has been shown clearly that the obtained testing LM result outperforms other algorithms such as SCG, GDX, RP and CGF. To evaluate the results of using PCA to reduce features dimensionality, hyperglycaemia detections utilizing neural network which did not use PCA in the first stage, are also performed. As expected, multicollinearity problem in the data set led to a worse performance, which means PCA shows significant contributions to the performance of hyperglycaemia detection.

In order to enhance the performance of hyperglycaemia detection gained from chapter 4, chapter 5 was pursuing another advanced technique, which has the potential to increase classification accuracy. It is the implementation of particle swarm optimisation (PSO) into an ANN as training algorithm (as known as PSO-NN). Investigations have been carried out based on the structure of the two-stage classifier, which has been explored in the previous chapter, and PSO is used to train the network. It was revealed that PSO-NN approach produced noticeable improvements in classification performance compared to the results of any individual classifiers, which have been introduced before. The PSO-NN clearly outperformed other approaches in the same data set and gave the highest geometric mean of 77.58%. With any biomedical applications, the outcome of sensitivity and specificity are expected to meet the requirements (see section 3.3.2.1). The best testing performance of PSO-NN, which is 82.35% and 73.08% in terms of sensitivity and

specificity, respectively, is much more than required. This confirmed the value and effectiveness of the proposed method for the classification of hyperglycaemia.

6.2.2 Main findings

The first important finding in this work is the exploration of the effect of hyperglycaemia on the ECG parameters. There has not been much information in the literature about ECG signs of hyperglycaemia. Several studies have reported about the correlations between ventricular repolarisation parameters, such as QT interval, RT interval and T_PT_E interval, however it is still not figured out properly. This research carried out a new clinical study of a group of ten adolescent T1DM patients who only suffer from hyperglycaemia and normoglycemic to evaluate the effect of hyperglycaemia to the abnormalities of ventricular repolarisation variables. Reduced heart rate variability and increased PR interval have also been seen among hyperglycaemia patients. It extends to further observations by going across the spectrum of hyperglycaemic and normoglycaemic events. Consistent with previous studies, this work reveals that high blood glucose levels correlate with lower HRV in diabetic patients. It must be emphasised that this study does not intend to explore the clinical reasons why hyperglycaemic conditions can result in changes of ECG parameters. Nevertheless, the study has given suggestions about the clinical manifestations of hyperglycaemia and hypoglycaemia.

Secondly, those ECG parameters have been used as input for exploiting the benefits of ANN approach for hyperglycaemia detection. The ANN classifier using Levenberg-Marquardt (LM) as a training algorithm produces good classification results, which are better than traditional classifiers such as LDA and KNN.

Another major finding is that this research has identified the most important ECG variables among sixteen cardiac intervals and HRV measures that are responsible for the variation of high blood glucose levels in subjects with diabetes. Multicollinearity has been seen to exist with high correlation among ECG variables, which could trigger incorrect hyperglycaemia classification. A new group of appropriate variables should be determined for better performance of hyperglycaemia detection. PCA has been employed to transform sixteen original ECG parameters

into a subgroup of nine variables to diminish the computational requirement and overcome the problem of multicollinearity. Exploration of the relationships among the independent variables can be used to filter the data so that only the significant independent ECG variables responsible for the high blood glucose levels can be determined.

Fourthly, a newly developed method of integration of PCA for feature reduction and ANN classifier is an efficient method for hyperglycaemia detection. Such a hybrid approach ensures selection of the best ECG variables and is capable of improving the results. This method using LM algorithm outperforms both non-feature extraction approach and non-LM ANN classifier. It shows great capability to avoid multicollinearity with very high obtained sensitivity and geometric mean.

Last but not least, the integration of PSO based on ANN classifier (PSO-NN) is proposed for hyperglycaemia detection and has proved to be a very robust and effective method. This method can exploit advantages of both PSO and ANN techniques to improve the classification performance. It was found that this integrated approach earns the best classification results, and outperforms other introduced methods, with the sensitivity, specificity and geometric mean of 82.35%, 73.08% and 77.58%, respectively.

6.3 FUTURE DIRECTIONS

Concerning the context of hyperglycaemia detection, instead of limiting only two states, hyperglycaemia and normoglycemia, the first and foremost future work should extend to three states in a row, hypoglycaemia, normoglycaemia and hyperglycaemia. That means the study will involve the interference of hypoglycaemia to the effect of hyperglycaemia on ECG parameters and conduct hyperglycaemia detection based on a database of three groups of blood glucose levels (BGLs). Two thresholds of blood glucose will be set to collect data sets according to each group, BGLs $\leq$ 3.33mmol/l for hypoglycaemic state; 3.33 mmol/l $<$ BGLs $<$ 8.33 mmol/l for normoglycaemic state; and BGLs $\geq$ 8.33 mmol/l for

hyperglycaemic group. This expanding will lead to a more complex classification system but offers closer to problems often occurring in real life of diabetic patients.

In terms of the algorithms that are employed for hyperglycaemia detection, more advanced neural networks could be further designed and developed to improve the performance. The utilisation of PSO in ANN architecture has not been explored utterly. A hybrid of PSO with Levenberg-Marquardt (LM) neural network could be proposed to exploit the advantages of both algorithms. In particular, the LM algorithm can achieve faster convergent speed around the global optimum and the PSO can avoid being trapped in the local minimum. A hybrid algorithm with a transition from PSO search to LM training might be an effective way to train the neural network, which uses the PSO algorithm to do a global search in the beginning stage, and then uses the LM algorithm to do the local search around the global optimum. This hybrid learning algorithm can optimise the weights and threshold selections, improve the learning ability and overall performance of a feed forward neural network. Besides, other advanced modified PSO algorithms, such as Dynamic PSO and Chaos PSO (Huynh *et al.*, 2012), may try to optimise ANN parameters. Constriction factor method (Li-ping *et al.*, 2005) is another potential direction on determining optimal choice of PSO parameters rather than manually tuning. Moreover, with regards to the optimisation techniques, apart from PSO and GA, other evolutionary algorithm such as Differential Evolution (DE) may be applied for ANN training. While GA is more well-established because of its much earlier introduction, the more recent PSO and DE algorithms have started to attract more attention especially for continuous optimisation problems. Similar to PSO, DE was proposed for global optimisation over continuous search space. Its theoretical framework is simple and requires a relatively few control variables but performs well in convergence. Therefore, utilising DE could be a good idea to optimize the performance of hyperglycaemia detection. In addition, other computational intelligences, for example support vector machine and fuzzy reasoning system, which have been well applied to hypoglycaemia detection, might be well adapted to hyperglycaemia detection with promising outcome.

Regarding the mechanism of hyperglycaemia contributing to the alterations of ECG parameters, the phenomenon has not been clearly understood. Previous studies have not supported enough evidence to claim adequate explanation. This research involves ten T1DM patients only and although it has pointed out obviously the significant difference in the changes of electrocardiogram between hyperglycaemia and normoglycaemia, further work in a larger population is absolutely necessary to investigate and clarify those observations. Expanding the database to a broader range including hypoglycaemia, normoglycaemia and hyperglycaemia also gives a greater chance to confirm the findings in this thesis and may take a clear look at possible causes of inverse effects on ventricular repolarisation in diabetics between hypoglycaemic and hyperglycaemic states.

Moreover, the results of this study might require further validation using different ECG feature extraction methods to yield higher performance. Minimising error in ECG measurement and feature extraction might reduce misclassified events. Further validation and improvement of classification methods in this thesis could assist in the application of a real time device for hyperglycaemia detection.

APPENDIX A. ARITIFICIAL NEURAL NETWORKS

Artificial neural networks (ANNs) are computational models inspired by the functioning of the human brain. They consist of simple but highly interconnected computing devices, each of which imitates the biological neurons. Research in the field of artificial neural networks has attracted increasing attention in recent years. A very important feature of artificial neural networks is their adaptive nature which makes such computational models very appealing in applications, especially where there is little or incomplete understanding of the problem to be solved but where training data is readily available.

In all neural network models, a single neuron is the basic building block of the network. The operation of a neuron is modelled by mathematical equations and the individual neurons are connected together as a network. Each neural network has its learning laws according to which it is capable of adjusting parameters of the neurons.

The neuron

The neuron is a program that learns concepts, i.e. it can learn to respond with *True* (1) or *False* (0) for inputs we present to it, by repeatedly "studying" examples presented to it. The structure of a single neuron is very simple. There are two inputs, a bias and an output. A simple schematic diagram is shown in figure A.1.

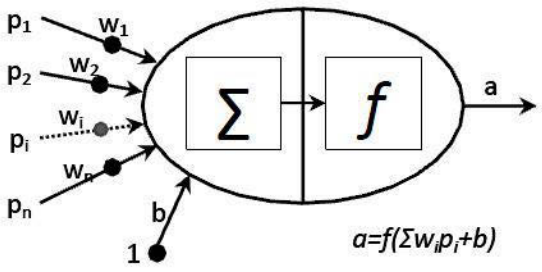


Figure A.1: An artificial neuron with one input and bias

The scalars inputs p_i are transmitted through connections that multiply their strength by the scalar weight w_i to form the product $w_i p_i$, again a scalar. All the

weighted inputs $w_i p_i$ are added and to $\sum w_i p_i$ we also add the scalar bias b. The result is the argument of the transfer function f, which produces the output a. The bias is much like a weight, except that is has a constant input of 1. Each of the inputs and the bias are connected to the main neuron by a weight. A weight is generally a real number between 0 and 1. When the input number is fed into the neuron, it is multiplied by the corresponding weight. After this, the weights are all summed up and fed through a hard-limiter. Basically, a hard-limiter is a function that defines the threshold values for 'firing' the neuron. For example, the limiter could be:

$$f(x) = \begin{Bmatrix} x \le 0 \to 0 \\ x \ge 1 \to 1 \end{Bmatrix} \tag{A.1}$$

Actually, both $w_i$and $b_i$are adjustable scalar parameters of the neuron and such parameters can be adjusted so that the network exhibits some desired or interesting behaviours. The way a neuron learns to distinguish patterns is through modifying its weights, the concept of a learning rule must be introduced. In the neuron, the most common

form of learning is by adjusting the weights by the difference between the desired output and the actual output. Mathematically, this can be written:

$$\Delta w_i = x_i \delta \tag{A.2}$$

where δ = *(desired output) - (actual output)*

Transfer function

There are different types of transfer functions. Many transfer functions have been included in Matlab neural network toolbox. The most commonly used functions are log-sigmoid, tan-sigmoid and linear transfer functions.

- Log-sigmoid transfer function

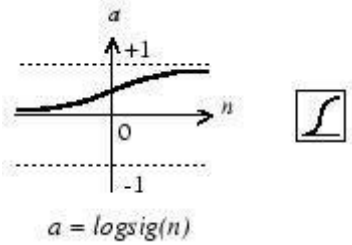


Figure A.2: Log-sigmoid transfer function

$$\text{Math expression:} \quad y = \frac{1-e^{-\alpha}}{1+e^{-\alpha}} \tag{A.3}$$

- Tan-sigmoid transfer function

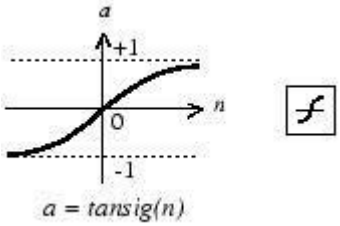


Figure A.3: Tan-sigmoid transfer function

$$\text{Math expression:} \quad y = \frac{1-e^{-\alpha}}{1+e^{-\alpha}} \tag{A.4}$$

- Linear (purelin) transfer function

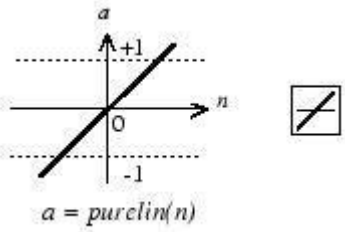


Figure A.4: Linear transfer function

$$\text{Math expression:} \quad y = f(\alpha) = \alpha \tag{A.5}$$

Multilayer networks often use log-sigmoid and tan-sigmoid transfer functions. The sigmoid transfer function squashes the input, which may have any value between plus and minus infinity into the range of 0 and 1. This transfer function is commonly used in back-propagation networks, in part because it is differentiable.

Single-layer feed-forward network

A layered neural network is a network of neurons organised in the form of layers. Figure 3.13 shows the simplest form of a layered network, which has an input layer of source nodes that projects onto an output of neurons but not vice versa. A layer of a network includes the combination of the weights, the multiplication and summing operation (here realised as a vector product W_p), the bias b, and the transfer function f. The array of inputs, vector p, will not be included in or called a layer.

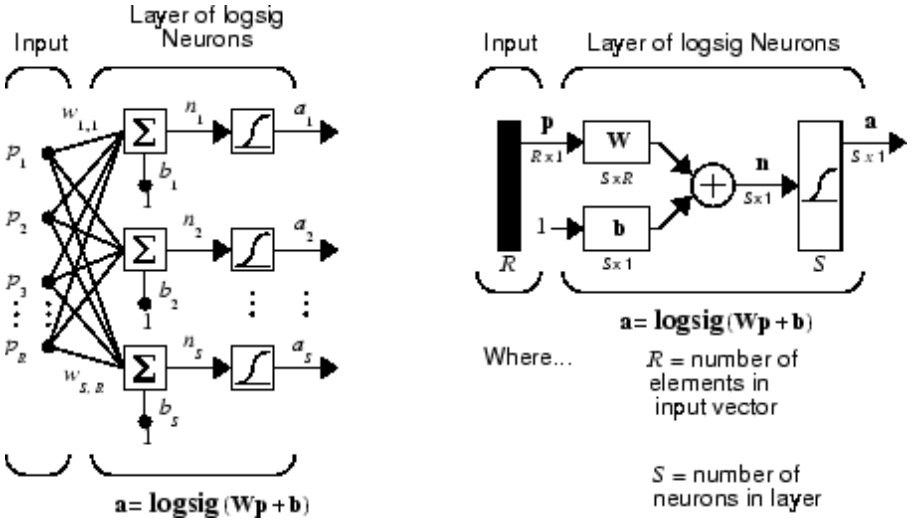


Figure A.5: Single-layer network of S *logsig* neurons

The input vector elements enter the network through the weight matrix W.

$$W = \begin{Bmatrix} w_{1,1}w_{1,2} \dots \dots w_{1,R} \\ w_{2,1}w_{2,2} \dots \dots w_{2,R} \\ \vdots \\ w_{S,1}w_{S,2} \dots \dots w_{S,R} \end{Bmatrix} \quad \text{(A.6)}$$

The row indices on the elements of matrix W indicate the destination neuron of the weight and the column indices indicate which source is the input for that weight.

In other words, this network is strictly of a feed forward type. The designation "single-layer" refers to the output layer of computation nodes. The input layer of source nodes does not count, because no computation is performed there.

A one-layer network with R input elements and S neurons are shown in figure A.5. In this network each element of the input vector p is connected to each neuron

input through the weight matrix W_p. The *ith* neuron has a summer that gathers its weighted inputs and bias to form its own scalar output *n(i)*. The various *n(i)* taken together form an *S*-element net input vector *n*. The sum, *n*, is the argument of the transfer function *f*.

$$n = w_{1,1}p_1 + w_{1,2}p_2 + \cdots + w_{1,R}p_R + b \quad \text{(A.7)}$$

Finally, the neuron layer outputs form a column vector *a*. It is common for the number of inputs to a layer to be different from the number of neurons. A layer is not constrained to have the number of its inputs equal to the number of its neurons.

Multi-layer feed-forward network

Multi-layer network may distinguish itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons or hidden units. The function of hidden neurons is to intervene between the external input and the network output. By adding one or more hidden layers, the network is enabled to extract higher-order statistics and is particularly valuable when the size of the input layer is large.

Each neuron in the hidden layer is connected to a local set of source nodes that lie in its immediate neighbourhood. Likewise, each neuron in the output layer is connected to local set variations of the source signal.

A network can have several layers. Each layer has a weight matrix W, a bias vector *b*, and an output vector *a*. To distinguish between the weight matrices, output vectors and so on, for each of these layers, we will append the number of the layer to the names for each of these variables. For instance, the weight matrix and output vector for the first layer are denoted as *W1* and *A1*, for the second layer these variables are designated as *W2*, *A2* and so on. The network shown above has *R1* inputs, *S1* neurons in the first layer, *S2* neurons in the second layer, etc. It is common for different layers to have different numbers of neurons. A constant input 1 is fed to the biases for each neuron.

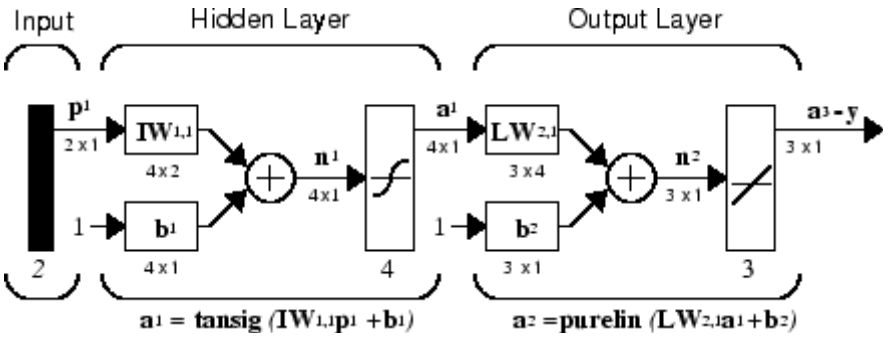


Figure A.6: Two-layer *tansig/purelin* network

The outputs of each intermediate layer are the inputs to the following layer. Thus layer 2 can be analysed as a one-layer network with *S1* inputs, S2 neurons, and an *S1xS2* weight matrix *W2*. The inputs to layer 2 is *a1*, the output is *a2*. All the vectors and matrices of layer 2 can be treated as a single layer network on its own. This approach can be taken with any layer of the network. The layers of a multi-layer network play different roles. A layer that produces the network output is called an output layer. All other layers are called hidden layers (Demuth H., 2001). Multi-layer networks are quite powerful. For instance, a network of three layers is used extensively in back-propagation neural network. Normally, no more than three layers are required in neurons like feed-forward networks, because a three-layer network can generate complex decision regions.

Backpropagation

Backpropagation was created by generalising the Widrow-Hoff learning rule to multi-layer networks and nonlinear differentiable transfer function created backpropagation. Input vectors and the corresponding output vectors are used to train the network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way as defined. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities.

Standard backpropagation is a gradient decent algorithm, as is the Widrow-Hoff learning rule, in which the network weights are moved along the negative of the gradient of the performance function. The term backpropagation refers to the manner

in which the gradient is computed for nonlinear multi-layer networks. There are numbers of variations on the basic algorithm which are based on other standard optimisation techniques, such as conjugate gradient and Newton methods.

The backpropagation neural network is a feed-forward network that usually has hidden layers. The activation function for this type of network is generally the sigmoid function. Since the activation function for these nodes is sigmoid function above, the output from each node is given by:

$$\sigma_i^k = F(a_i^k) \tag{A.8}$$

where a_i is the total input to node *i*, which is given by:

$$\sigma_i^k = \sum_{j=1}^{n} w_{ij} a_j^K + \theta_i \tag{A.9}$$

Weight w_{ij} is the weight of the connection form node *j* to node *i*. Now, as for perception, we will minimise the error in the network by using the gradient descent algorithm to adjust the weights. So the change in the weight from node *j* to node *i* is given by:

$$\Delta_K W_{ij} = -\alpha \frac{\partial E^K}{\partial W_{ij}} \tag{A.10}$$

where E^K is the mean square error for the K_{th} pattern. The error for a hidden node *i* is calculated from the errors of the nodes in the next layer to which node *i* is connected. This is how the error of the network is back propagated (Hasnain S.K.U, 2001).

So, putting them all together, the change for the weight w_{ij}, where node *i* is in a hidden layer, is given by:

$$\begin{aligned} \Delta_k w_{ij} &= \alpha \delta_i^k \sigma_j^k \\ &= \alpha \left(F'\left(a_k^k\right) \sum_{n=1}^{N_{p+1}} \delta_n^k w_{ni} \right) \sigma_j^k \\ &= \alpha \left(\sigma_i^k \left(1 - \sigma_i^k\right) \sum_{n=1}^{N_{p+1}} \delta_n^k w_{ni} \right) \sigma_j^k \end{aligned} \tag{A.11}$$

The changes in the weights of the network, which allow the network to learn, are now totally defined. This generalised delta rule for backpropagation neural networks defines how the weights between the outputs layer and the hidden layer change, and how the weights between other layers change also. This network is called backpropagation because the errors in the network are fed backward, or back propagated, through the network.

Generalisation is perhaps the most useful feature of a backpropagation network. Since the network uses supervised training, a set of input patterns can be organised into groups and fed to the network. The network will "observe" the patterns in each group, and will learn to identify the characteristics that separate the groups. Often, these characteristics are such that a trained network will be able to identify the correct groups, even if the patterns are noisy. The network learns to ignore the irrelevant data in the input patterns.

APPENDIX B. PARTICLE SWARM OPTIMIZATION

Static inertia weight and constriction coefficient

There was a weakness inherent in velocity update equations (5.2) and (5.3) that was fixed by the introduction of an inertia weight. For the following derivation, let $k = 0$ be the iteration at which particles have their positions and, optionally, their velocities randomly initialized. Then for any particle i, the velocity at iteration $k = 1$ is

$$v_i(1) = v_i(0) + c_1 rand_1(0).\left[p_{best_i}(0) - x_i(0)\right] + c_2 rand_2(0).\left[g_{best_i}(0) - x_i(0)\right] \quad \text{(B.1)}$$

Since a particle has only one position, $x_i(0)$, from which to choose in order to determine its personal best, $p_{best_i}(0)$, of necessity $p_{best_i}(0) = x_i(0)$, and the middle term of equation (B.1) is zero, so the particle's velocity at iteration $k = 1$ can more succinctly be expressed as

$$v_i(1) = v_i(0) + c_2 rand_2(0).\left[g_{best_i}(0) - x_i(0)\right] \quad \text{(B.2)}$$

Using (5.2) again, the velocity of particle i at iteration $k = 2$ is

$$v_i(2) = v_i(1) + c_1 rand_1(1).\left[p_{best_i}(1) - x_i(1)\right] + c_2 rand_2(1).\left[g_{best_i}(1) - x_i(1)\right] \quad \text{(B.3)}$$

Substituting the value found in (B.2) for $v_i(1)$, the velocity at the second iteration following initialisation becomes

$$v_i(2) = v_i(0) + c_1 rand_1(1).\left[p_{best_i}(1) - x_i(1)\right] + c_2\left\{rand_2(0).\left[g_{best_i}(0) - x_i(0)\right] + rand_2(1).\left[g_{best_i}(1) - x_i(1)\right]\right. \quad \text{(B.4)}$$

With the substituted values in bold for emphasis. By velocity update equation (5.2), the velocity of particle i at iteration $k = 3$ is

$$v_i(3) = v_i(2) + c_1 rand_1(2).\left[p_{best_i}(2) - x_i(2)\right] + c_2 rand_2(2).\left[g_{best_i}(2) - x_i(2)\right] \quad \text{(B.5)}$$

Substituting for $v_2(1)$ the value found in (B.3), the velocity at the third iteration following initialisation becomes:

$$v_i(3) = v_i(0) + c_1\{rand_{1i}(1)g_{best}[p_{best_i}(1) - x_i(1)] + rand_{1j}(2)[p_{best_i}(2) - x_i(2)]\}$$
$$+ c_2\{rand_{2i}(0)g_{best}[g_{best_i}(0) - x_i(0)]$$
$$+ rand_{2i}(1)g_{best}[g_{best_i}(1) - x_i(1)] + rand_{2j}(2)[g_{best}(2) - x_i(2)]\} \quad \text{(B.6)}$$

By mathematical induction, it can be seen that

$$v_i(k+1) = v_i(0) + c_1 \sum_{a=1}^{k} rand_{1_i}(a).[p_{best_i}(a) - x_i(a)] + c_2 \sum_{a=1}^{k} rand_{2_i}(a).[g_{best_i}(a) - x_i(a)] \quad \text{(B.7)}$$

Because the personal bests and global best can only improve over time, $v_i(k+1)$ should rely more heavily upon recent bests than upon early values. Yet (B.7) shows that early information in $p_{best_i}(a \ll k)$ and $g_{best_i}(a \ll k)$ is given just as much opportunity to affect $v_i(k+1)$ as is the higher quality information of later iterations since the information of all iterations is summed without any weighting scheme by which to increase the relative importance of the higher quality information of later iterations. The problem is remedied by introducing either an inertia weight, $\omega \in (0,1)$.

$$v_i(k+1) = \omega v_i(k) + c_1 rand_{1_i}(k).\left(p_{best_i}(k) - x_i(k)\right) + c_2 rand_{2_i}(k).\left(g_{best_i}(k) - x_i(k)\right) \quad \text{(B.8)}$$

The same process that led to (B.7) beginning with velocity update equation (5.2) leads to (B.9) when beginning with velocity update equation (B.8) with inertia weight.

$$v_i(k+1) = \omega^{k+1} v_i(0) + c_1 \sum_{a=1}^{k} \omega^{k-a}\, rand_{1_i}(a).\left(p_{best_i}(a) - x_i(a)\right) + c_2 \sum_{a=0}^{k} \omega^{k-a}\, rand_{2_i}(a).\left(g_{best_i}(a) - x_i(a)\right) \quad \text{(B.9)}$$

So long as the inertia weight has a magnitude less than one, (B.9) shows that past personal bests are expected to have less effect on a particle's velocity at iteration k+1 than more recent personal bests due to the effect of multiplication at each iteration by the inertia weight, ω. This makes sense conceptually since recent bests - both global and personal - are expected to be higher quality than past bests. However, past bests could still have more effect on a particle's overall velocity than recent bests for a while at the beginning of the search since $g_{best_i}(a) - x_i(a)$, at least, is generally more significant in early iterations when the swarm is more spread out.

Additionally, a particle's initial velocity, which is not derived from any information, but randomly initialised to lie between the upper and lower velocity clamping values, becomes less effective overtime. This too makes sense because its main benefit is in early iterations where it provides momentum by which to propel the best particle, but after some time it effectively becomes noise diluting actual information.

Setting $\omega = 1$ would make velocity update equation (B.8) with inertia weight equivalent to velocity update equation (5.2) without inertia weight so that (B.8) can be accepted without a rigorous proof demonstrating its superiority to (5.2) since it simply provides more options. So long as $\omega \in (0,1)$, velocity update (B.8) helps particles forget their lower quality past positions in order to be more affected by the higher quality information of late, which seems to make more sense conceptually.

Velocity clamping

Eberhart and Kennedy introduced velocity clamping, which helps particles take reasonably sized steps in order to comb through the search space rather than bouncing about excessively (Eberhart *et al.*, 1995). Eberhart showed clamping to improve performance even when parameters are selected according to a simplified constriction model. Consequently, velocity clamping has become a standard feature of PSO.

Velocity clamping is done by first calculating the range of the search space on each dimension, which is done by subtracting the lower bound from the upper

bound. For example, if each dimension of the search space is defined by lower and upper bounds [-100, 100], the range of the search space is 200 per dimension. Velocities are then clamped to a percentage of that range according to

$$v_j^{max} = v_{\max\_perc}.range_j(\Omega)$$

$$v_{\max\_perc} \in (0,1]$$

where

$$range_j(\Omega) = x_j^U - x_j^L$$

for$j = 1,2,\ldots,n$ and search space Ω

$$\Omega = [x_1^L, x_1^U] \times [x_2^L, x_2^U] \times \ldots \times [x_n^L, x_n^U] \subset R^n$$

where x_j^U and x_j^L are, respectively, the lower and upper bounds of the search space along dimension j.

The same maximum velocity should be applied in both the positive and negative directions in order to avoid biasing the search in either the positive or negative direction. The following pseudo code shows how velocities proposed by velocity update equation (B.8) are clamped prior to usage in position update equation (5.3).

If $v_{ij}(k+1) > v_j^{max}$

$v_{ij}(k+1) = v_j^{max}$

else if $v_{ij}(k+1) < -v_j^{max}$

$v_{ij}(k+1) = -v_j^{max}$

end

Figure B.1: Velocity clamping pseudo code

As noted by (Engelbrecht, 2007), clamping a particle's velocity changes not only the step size, but usually also the particle's direction since changing any component of a vector changes that vector's direction unless each component should happen to be reduced by the same percentage. This should not be thought of as a

problem, however, since each dimension is to be optimised independently, and the particle still moves toward the global best on each dimension, though at a less intense speed. Since the maximum iterative movement toward global best on any dimension is clamped, particles may be thought of as combing the search space a bit more thoroughly than were their velocities unclamped.

Though the same velocity clamping percentage of 50% is used in most papers for the sake of comparison, the value does not appear to have been optimised yet. The velocity clamping percentage, v_{max_perc}, is usually chosen within range $0.1 \leq v_{max\_perc} \leq 0.5$.

The lbest Model

For the *lbest* model, a swarm is divided into overlapping neighbourhoods of particles. For each neighbourhood N_i, the best particle is determined, with position g_{best}. This particle is referred to as the *neighbourhood best* particle. Let the indices of the particles wrap around at *s* and the neighbourhood size is *l*. The update equations are:

$$N_i = \{p_{best_{i-l}}(t), p_{best_{i-l+1}}(t), \dots, p_{best_{i-1}}(t), p_{best_i}(t), p_{best_{i+1}}(t), \dots, p_{best_{i+l-1}}(t), p_{best_{i+l}}(t)\} \quad \text{(B.10)}$$

$$g_{best_i}(t+1) \in \left\{N_i | f\left(g_{best_i}(t+1)\right) = min\{f\left(p_{best_i}(t)\right), \forall p_{best_i} \in N_i\}\right\} \quad \text{(B.11)}$$

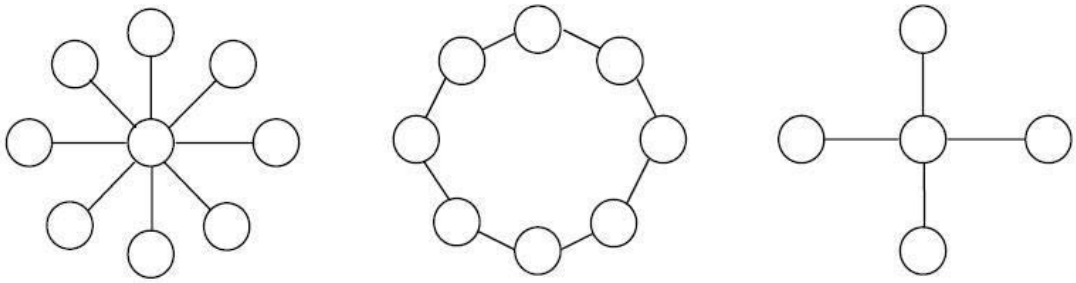
$$v_{i,j}(t+1) = wv_{i,j}(t) + c_1 rand_{1,j}(t)\left(p_{best_{i,j}}(t) - x_{i,j}(t)\right) + c_1 rand_{2,j}(t)\left(g_{best_{i,j}}(t) - x_{i,j}(t)\right) \quad \text{(B.12)}$$

$$x_i(t+1) = x_i(t) + v(t+1) \quad \text{(B.13)}$$

The position update equation is the same as given in equation (B.13). Neighbours represent the social factor in PSO. Neighbourhoods are usually determined using particle indices, however, topological neighbourhoods can also be used (Suganthan, 1999). It is clear that *gbest* algorithm is a special case of *lbest* algorithm with $l = s$; that is, the neighbourhood is the entire swarm. While the *lbest* approach results in a larger diversity, it is still slower than the *gbest* approach.

PSO Neighbourhood topologies

Different neighbourhood topologies have been investigated (Kennedy, 1999; Kennedy *et al.*, 2002). Two common neighbourhood topologies are the *star* (or *wheel*) and *ring* (or *circle*) topologies. For the star topology one particle is selected as a *hub*, which is connected to all other particles in the swarm. However, all the other particles are only connected to the hub. For the ring topology, particles are arranged in a ring. Each particle has some number of particles to its right and left as its neighbourhood. Recently, Kennedy and Mendes [2002] proposed a new PSO model using a *Von Neumann* topology. For the Von Neumann topology, particles are connected using a grid network (2-dimensional lattice) where each particle is connected to its four neighbour particles (above, below, right and left particles). Figure B.1 illustrates the different neighbourhood topologies.



(a) Star Topology (b) Ring Topology (c) Von Neumann Topology

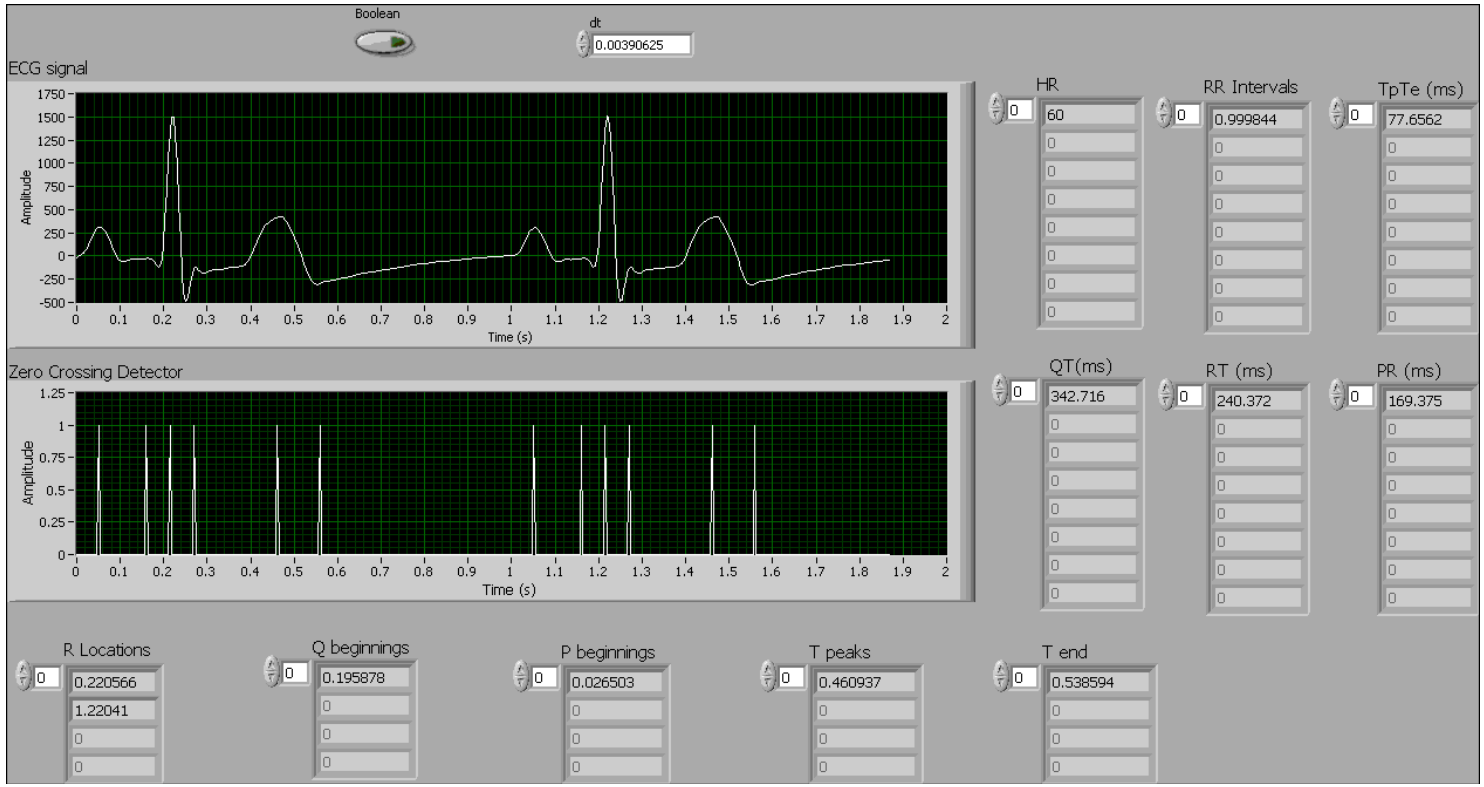
Figure B.2: A diagrammatic representation of neighborhood topologies

The choice of neighbourhood topology has a profound effect on the propagation of the best solution found by the swarm. Using the *gbest* model the propagation is very fast (i.e. all the particles in the swarm will be affected by the best solution found in iteration t, immediately in iteration $t+1$). This fast propagation may result in the premature convergence problem. However, using the ring and Von Neumann topologies will slow down the convergence rate because the best solution found has to propagate through several neighbourhoods before affecting all particles in the swarm. This slow propagation will enable the particles to explore more areas in the search space and thus decreases the chance of premature convergence.

APPENDIX C. LABVIEW PROGRAM

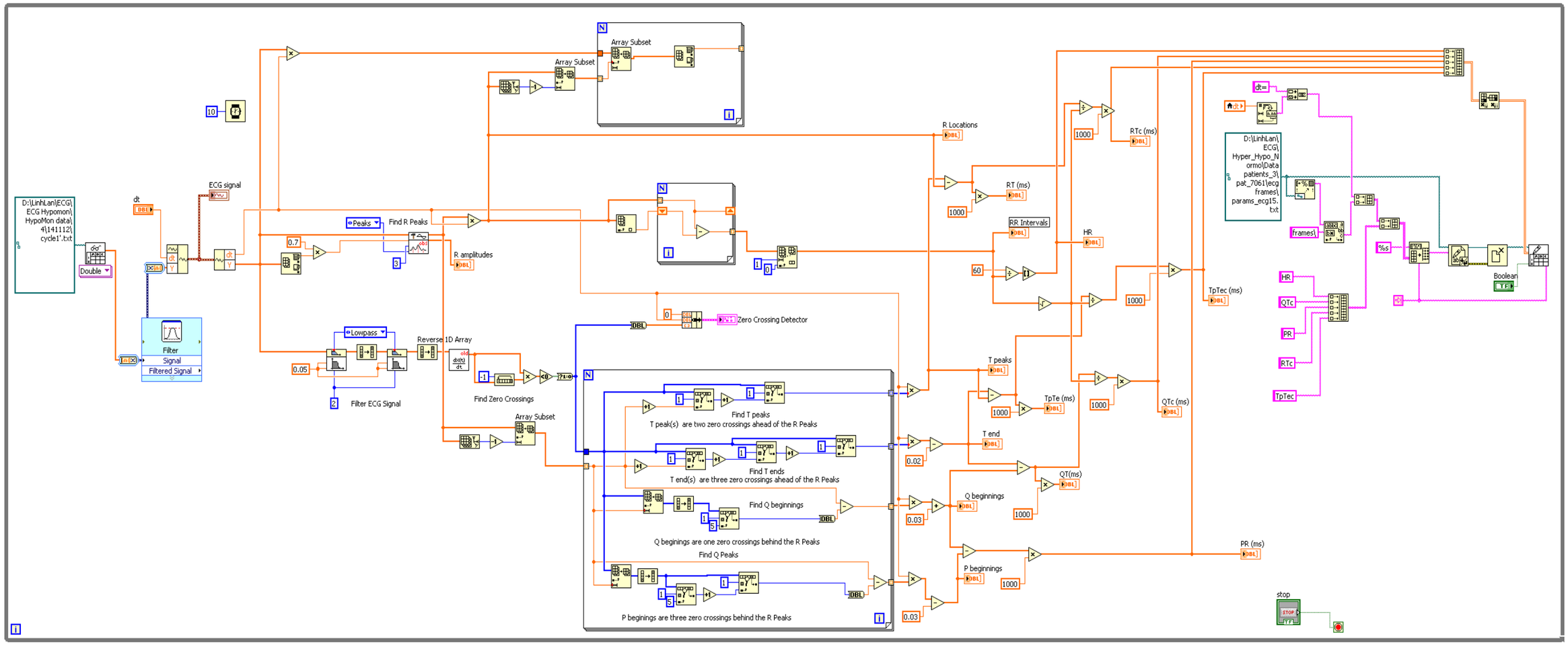
ECG_PQRT.vi

Front panel



ECG_PQRT.vi

Block diagram



APPENDIX D. MATLAB CODE

```
%--- Neural network----
close all, clear all , clc
xx=importdata('16 variables_nml.txt');
t=importdata('9 variables_target_1opnode.txt');
[I N] = size(xx);
[O N] =size(t);

% Neural Net Model
Neq = N*O; % No. of equations;
%Nw = (I+1)*H+(H+1)*O No of unknown weights
Hub = floor((N-1)*O/(I+O+1)); % =1 Maximun H for Neq >= Nw
Ntrials = 50;

% In case want to duplicate the run
rng(1234,'twister');

for j = 1:10 %Validation set stopping allows H > Hub
    H = 8
    Nw = (I+1)*H+(H+1)*O;
    for i = 1:Ntrials
        net = newpr(xx,t,H);
        net.trainFcn='trainpso';
        net.layers{1}.transferFcn = 'tansig'; %sigmoidal transer function for hidden
        layers
        net.layers{2}.transferFcn = 'purelin'; %linear transfer function for output
        layer
        net.trainParam.goal = 0.001; % equivalent to PSO true global minimum.
        net.trainParam.epochs = 1000; % equivalent to number of PSO iterations.
```

```
net.trainParam.max_fail = 10; % this is the maximum number of consecutive
PSO iterations where no improved performance is found.


% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 35/100;
net.divideParam.valRatio = 35/100;
net.divideParam.testRatio = 30/100;

% Train the Network
[net tr Y E] = train(net,xx,t);

layer_weights=net.lw{2,1}
bias2=net.b{2};
Nepochs(i) = tr.epoch(end);
MSE = tr.perf(end);
R2(i,j) = 1-MSE;
Nerr = numel(find(round(Y)~=t));
PctErr(i,j) = 100*Nerr/N;

e=t-Y;
mse_1=mse(e);
rmse=sqrt(mse_1);
mae_1=mae(e);

%Confusion matrices
trainX=xx(:,tr.trainInd);
trainT=t(:,tr.trainInd);
trainY=sim(net,trainX);
trainIndices=vec2ind(trainY);
[cTrain,cmTrain]=confusion(trainT,trainY);
```

```
senTrain=cmTrain(1,1)/(cmTrain(1,1)+cmTrain(1,2));
specTrain=cmTrain(2,2)/(cmTrain(2,1)+cmTrain(2,2));

valX=xx(:,tr.valInd);
valT=t(:,tr.valInd);
valY=sim(net,valX);
valIndices=vec2ind(valY);
[cVal,cmVal]=confusion(valT,valY);
senVal=cmVal(1,1)/(cmVal(1,1)+cmVal(1,2));
specVal=cmVal(2,2)/(cmVal(2,1)+cmVal(2,2));

testX=xx(:,tr.testInd);
testT=t(:,tr.testInd);
testY=sim(net,testX);
testIndices=vec2ind(testY);
[cTest,cmTest]=confusion(testT,testY);
senTest=cmTest(1,1)/(cmTest(1,1)+cmTest(1,2));
specTest=cmTest(2,2)/(cmTest(2,1)+cmTest(2,2));

%Kappa coefficient
observed_agreement=(cmTest(1,1)+cmTest(2,2))/(cmTest(1,1)+cmTest(1,2)+
cmTest(2,1)+cmTest(2,2));
actualPos=(cmTest(1,1)+cmTest(1,2))/(cmTest(1,1)+cmTest(1,2)+cmTest(2,
1)+cmTest(2,2));
actualNeg=(cmTest(2,1)+cmTest(2,2))/(cmTest(1,1)+cmTest(1,2)+cmTest(2,
1)+cmTest(2,2));
predictedPos=(cmTest(1,1)+cmTest(2,1))/(cmTest(1,1)+cmTest(1,2)+cmTest
(2,1)+cmTest(2,2));
predictedNeg=(cmTest(1,2)+cmTest(2,2))/(cmTest(1,1)+cmTest(1,2)+cmTes
t(2,1)+cmTest(2,2));
chance_agreement=actualPos*predictedPos+actualNeg*predictedNeg;
kappa=(observed_agreement-chance_agreement)/(1-chance_agreement);
```

```
  if senTrain >= .5 && specTrain >= .5 && senTest >= .5 && specTest >= .5 &&
kappa >= .5
        fprintf(1, 'ith= %3.0f \n',i)
        fprintf(1, 'senTrain= %3.4f \n',senTrain)
        fprintf(1, 'specTrain= %3.4f \n',specTrain)
        fprintf(1, 'cmTrain= %3.4f \n',cmTrain)
        fprintf(1, 'senVal= %3.4f \n',senVal)
        fprintf(1, 'specVal= %3.4f \n',specVal)
        fprintf(1, 'cmVal= %3.4f \n',cmVal)
        fprintf(1, 'senTest= %3.4f \n',senTest)
        fprintf(1, 'specTest= %3.4f \n',specTest)
        fprintf(1, 'cmTest= %3.4f \n',cmTest)
        fprintf(1, 'mae= %3.4f \n',mae_1)
        fprintf(1, 'mse= %3.4f \n',mse_1)
        fprintf(1, 'rmse= %3.4f \n',rmse)
        fprintf(1, 'kappa= %3.4f \n',kappa)
        disp(' ');
    end

  %ROC
    [tpr,fpr,thresholds]=roc(trainT,trainY);
    tpr
    fpr
    auc=trapz(fpr,tpr)
    [tpr,fpr,thresholds] = roc(testT,testY)
    combTrValT=[trainT,valT];
    combTrValY=[trainY,valY];
    figure(i)
    plotroc(t,Y)
```

```
    xlabel('Sensitivity');
    ylabel('1 - Specificity');
    title('ROC plot');
end
end

Tabulations
format short
Nepochs = Nepochs
R2 = R2
format bank
PctErr = PctErr



%---Particle swarm optimisation---
%Control_Panel.m
%   Neural Network training using PSO settings based on Tricia Rambharose's  %
modification of existing neural network training algorithms,
%   provided in Matlab's Neural Network toolbox, using ideas from Brian
% Birge's PSO toolbox @

OnOff_Tricias_NN_training = logical(1);
if ~OnOff_Tricias_NN_training
   clc
   close all
   clear all
   OnOff_Tricias_NN_training = logical(0);
end

%BASIC SWITCHES & PSO ALGORITHM SELECTION : ON (1), OFF (0)%
OnOff_Autosave_Workspace_Per_Grouping = logical(1);
```

```
OnOff_Autodelete_Grouping_Data = logical(1);
OnOff_Autosave_Workspace_Per_Trial = logical(1);
OnOff_Autodelete_Trial_Data = logical(1);
OnOff_Autosave_Workspace_Per_Column = logical(1);
OnOff_Autosave_Workspace_Per_Table = logical(1);
   %TERMINATION CRITERIA FOR REGPSO & PSO ALGORITHMS%

OnOff_NormR_stag_det = logical(0);

OnOff_func_evals = logical(0);
   %REGPSO & PSO SUCCESS MEASURES%
   OnOff_SuccessfulUnsuccessful = logical(1); %
if OnOff_SuccessfulUnsuccessful
   OnOff_iter_success = logical(1);
   OnOff_Terminate_Upon_Success = logical(0); %
end
   %PSO ALGORITHM SELECTION%

OnOff_lbest = logical(1);
if OnOff_lbest
   lbest_neighb_size = 2;
end
OnOff_w_linear = logical(1);
OnOff_v_clamp = logical(1);
OnOff_v_reset = logical(0);
OnOff_position_clamping = logical(0);
OnOff_RegPSO = logical(0);

if OnOff_RegPSO
   Reg_Method = 1;
else
```

```
    Reg_Method = 0;
end
OnOff_GCPSO = logical(0);
if OnOff_GCPSO
   GCPSO_threshold_num_successes = 15;
   GCPSO_threshold_num_failures = 5;
end
OnOff_MPSO = logical(0);
if OnOff_MPSO
    OnOff_MPSO_zhist = logical(1);
    MPSO_max_starts = 2;
    if OnOff_func_evals
        MPSO_max_FEs = 2000000;
    else
        MPSO_max_iters = 100000;
    end
end
OnOff_OPSO = logical(0);
if OnOff_OPSO
    p0 = 0.5;
end
OnOff_Cauchy_mutation_of_global_best = logical(0);

    %MISCELLANEOUS FEATURES%
OnOff_user_input_validation_required = logical(0);
OnOff_progress_meter = logical(1);
OnOff_asymmetric_initialization = logical(0);

    %PSO HISTORIES TO BE MAINTAINED%
OnOff_fghist = logical(0);
OnOff_ghist = logical(1);
```

```
if OnOff_lbest
    OnOff_lhist = logical(1);
end
OnOff_fphist = logical(0);
OnOff_phist = logical(1);
OnOff_fhist = logical(1);
OnOff_xhist = logical(0);
OnOff_vhist = logical(1);
OnOff_v_cog_hist = logical(0);
OnOff_v_soc_hist = logical(0);

%PARTICLE SWARM ALGORITHM SETTINGS%
if OnOff_Tricias_NN_training
    num_trials = 1;
else
num_trials = 1;
end
vmax_perc = .2;
global dim
if OnOff_Tricias_NN_training
    dim = length(NN_wb);
else
    dim = 2;
end
np = 30;
c1 = 2;
c2 = 2;
w = 0.72984;
if OnOff_func_evals
    max_FEs_per_grouping = 200; %
    if OnOff_w_linear
```

```
        max_iter_per_grouping = max_FEs_per_grouping/np;
    end
else
    max_iter_per_grouping =  35;
end
format short e;

%PSO GRAPHING SWITCHES AND SETTINGS%
OnOff_graphs = logical(1);
if OnOff_graphs
    OnOff_graph_fg_mean = logical(1);
    OnOff_swarm_trajectory = logical(0);
    OnOff_graph_ObjFun_f_vs_2D = logical(1);
    OnOff_phase_plot = logical(1);
    OnOff_graph_fg = logical(0);
    OnOff_graph_g = logical(1); %Graph History of the Global Best
    OnOff_graph_p = logical(0);
    OnOff_graph_f = logical(0);
    OnOff_graph_x = logical(1);
    OnOff_graph_v = logical(0);
    OnOff_title = logical(1);
    OnOff_semilogy = logical(0);
    OnOff_reuse_figures = logical(0);
    OnOff_autosave_figs = logical(1);
    OnOff_autosave_figures_to_another_format = logical(1);
    if OnOff_autosave_figures_to_another_format
        GraphParams_autosave_format = 'png';
    end
    OnOff_Close_All_Graphs = logical(0);

    Figure_Position = [1 76 1920 1058];
```

```
    GraphParams_labeltitle_font_size = 40;
    GraphParams_axes_font_size = 35;
    GraphParams_marker_size = 40;
    GraphParams_text_font_size = 30;
    if OnOff_swarm_trajectory || OnOff_graph_ObjFun_f_vs_2D || OnOff_phase_plot
        GraphParams_meshgrid_quality = 450;
        if OnOff_swarm_trajectory
            OnOff_contourf = logical(1);
            OnOff_zoom_on_final_graph = logical(1);
            GraphParams_snapshot_per_how_many_iterations = 10;
            GraphParams_line_width = 2;
            OnOff_plot_indices = logical(1); %Plot particles' index numbers.
            OnOff_plot_box_4_new_search_space = logical(1);
            if OnOff_plot_box_4_new_search_space
                GraphParams_fraction_of_line_width = 0.0543;
            end
            OnOff_mark_personal_bests = logical(1);
            OnOff_mark_global_best_always = logical(1);
            OnOff_mark_global_best_on_zoomed_graph = logical(1);
            GraphParams_swarm_traj_snapshot_mode = 4;
            if GraphParams_swarm_traj_snapshot_mode == 4
                GraphParams_SwarmTrajMode4Factor = 1.6;
            end %suggested values: 1.6 - 1.8
        end
    end
end

%trainpso.m
function [net, tr] = trainpso(net, tr, trainV, valV, testV, varargin)
```

```
global first_pso

first_pso = true ;
if strcmp(net,'info')
  info.function = mfilename;
  info.title = 'Particle Swarm Optimization NN training';
  info.type = 'Training';
  info.version = 6;
  info.training_mode = 'Supervised';
  info.gradient_mode = 'No Gradient';
  info.uses_validation = false;
  info.param_defaults.show = 1000;
  info.param_defaults.epochs = 10;
  info.param_defaults.time = inf;
  info.param_defaults.goal = 0.3;
  info.param_defaults.max_fail = 6;
  info.param_defaults.max_perf_inc = 1.04;
  info.param_defaults.showCommandLine = true;
  info.param_defaults.showWindow = false;
  info.param_defaults.plotPSO = true;
  net = info;
  return
end

if ischar(net)
  switch (net)
    case 'name', info = feval(mfilename,'info'); net = info.title;
    case 'pnames', info = feval(mfilename,'info'); net =
fieldnames(info.param_defaults);
    case 'pdefaults', info = feval(mfilename,'info'); net = info.param_defaults;
    case 'gdefaults', if (tr==0), net='calcgrad'; else net='calcgbtt'; end
```

```
    otherwise, error('NNET:Arguments','Unrecognised code.')
  end
  return
end

epochs = net.trainParam.epochs;
goal = net.trainParam.goal;
max_fail = net.trainParam.max_fail;
max_perf_inc = net.trainParam.max_perf_inc;
show = net.trainParam.show;
time = net.trainParam.time;

if (~isa(epochs,'double')) || (~isreal(epochs)) || (any(size(epochs)) ~= 1) || ...
  (epochs < 1) || (round(epochs) ~= epochs)
  error('NNET:Arguments','Epochs is not a positive integer.')
end
if (~isa(goal,'double')) || (~isreal(goal)) || (any(size(goal)) ~= 1) || ...
  (goal < 0)
  error('NNET:Arguments','Goal is not zero or a positive real value.')
end
if (~isa(max_fail,'double')) || (~isreal(max_fail)) || (any(size(max_fail)) ~= 1) || ...
  (max_fail < 1) || (round(max_fail) ~= max_fail)
  error('NNET:Arguments','Max_fail is not a positive integer.')
end
if (~isa(max_perf_inc,'double')) || (~isreal(max_perf_inc)) || (any(size(max_perf_inc))
~= 1) || ...
  (max_perf_inc < 1)
  error('NNET:Arguments','Max_perf_inc is not a positive real value greater or equal
to 1.0.')
end
if (~isa(show,'double')) || (~isreal(show)) || (any(size(show)) ~= 1) || ...
```

```
  (isfinite(show) && ((show < 1) || (round(show) ~= show)))
  error('NNET:Arguments','Show is not "NaN" or a positive integer.')
end
if (~isa(time,'double')) || (~isreal(time)) || (any(size(time)) ~= 1) || ...
  (time < 0)
  error('NNET:Arguments','Time is not zero or a positive real value.')
end

Q = trainV.Q;
TS = trainV.TS;
startTime = clock;
NN_wb = getx(net);
original_net = net;

tr.states = {'epoch','time','perf'};
if net.trainParam.plotPSO
    fig = figure;
end

Control_Panel;

NN_wb = g(1, :);
net = setx(net, NN_wb);
[perf,El,trainV.Y,Ac,N,Zb,Zi,Zl] =
calcperf2(net,NN_wb,trainV.Pd,trainV.Tl,trainV.Ai,Q,TS);
tr = tr_clip(tr);
end
```

REFERENCES

ADA, AMERICAN DIABETES ASSOCIATION, 2004. Hyperglycemic Crises in Diabetes. *Diabetes Care* **27**(suppl 1): s94-s102.

AKIRAV, EITAN M.*, et al.*, 2011. Detection of β cell death in diabetes using differentially methylated circulating DNA. *Proceedings of the National Academy of Sciences* **108**(47): 19018-19023.

AKSNES, T. A.*, et al.*, 2008. Impact of new-onset diabetes mellitus on development of atrial fibrillation and heart failure in high-risk hypertension (from the VALUE Trial). *Am J Cardiol* **101**(5): 634-638.

ALDHOUS, P, 1992. Race quickens for non-stick blood monitoring technology. *Science* **258**(5084): 892-893.

ALEXAKIS, C.*, et al.*, 2006. A knowledge-based electrocardiogram-monitoring system for detection of the onset of nocturnal hypoglycaemia in type 1 diabetic patients. *Computers in Cardiology.*

ALTMAN, D G and BLAND, J M, 1994. Statistics Notes: Diagnostic tests 1: sensitivity and specificity. *BMJ* **308**(6943): 1552.

ASSOCIATION, AMERICAN DIABETES, 2012. Standards of Medical Care in Diabetes—2012. *Diabetes Care* **35**(Supplement 1): S11-S63.

ASTLE, S. M., 2005. Restoring eletrolyte balance. *Rn* **68**(5): 34-39.

BALAKRISHNAMA S., GANAPATHIRAJU A., 1998. Linear discriminant analysis, a brief tutorial. Institute for singal and information processing, Mississippi State University.

BARRON, H. V. and LESH, M. D., 1996. Autonomic nervous system and sudden cardiac death. *J Am Coll Cardiol* **27**(5): 1053-1060.

BATHULA, R.*, et al.*, 2010. Indian Asians have poorer cardiovascular autonomic function than Europeans: this is due to greater hyperglycaemia and may contribute to their greater risk of heart disease. *Diabetologia* **53**(10): 2120-2128.

BEN-HAROUSH, A.*, et al.*, 2004. Epidemiology of gestational diabetes mellitus and its association with Type 2 diabetes. *Diabet Med* **21**(2): 103-113.

BHARMI, RUPINDER, "System and method for detecting hypoglycemia based on a paced depolarization integral using an implantable medical device", US Patent, US7590443 B2, granted 15 Sep 2009.

BIGGER, J. T., JR.*, et al.*, 1992. Frequency domain measures of heart period variability and mortality after myocardial infarction. *Circulation* **85**(1): 164-171.

BIRGE, B., 2003. PSOt - a particle swarm optimization toolbox for use with Matlab. *Swarm Intelligence Symposium, 2003. SIS '03. Proceedings of the 2003 IEEE*,

BIRGE, BRIAN, 2005. Paritcle Swarm Optimization Toolbox. Matlab Central. Available from: http://www.mathworks.com.au/matlabcentral/fileexchange/7506-particle-swarm-optimization-toolbox.

BISHOP, D. K., *et al.*, 2010. A disposable tear glucose biosensor-part 1: design and concept testing. *J Diabetes Sci Technol* **4**(2): 299-306.

BOERINGER, D. W. and WERNER, D. H., 2004. Particle swarm optimization versus genetic algorithms for phased array synthesis. *Antennas and Propagation, IEEE Transactions on* **52**(3): 771-779.

BONOW, R.O., *et al.*, 2011. In: Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine, 2-Volume Set. Elsevier Health Sciences.

BREALEY, D. and SINGER, M., 2009. Hyperglycemia in critical illness: a review. *J Diabetes Sci Technol* **3**(6): 1250-1260.

BROWNLEE, M., 2005. In: The pathobiology of diabetic complications: a unifying mechanism. *Diabetes*. **54**(6):1615-25.

BURGE, MARK R., *et al.*, 2008. Continuous Glucose Monitoring: The Future of Diabetes Management. *Diabetes Spectrum* **21**(2): 112-119.

CAPES, S. E., *et al.*, 2000. Stress hyperglycaemia and increased risk of death after myocardial infarction in patients with and without diabetes: a systematic overview. *Lancet* **355**(9206): 773-778.

CARNETHON, MERCEDES R., *et al.*, 2003. Prospective Investigation of Autonomic Nervous System Function and the Development of Type 2 Diabetes: The Atherosclerosis Risk In Communities Study, 1987–1998. *Circulation* **107**(17): 2190-2195.

CH, SUDHEER and MATHUR, SHASHI, 2012. Particle swarm optimization trained neural network for aquifer parameter estimation. *KSCE Journal of Civil Engineering* **16**(3): 298-307.

CHAN, N. N. and HUREL, S. J., 2002. Potential impact of a new blood glucose monitoring device: the GlucoWatch® Biographer. *Practical Diabetes International* **19**(4): 97-100.

CHEMLA, D., *et al.*, 2005. Comparison of fast Fourier transform and autoregressive spectral analysis for the study of heart rate variability in diabetic patients. *Int J Cardiol* **104**(3): 307-313.

CHRISTENSEN, T. F., *et al.*, 2010. QT interval prolongation during spontaneous episodes of hypoglycaemia in type 1 diabetes: the impact of heart rate correction. *Diabetologia* **53**(9): 2036-2041.

COMMITTEE, EXPERT, 2003. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care* **26**(1): S5-20.

CRYER, P. E., 2004. Diverse causes of hypoglycemia-associated autonomic failure in diabetes. *N Engl J Med* **350**(22): 2272-2279.

DASKALAKI, E., *et al.*, 2012. Real-time adaptive models for the personalized prediction of glycemic profile in type 1 diabetes patients. *Diabetes Technol Ther* **14**(2): 168-174.

DEKKER, J. M., *et al.*, 1996. QTc duration is associated with levels of insulin and glucose intolerance. The Zutphen Elderly Study. *Diabetes* **45**(3): 376-380.

DEMUTH H., BEALE M., 2001. Neural network Toolbox for use with Matlab. Available from: http://au.mathworks.com/products/neural-network/

EBERHART, R. and KENNEDY, J., 1995. A new optimiser using particle swarm theory. *Micro Machine and Human Science, 1995. MHS '95., Proceedings of the Sixth International Symposium on*, pp. 39-43.

ECKERT, B. and AGARDH, C. D., 1998. Hypoglycaemia leads to an increased QT interval in normal men. *Clin Physiol* **18**(6): 570-575.

ELECTROPHYSIOLOGY, TASK FORCE OF THE EUROPEAN SOCIETY OF CARDIOLOGY THE NORTH AMERICAN SOCIETY OF PACING, 1996. Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use. *Circulation* **93**(5): 1043-1065.

ENGELBRECHT, A.P., 2007. In: Computational Intelligence: An Introduction. 2 ed., Wiley.

EVANS, RONALD WILLIAM, "Systems and methods for detecing symtoms of hypoglycemia", US Patent, US7052472 B1, granted May 30, 2006.

EVERS, G. I. and BEN GHALIA, M., 2009. Regrouping particle swarm optimization: A new global optimization algorithm with improved performance consistency across benchmarks. *Systems, Man and Cybernetics. SMC 2009. IEEE International Conference on*,

FERRANTE DO AMARAL, CARLOS EDUARDO and WOLF, BENHARD, 2008. Current development in non-invasive glucose monitoring. *Medical Engineering & Physics* **30**(5): 541-549.

FREIRE, A. X., *et al.*, 2005. Admission hyperglycemia and other risk factors as predictors of hospital mortality in a medical ICU population. *Chest* **128**(5): 3109-3116.

GALASSETTI, PIETRO, 2007. Breath analysis offers potential for non-invasive blood sugar monitoring in diabetes.*http://today.uci.edu/iframe.php?p=/news/release_detail_iframe.asp?key=1667*.

GANDHI, G. Y., *et al.*, 2005. Intraoperative hyperglycemia and perioperative outcomes in cardiac surgery patients. *Mayo Clin Proc* **80**(7): 862-866.

GAO, LIANG, *et al.*, 2006. Combining Particle Swarm Optimization and Neural Network for Diagnosis of Unexplained Syncope. Computational Intelligence and Bioinformatics. Huang, D.-S., K. Li and G. Irwin, Springer Berlin Heidelberg. **4115:** 174-181.

GARSON, A., JR., 1993. How to measure the QT interval--what is normal? *Am J Cardiol* **72**(6): 14B-16B.

GORDIN, D.*, et al.*, 2008. Acute hyperglycaemia disturbs cardiac repolarization in Type 1 diabetes. *Diabet Med* **25**(1): 101-105.

GOWDAR, D., 2010. In: Automatic Detection of QT and Related Heart Beat Intervals. VDM Verlag Publisher, Germany.

GROSS, T. M.*, et al.*, 2000. Performance evaluation of the MiniMed continuous glucose monitoring system during patient home use. *Diabetes Technol Ther* **2**(1): 49-56.

HASNAIN S.K.U, ASIM S.M., 2001. Artificial Neural Networks in Cardiology - ECG Wave Analysis and Diagnosis Using Backpropagation Neural Networks. *IEEE Engineering in Medicine and Biology Society Conference*, Melbourne, Australia.

HIRSHBERG, E.*, et al.*, 2008. Alterations in glucose homeostasis in the pediatric intensive care unit: Hyperglycemia and glucose variability are associated with increased mortality and morbidity. *Pediatr Crit Care Med* **9**(4): 361-366.

HOMOUD, M. K., 2009. ACP Journal Club: Prolonged PR intervals were associated with increased risk for atrial fibrillation, pacemaker implantation, and mortality. *Ann Intern Med* **151**(10): 0003-4819.

HUYNH, DUY C. and DUNNIGAN, MATTHEW W., 2012. Advanced particle swarm optimisation algorithms for parameter estimation of a single–phase induction machine. *International Journal of Modelling, Identification and Control* **15**(4): 227-240.

IRELAND, R. H.*, et al.*, 2000. Measurement of high resolution ECG QT interval during controlled euglycaemia and hypoglycaemia. *Physiol Meas* **21**(2): 295-303.

JAISWAL, M.*, et al.*, 2013. Reduced heart rate variability among youth with type 1 diabetes: the SEARCH CVD study. *Diabetes Care* **36**(1): 157-162.

JOHNSON, KIMBALL, 2012. Diabetes and Continuous Glucose Monitoring.WebMD. Available: *http://diabetes.webmd.com/continuous-glucose-monitoring*.

JOLLIFFE, I.T., 2002. In: Principal Component Analysis. Springer.

JOSLIN, E.P. and KAHN, C.R., 2005. In: Joslin's Diabetes Mellitus: Edited by C. Ronald Kahn ... [et Al.]. Lippincott Williams & Wilkins.

KARABAG, T.*, et al.*, 2011. Prolonged P wave dispersion in pre-diabetic patients. *Kardiol Pol* **69**(6): 566-571.

KATO, T.*, et al.*, 2006. What are arrhythmogenic substrates in diabetic rat atria? *J Cardiovasc Electrophysiol* **17**(8): 890-894.

KAUFMAN, F. R.*, et al.*, 2001. A pilot study of the continuous glucose monitoring system: clinical decisions and glycemic control after its use in pediatric type 1 diabetic subjects. *Diabetes Care* **24**(12): 2030-2034.

KEENAN, D. B., *et al.*, 2009. Delays in minimally invasive continuous glucose monitoring devices: a review of current technology. *J Diabetes Sci Technol* **3**(5): 1207-1214.

KENNEDY, J., 1999. Small worlds and mega-minds: effects of neighborhood topology on particle swarm performance. *Evolutionary Computation. CEC 99. Proceedings of the 1999 Congress on*, vol. 3, pp. 1931-1938.

KENNEDY, J. and MENDES, R., 2002. Population structure and particle swarm performance. *Evolutionary Computation. CEC '02. Proceedings of the 2002 Congress on*, vol. 2, pp. 1671-1676.

KLONOFF, DAVID C, 1997. Noninvasive Blood Glucose Monitoring. *Diabetes Care* **20**(3): 433-437.

KUDAT, H., *et al.*, 2006. Heart rate variability in diabetes patients. *J Int Med Res* **34**(3): 291-296.

LE COMPTE, A. J., *et al.*, 2010. Blood glucose prediction using stochastic modeling in neonatal intensive care. *IEEE Trans Biomed Eng* **57**(3): 509-518.

LEE, S., *et al.*, 2003. Effects of adrenaline and potassium on QTc interval and QT dispersion in man. *Eur J Clin Invest* **33**(2): 93-98.

LEE, Y. S., *et al.*, 2008. Bounded PSO Vmax Function in Neural Network Learning. *Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on*, vol. 1, pp. 474-479.

LI-PING, ZHANG, *et al.*, 2005. Optimal choice of parameters for particle swarm optimization. *Journal of Zhejiang University Science A* **6**(6): 528-534.

LIN, J. N., *et al.*, 2009. Risk factors for mortality of bacteremic patients in the emergency department. *Acad Emerg Med* **16**(8): 749-755.

LING, S. S. and NGUYEN, H. T., 2011. Genetic-algorithm-based multiple regression with fuzzy inference system for detection of nocturnal hypoglycemic episodes. *IEEE Trans Inf Technol Biomed* **15**(2): 308-315.

LORSHEYD, A., *et al.*, 2005. PR and OTc interval prolongation on the electrocardiogram after binge drinking in healthy individuals. *Neth J Med* **63**(2): 59-63.

MALIN, STEPHEN F., *et al.*, 1999. Noninvasive Prediction of Glucose by Near-Infrared Diffuse Reflectance Spectroscopy. *Clinical Chemistry* **45**(9): 1651-1658.

MALLIANI, A., *et al.*, 1991. Cardiovascular neural regulation explored in the frequency domain. *Circulation* **84**(2): 482-492.

MALULI, HAYAN AL and MESHKOV, ARNOLD B., 2013. A short story of the short QT syndrome. *Cleveland Clinic Journal of Medicine* **80**(1): 41-47.

MARDIA, KANTI, *et al.*, 1995. In: Multivariate Analysis (Probability and Mathematical Statistics). Academic Press.

MARFELLA, R.*, et al.*, 2000. The effect of acute hyperglycaemia on QTc duration in healthy man. *Diabetologia* **43**(5): 571-575.

MARFELLA, R.*, et al.*, 2001. Hyperglycemia and QT interval: time for re-evaluation. *Diabetes Nutr Metab.***14**(2): 63-5.

MCMILLIN, J. MICHAEL, 1990. Blood Glucose. In: Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd. Boston, Butterworths. Chapter 141.

MOWERY, N. T.*, et al.*, 2009. Stress insulin resistance is a marker for mortality in traumatic brain injury. *J Trauma* **66**(1): 145-151.

MÜLLER, B.*, et al.*, 1995. In: Neural Networks: An Introduction. Springer Berlin Heidenberg.

NGUYEN, D. and WIDROW, B., 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *Neural Networks, IJCNN International Joint Conference on*, vol 3., pp. 21-26.

NGUYEN, H. T.*, et al.*, 2006. Neural-Network Detection of Hypoglycemic Episodes in Children with Type 1 Diabetes using Physiological Parameters. *Engineering in Medicine and Biology Society. EMBS '06. 28th Annual International Conference of the IEEE*, pp. 6053-6056.

NGUYEN, H. T., NEJHDEH GHEVONDIAN, "Non-invasive method and apparatus for determining onset of physiological conditions", US Patent, US 7,450,986 B2, granted 11 November 2008.

NGUYEN, H. T.*, et al.*, 2009. Real-time detection of nocturnal hypoglycemic episodes using a novel non-invasive hypoglycemia monitor. *Conf Proc IEEE Eng Med Biol Soc* **5**(10): 5335144.

NGUYEN, L. L.*, et al.*, 2012. Identification of hypoglycemia and hyperglycemia in type 1 diabetic patients using ECG parameters. *Conf Proc IEEE Eng Med Biol Soc* **9**(10): 2716-2719.

NGUYEN, L.*, et al.*, 2013. Effects of hyperglycemia on variability of RR, QT and corrected QT intervals in Type 1 diabetic patients. *Conf Proc IEEE Eng Med Biol Soc* **22**(10): 1819-1822.

NISHIZAKI, M.*, et al.*, 2002. Effects of glucose-induced insulin secretion on ventricular repolarization in patients with congenital long QT syndrome. *Circ J* **66**(1): 35-40.

NORITAKE, M.*, et al.*, 1992. Diurnal change in heart rate variability in healthy and diabetic subjects. *Intern Med* **31**(4): 453-456.

NOVAK, B. J.*, et al.*, 2007. Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes. *Proceedings of the National Academy of Sciences* **104**(40): 15613-15618.

NURYANI, N.*, et al.*, 2012. Electrocardiographic signals and swarm-based support vector machine for hypoglycemia detection. *Ann Biomed Eng* **40**(4): 934-945.

PAGANI, M. and MALLIANI, A., 2000. Interpreting oscillations of muscle sympathetic nerve activity and heart rate variability. *J Hypertens* **18**(12): 1709-1719.

PAPPADA, S. M., *et al.*, 2010. Development of a neural network model for predicting glucose levels in a surgical critical care setting. *Patient Saf Surg* **4**(1): 1754-9493.

PARVIZ, BABAK, 2009. Augmented Reality in a Contact Lens: A new generation of contact lenses built with very small circuits and LEDs promises bionic eyesight. IEEE Spectrum, Biomedical. http://spectrum.ieee.org/biomedical/bionics/augmented-reality-in-a-contact-lens.

PRAKASH, A. and MATTA, B. F., 2008. Hyperglycaemia and neurological injury. *Curr Opin Anaesthesiol* **21**(5): 565-569.

RAMBHAROSE, TRICIA, 2011. Neural Network add-in for PSORT. Matlab Central. Available from: http://www.mathworks.com.au/matlabcentral/fileexchange/29565-neural-network-add-in-for-psort.

RAMCHANDANI, N. and HEPTULLA, R. A., 2012. New technologies for diabetes: a review of the present and the future. *Int J Pediatr Endocrinol* **26**(1): 1687-9856.

RAO, G., *et al.*, 1993. Reverse iontophoresis: development of a noninvasive approach for glucose monitoring. *Pharm Res* **10**(12): 1751-1755.

REBRIN, K., *et al.*, 1999. Subcutaneous glucose predicts plasma glucose independent of insulin: implications for continuous monitoring. *Am J Physiol* **277**(3 Pt 1): E561-571.

ROBINSON, R. T., *et al.*, 2004. Changes in cardiac repolarization during clinical episodes of nocturnal hypoglycaemia in adults with Type 1 diabetes. *Diabetologia* **47**(2): 312-315.

SALVI, V., *et al.*, 2011. Comparison of 5 methods of QT interval measurements on electrocardiograms from a thorough QT/QTc study: effect on assay sensitivity and categorical outliers. *J Electrocardiol* **44**(2): 96-104.

SANDLER, B.P., 1951. In: Diet Prevents Polio. Lee Foundation for Nutritional Research. Milwakuee, WI.

SHARMA, S. and MUKHERJEE, S., 1996. In: Applied Multivariate Techniques. John Wiley & Sons Canada, Limited.

SHIN, S. J., *et al.*, 1989. Assessment of autonomic regulation of heart rate variability by the method of complex demodulation. *IEEE Trans Biomed Eng* **36**(2): 274-283.

SIEG, ANKE, *et al.*, 2004. Noninvasive Glucose Monitoring by Reverse Iontophoresis in Vivo: Application of the Internal Standard Concept. *Clinical Chemistry* **50**(8): 1383-1390.

SINGH, J. P., *et al.*, 2000. Association of hyperglycemia with reduced heart rate variability (The Framingham Heart Study). *Am J Cardiol* **86**(3): 309-312.

SPARACINO, G., *et al.*, 2008. Continuous glucose monitoring time series and hypo/hyperglycemia prevention: requirements, methods, open problems. *Curr Diabetes Rev* **4**(3): 181-192.

SUGANTHAN, P. N., 1999. Particle swarm optimiser with neighbourhood operator. *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*, vol. 3, pp. 1958-1962.

SUYS, B., *et al.*, 2006. Glycemia and corrected QT interval prolongation in young type 1 diabetic patients: what is the relation? *Diabetes Care* **29**(2): 427-429.

TAMADA, J. A., *et al.*, 1995. Measurement of glucose in diabetic subjects using noninvasive transdermal extraction. *Nat Med* **1**(11): 1198-1201.

TAMADA, J. A., *et al.*, 1999. Noninvasive glucose monitoring: comprehensive clinical results. Cygnus Research Team. *Jama* **282**(19): 1839-1844.

TARVAINEN, M. P., *et al.*, 2013. Complexity of heart rate variability in type 2 diabetes - effect of hyperglycemia. *Conf Proc IEEE Eng Med Biol Soc* **61**(10): 5558-5561.

TATTERSALL, R. B. and GILL, G. V., 1991. Unexplained deaths of type 1 diabetic patients. *Diabet Med* **8**(1): 49-58.

TETKO, IGOR V., *et al.*, 1995. Neural network studies. 1. Comparison of overfitting and overtraining. *Journal of Chemical Information and Computer Sciences* **35**(5): 826-833.

TZAMALOUKAS, A. H. and AVASTHI, P. S., 1987. Temporal Profile of Serum Potassium Concentration in Nondiabetic and Diabetic Outpatients on Chronic Dialysis. *American Journal of Nephrology* **7**(2): 101-109.

TZAMALOUKAS, A. H., *et al.*, 2011. Abnormalities of serum potassium concentration in dialysis-associated hyperglycemia and their correction with insulin: review of published reports. *Int Urol Nephrol* **43**(2): 451-459.

VASHIST, SANDEEP KUMAR, 2012. Non-invasive glucose monitoring technology in diabetes management: A review. *Analytica Chimica Acta* **750**(0): 16-27.

WENSLEY, FRANCES, 2013. Heart Rate Variability (HRV) & Diabetes. *ithlete Blog*. Available: http://myithlete.com/blog/?p=1297#.U4QKbXYWfF8.

WIKIPEDIA, 2013. Hyperglycemia.*http://en.wikipedia.org/wiki/Hyperglycemia*.

YAGHOUBY, F., AYATOLLAHI, A.,SOLEIMANI, R., 2009. Classification of Cardiac Abnormalities Using Reduced Features of Heart Rate Variability Signal. *World Applied Sciences Journal* **6**(11): 1547-1554.

YAO, H., *et al.*, 2011. A contact lens with embedded sensor for monitoring tear glucose level. *Biosens Bioelectron* **26**(7): 3290-3296.

YAZICI, M.*, et al.*, 2007. The effect of diabetes mellitus on the P-wave dispersion. *Circ J* **71**(6): 880-883.

YUHUI, SHI and EBERHART, R., 1998. A modified particle swarm optimiser. *Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence.,* pp. 69-73.

ZHANG, J.*, et al.*, 2011. Noninvasive diagnostic devices for diabetes through measuring tear glucose. *J Diabetes Sci Technol* **5**(1): 166-172.

ZHANG, Y.*, et al.*, 2003. Impairment of human ether-a-go-go-related gene (HERG) K+ channel function by hypoglycemia and hyperglycemia. Similar phenotypes but different mechanisms. *J Biol Chem* **278**(12): 10417-10426.