

第一章 緒 論

1.1 前言

根據行政院衛生署統計資料[1]顯示，心臟疾病在民國 88 至 90 三個年度分居台灣地區十大死亡原因的第四及第三位，如表 1。國人的生活飲食自從 80 年代中期以後脫離開發中國家的經濟型態以來，心臟血管疾病就一直高居不下，這與國人飲食習慣和生活作息由農業社會轉為工商社會，壓力增加且缺乏適量的運動有關。因此早期發現心臟疾病並能即早治療，已是當務之急了。

表 1 臺灣地區歷年每十萬人口死亡率按主要死亡原因別分

年別 (民國)	80	81	82	83	84	85	86	87	88	89	90
每十萬人口 死亡率(人)	58.79	62.91	59.99	56.93	52.93	52.58	49.71	50.51	51.33	47.56	49.25
十大死因排 名	3	3	3	2	4	4	4	3	4	3	3

資料來源：行政院衛生署衛生統計資訊網

心電圖(Electrocardiogram, ECG)為診斷心臟疾病重要的輔助工具之一，主要是它具有快速、非侵入性及易取得的優點。心電圖是心臟收縮時所產生的心臟電氣活動的一種記錄，利用電極置於身體特定部位，依特定組合所產生觀察心臟的十二個不同的角度[2]。

1.2 研究背景

心電圖的量測具有量測簡易、非侵入性 (non-invasive)、含豐富資訊及判讀很多優點，因此心電圖學在心臟專科中是一項不可或缺的

必備知識。過去一直有很多學者專家對心電圖自動診斷系統從事研究，此系統可概分為特徵值的擷取與辨識兩部分[3]：

1. 特徵值擷取：

特徵值擷取可分為時域（Time Domain）與頻域（Frequency Domain）兩類：

- (1).時域特徵值擷取，是直接藉由心電圖形來擷取其特徵值。這包括了取用波形的寬度、高度、面積、斜率、區間、區段、及比值等，以提供辨識時使用。
- (2).頻域特徵值擷取則是透過數學轉換，如快速傅立葉轉換(FFT)、小波轉換(Wavelet)，將時域心電圖波形轉換到頻域，再找出頻譜特徵值，以供病症辨識時使用。

2.波形辨識與分類：

波形辨識與分類方法有決策邏輯、圖形識別、決策規則、多變量分析及類神經網路等。多變量分析具有運算簡單，且有一定的分類能力特點，而類神經網路具有學習能力、容錯性和大量平行運算之優點，故都常被用來做波形辨識與分類。

過去有很多學者提出各種特徵值擷取及辨識方法，如 1985 年，Tompkins 等學者[4]，將 MIT/BHI 資料庫之心電圖經帶通濾波器濾波，在時域上直接取心電圖特徵值，辨識率可達 99.3% 以上。1990 年，童淑芬 [5]從 MIT/BHI 資料庫取心電圖信號，經帶通濾波器濾波，再經快速傅立葉轉換(FFT)，利用倒傳遞類神經網路(backpropagation neural network)，對 6 種病症進行分類，辨識率可達 92.5% 以上。1994 年林俊榮[6]，利用自組織映射圖 SOM(self organizing map)類神經網路模式，做病症辨識，在時域直接由心電圖取特徵值，做 5 種病症辨識，辨識率達 98% 以上。1999 年，Minami 等學者[7]，

將 QRS 波利用快速傅立葉轉換，將信號由時域轉成頻率，再取 5 個頻譜成份 (spectral component) 加至類神經網路，做三種病症辨識，辨識率達 98% 以上。2001 年，Osowski 等學者[8]，用模糊混合類神經網路(Fuzzy Hybrid Neural Network)做心電圖辨識，對 7 種病症做分類其辨識率 96.06% 以上。2002 年楊朝樑[3]使用主要元素分析神經網路做心電圖病症辨識，辨識率達 98% 以上。

由前面之文獻可瞭解，不論是頻域或時域擷取特徵值都已經有專家與學者做出蠻好的成果。本研究嘗試將 ECG 信號經由小波轉換取其特徵值後，利用最近鄰法、區別分析、倒傳遞網路及徑向基網路等四種方式做分類與辨識，期望能有更佳成果。

1.3 研究目的

本研究以小波轉換、多變量分析及類神經網路為基礎，發展一個具準確性、即時性、簡單又快速之心電圖病症自動辨識系統，提供醫療人員判斷之參考，以減輕其工作負擔、提高醫療品質。

1.4 全文架構

本論文共分九章，第一章緒論，內容包括前言、研究背景、研究目的及全文結構。第二章心電圖原理，介紹心跳與電的關係、心臟結構、心臟的傳導系統及心臟十二導程心電圖。第三章小波轉換，介紹小波轉換緣起、小波發展歷史及小波分析基本理論。第四章分類法介紹歐氏距離、區別分析、倒傳遞網路及徑向基網路。第五章實驗設計，包括實驗流程圖、實驗設備、心電圖資料庫、心電圖樣本之取得、特徵值之產生及評估方式。第六章研究結果，包括歐氏距離法辨識結果、區別分析法辨識結果、倒傳遞網路法辨識結果及徑向基網路

法辨識結果。第七章討論不同的分類方式對辨識程度之影響及結果比較、各種分類結果之特性分析，以及與相關文獻研究之比較。第八章未來方向是對本研究提出改進方向及建議。

第二章 心電圖原理

2.1 心跳與電的關係

要了解心電圖產生之原理，首先要了解心跳與電的關係。驅使心臟動作的原動力就是自發生物性電力，而心電圖就是記錄心臟收縮與舒張之間，心臟電氣生理反應的一項記錄。它是偵測心臟各類病變最簡便而安全之方法。醫師根據心電圖，可以告訴病人各類心律不整 (Arrhythmias)、心肌梗塞 (Myocardial Infarctions)、心房或心室之肥大 (Atrial Enlargement or Ventricular Hypertrophy)、電軸偏轉 (Axis Deviation)、電解質不平衡 (Electrolyte Disturbances)、心肌病 (cardiomyopathy)、心肌炎(myocarditis)、心包炎(pericarditis)、特殊心臟藥物 (Druge) 造成之變化以及其他心臟疾病。

2.2 心臟結構

心臟是一個肌肉組成而中空的器官，負責將含有氧氣及養料的血液泵到身體各器官。心臟位於人體胸部中央，傾向左方，介於第二條至第六條左肋骨之間，被胸骨和肋骨包圍著。心臟長度約 10cm，重量約 270gm。心臟外表有一層膜囊，稱為心包囊(Pericardial sac)。這囊內含有少量水一般的流體，把心浸泡著作保護功能，如圖 2.1 所示。

雖然心臟能獨立地工作，但亦有兩套神經系統支配它的工作。這就是，交感神經(Sympathetic nerves)刺激心臟，而副交感神經則令心臟鎮定下來。

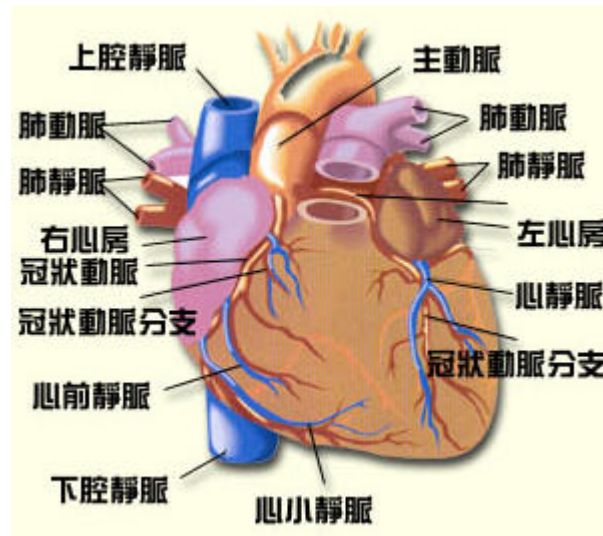


圖 2.1 心臟的結構

心臟也有自己的重要血液供應，即冠狀循環 (Coronary circulation)。兩條主要的冠狀動脈 (Coronary arteries) 恰好在主動脈離開心臟之處由主動脈分支而成。而它們自己亦有無數的分支，保證心臟有足夠的血液及氧氣供應。

心壁有三層組織，如圖 2.2：

- 1、心包 (Pericardium)：在心外的一個透明薄膜。
- 2、心內膜 (Endocardium)：與心包相似，同屬一透明薄膜，貼於心腔內壁。
- 3、心肌層 (Myocardium)：是一層厚厚的肌肉，它位於心包及心內膜之間。心肌層是心臟特有的一種肌肉，主管心臟的收縮。

心臟的心室心房系統如圖 2.3 所示，分為上下左右四個心室，上方的二個稱為心房 (Atria)，負責接收從右側靜脈 (Veins) 帶回來的血液。下方的二個稱為心室 (Ventricles)，負責泵出血液至身體其他部份。

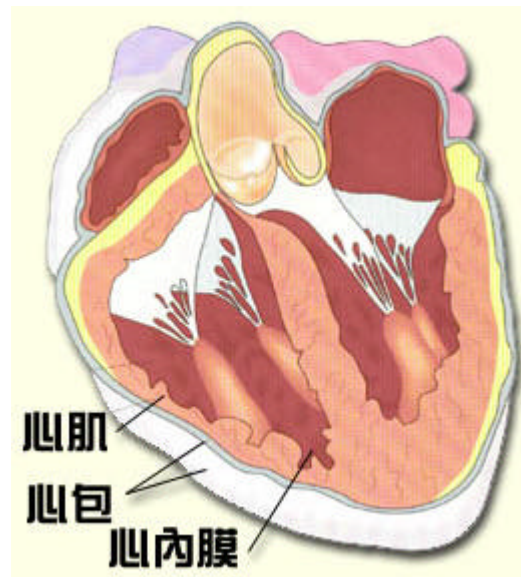


圖 2.2 心壁構造

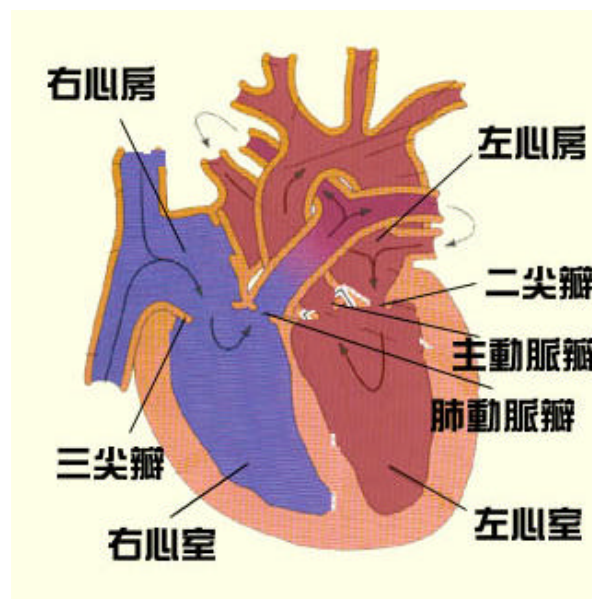


圖 2.3 心臟的心室心房

心臟內均有四個心瓣(Valves)，使血液只能往一個方向流動。心膈(Septum)則將心臟分為左右兩邊，右邊心臟負責接收全身的血液，然後輸送到肺部；左邊心臟負責接收由肺部的帶氧血液，然後輸送到身體各部份。

血液流動的途徑：右心房(Right atria) 三尖瓣(Tricuspid valve)
右心室(Right ventricle) 肺動脈瓣(Pulmonary valve) 肺動脈
(Pulmonary artery) 肺部 左心房 二尖瓣 左心室 主動脈瓣
主動脈 身體各部份。而每分鐘心臟會重複上述循環動作大約 70 次
左右，這就是脈搏。

2.3 心臟的傳導系統

心臟本身電的傳導活動記錄是可以測量的，而測量方法就是把電極安置於身體胸部不同部位，接收由心跳時所產生的微小電壓變化，通常記錄於紙帶上，並且這樣的過程並不會引起身體不適。心電圖是利用圖形描繪心臟跳動時電位之變化，心臟之正常與否可由心電圖予以判別。它可記錄心臟肌肉收縮前後電脈波的情況，並檢查心臟運作和功能[9]。

當靜脈的血液由上下腔靜脈流入右心房時，同時四條肺靜脈的血液也流入左心房，心房血液注滿時，自主神經系統(Autonomic nervous system)的激動(Impulse)，可刺激竇房結(Sino-atrial node)，送往心房的心肌，引致心房衝動，再刺激房室結(Atrio-ventricular node, AV node)，之後傳至希氏束(Bundle of His)，再傳到兩心室的肌肉而引起心室收縮(Ventricular systole)，如圖 2.4。

心臟就像一所自動泵，受心臟傳導系統的影響，規律地作週期運動，這規律運動稱為『心臟搏動週期(Cardiac cycle)』。心臟搏動每分鐘約七十二次，每週期(即每搏動一次)約 0.8 秒。

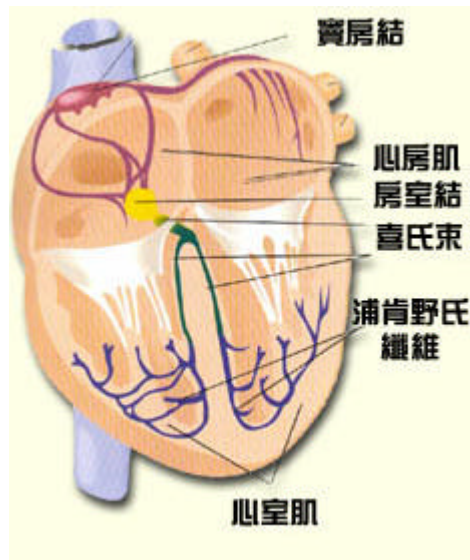


圖 2.4 心臟的傳導系統

心臟搏動週期包括：

1. 心房收縮 (Atrial Systole) 產生 P 波 (P wave) , 如圖 2.5 所示 , 約 0.1 秒。
2. 心室收縮 (Ventricular Systole) 產生 QRS 波組 (QRS complex) , 如圖 2.6 所示 , 約 0.3 秒。
3. 心室舒張 (Complete cardiac diastole) 產生 T 波 (T wave) , 如圖 2.7 所示 , 約 0.4 秒

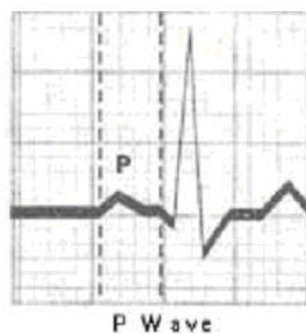


圖 2.5 心房收縮

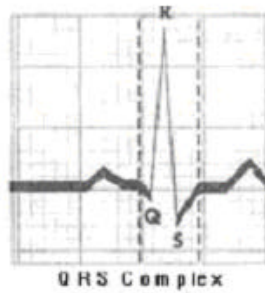


圖 2.6 心室收縮

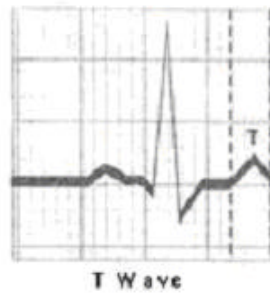


圖 2.7 心室舒張

心電圖對個別疾病診斷之靈敏度(sensitivity)及特異性(specificity)有極大之差異，因此醫師可能配合其他之檢查來作整體之判斷。

正常心電圖(圖 2.8)是有規律地，能清楚看到心房收縮(P)、心室收縮(QRS)及心室舒張(T)。

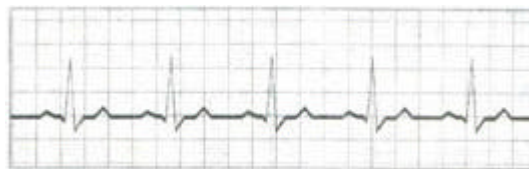


圖 2.8 正常心電圖

2.4 十二導程心電圖

要了解十二導程心電圖之前，首先得介紹心電圖學上的一些定

義：

1、標準肢導極定義：

如圖 2.9 所示為標準肢導極定義[10]的量測位置。

- (1). 第一肢導極(lead I)：左手為正極，右手為負極所構成，它的方位角是 0 度。
- (2). 第二肢導極(lead II)：左腳為正極，右手為負極所構成，它的方位角是 60 度。
- (3). 第三肢導極(lead III)：左腳為正極，左手為負極所構成，它的方位角是 120 度。

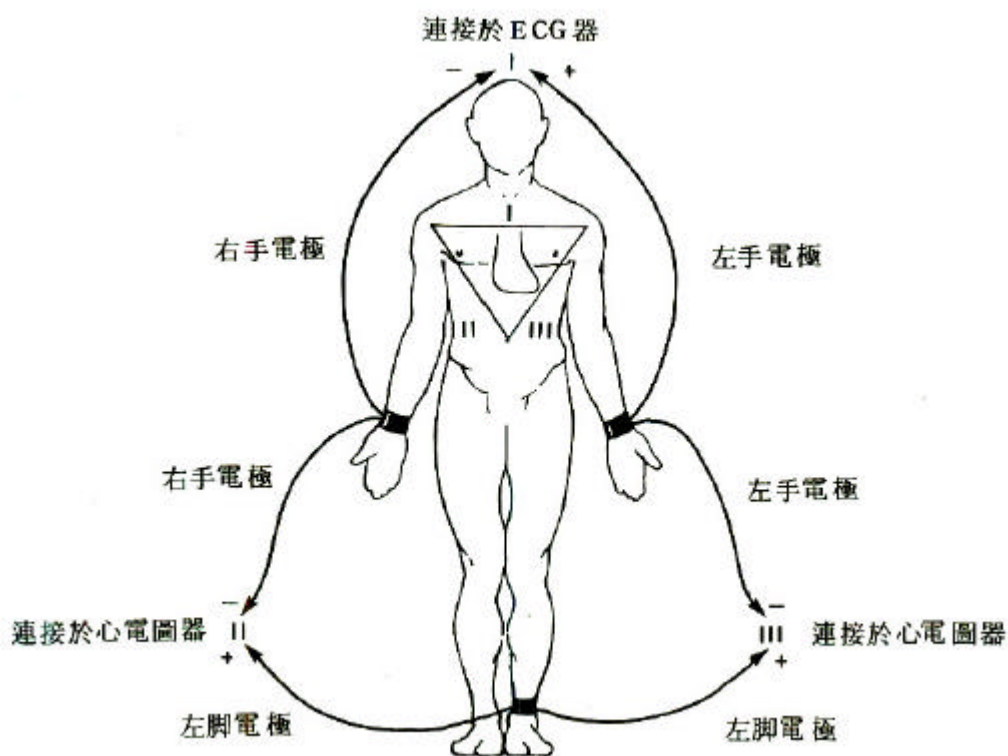


圖 2.9 標準肢導極量測位置

2.三個增大導極(augmented leads)定義[11]：

- (1). AVL 導極(Lead AVL)：量測位置如圖 2.10，左手為正極，而其他電極為負極所構成的，它的方位角是-30 度。

- (2). AVR 導極(Lead AVR , AVR)：量測位置如圖 2.11，右手為正極，而其他電極為負極所構成的，它的方位角是-150 度。
- (3). AVF 導極(Lead AVF , AVF)：量測位置如圖 2.12，左腳為正極，而其他電極為負極所構成的，它的方位角是 90 度。

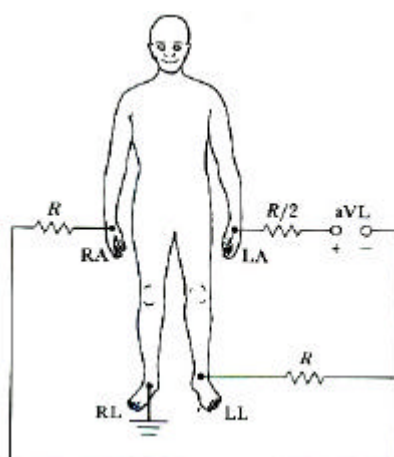


圖 2.10 AVL 導極量測位置

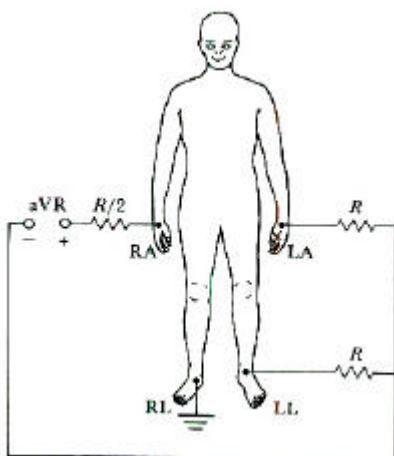


圖 2.11 AVR 導極量測位置

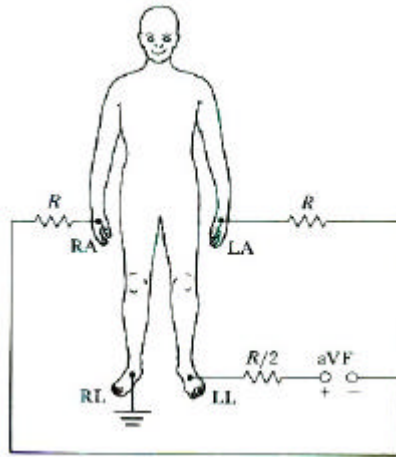


圖 2.12 aVF 導極量測位置

3.六個胸導極(chest lead) 定義：

量測位置如圖 2.13 所示，V1 至 V6 的六個電極位置分別如下：

- (1). V1 是位於胸骨右側的第四肋間。
- (2). V2 是位於胸骨左側的第四肋間。
- (3). V3 是位於 V2 和 V4 之間。
- (4). V4 是位於第五肋間的左鎖骨中線之處。
- (5). V5 是位於 V4 和 V6 之間。
- (6). V6 是位於第五肋間的腋中線(midaxillary line)之處。

圖 2.14 是完整的標準十二導極心電圖波形及其相對位置圖。

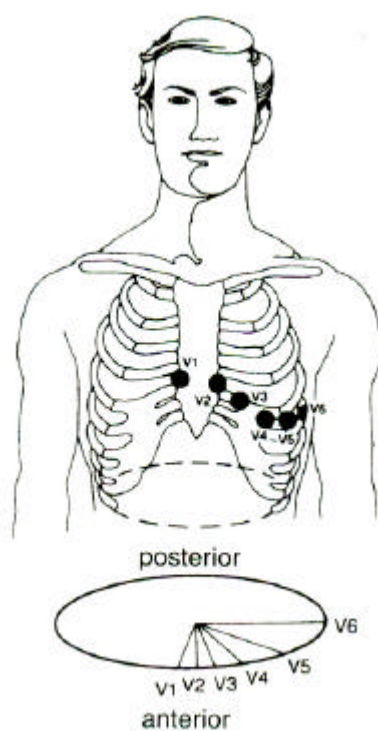


圖 2.13 六個胸導極(chest lead)量測位置

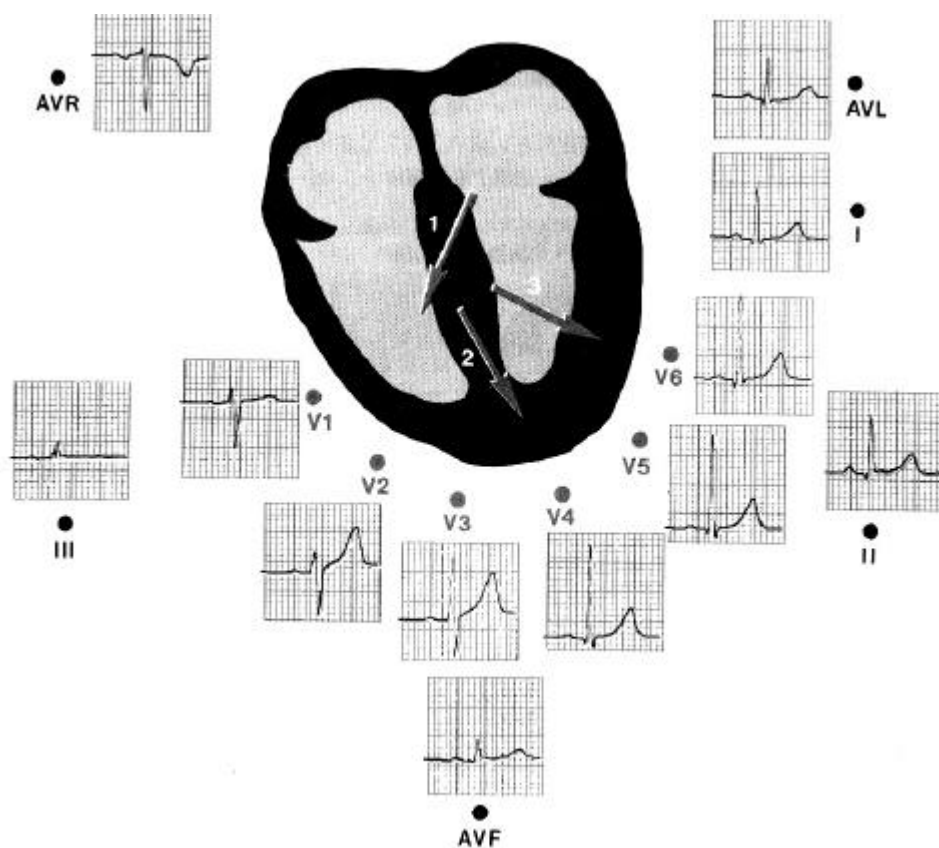


圖 2.14 標準十二導極心電圖波形及其相對位置圖

第三章 小波轉換介紹

本研究論文是以小波轉換為基礎，藉其分解後的信號建立特徵值，以便進行運算及分類。本章就小波轉換作簡單介紹。

3.1 數位訊號處理簡介

數位訊號處理 (Digital System Processing , DSP) 是一門新穎的學問，因為它所要處理的數位訊號本身就是一種新的發明，它是近四十年由於電子計算機所產生新的資料形態。我們生活四周已經充斥數位訊號的存在，不但在電腦和網路中看到、聽到的多媒體資訊，全都是數位訊號，從照相機的控制系統、汽車的電子噴射引擎、數位型大哥大，到現在已在推廣的高傳真電視，也都有數位訊號。雖然如此，並不代表我們已經能完全掌握數位訊號的特質。就好像汽車已經在陸地跑了一百多年[12]，但是我們至今還是不能完全掌握它的絕對安全性。同樣地數位訊號還有許多尚未開發的特性值得去研究，而且由於它的廣泛應用及價值，更使得這一類的研究與開發，更形急迫。

數位訊號處理的工具，可以概略分為數學工具與電腦工具兩種。前者則通常需要較多的數學知識，後者利用資料結構或演算法。以訊號壓縮為例，DCT (離散餘弦轉換：Discrete Cosine Transform) 和 DWT (離散小波轉換：Discrete Wavelet Transform) 都是數學工具，ADPCM (可適性差分脈衝碼調變) 和 LZW (多階層資料壓縮) 都是電腦工具。DCT 是傅立葉轉換的應用，而 DWT 是小波函數的應用。例如 JPEG (靜態影像壓縮標準：Joint Photographic Experts Group) 圖形格式應用 DCT 演算法，可以達到平均 1:10 以上的壓縮效果。但是，使用 DWT 的演算法，在同樣的視覺表現下，可以將壓縮比再提升一倍以上。ADPCM 可以將十六位元的聲音訊號做 1:4 的壓

縮，GIF（圖形交換格式）圖形格式應用 LZW 演算法，可以將二五六色的點陣圖做平均 1:5 的壓縮。電腦工具的成效，通常比較受限制。目前已經有一些 DWT 類的演算法走出了實驗室，而製成產品軟體與硬體都有。

3.2 小波轉換（Wavelet Transform）緣起

約在 1986 年，一群在調和分析領域的數學家，發現了小波函數。剛開始時，只有屬於數學、物理和地震波的研究者，注意到這個新發現。而小波函數和 DSP 之間相互的關聯也隨之被發現，很快地，其他領域的專家、科學家與工程師們，都加入了這個發展與研究的行列。也因為這些跨領域學者的共同開發，小波函數從一開始就是理論與應用並行，而且發展快速。

廣義來說，小波函數就是那些在時間領域與頻率領域都具有局部性的函數。依據海森堡測不準定理可知，任何函數的時間反應與頻率反應不可能同時都很狹窄。而從應用的角度來說，就是任何濾波器，無法同時在時間領域和頻率領域上都具有高準確度。小波函數也像其它函數一樣受到這種理論的限制。然而小波函數卻容許在此限制下盡可能地設計出高效率的濾波器。

小波轉換與數位信號處理在 Daubechies 及 Mallat 結合下 [13]，已有所改變。發展至今，小波轉換不論在數學、電腦科學、物理或工程上皆有舉足輕重的地位。小波轉換對於各種訊號分析已提供了完整的架構，本研究也將它應用在信號處理上。

3.3 小波發展歷史

就本研究建構的分析工具所採用的小波理論基礎，依年代先後編

排文獻如下[14]：

- 1807年：Joseph Fourier 提出頻域分析理論(Fourier synthesis)。
- 1910年：Haar 發展出Haar function。
- 1938年：Littlewood-Paley 對Fourier series 建立L-P 理論。
- 1970年代：Hardy space theory 的發展。
- 1982年：Strömberg 提出第一個正交小波，其特性為具有良好的時頻局部定位性，同時也證明小波的存在。
- 1984年：Morlet 將小波的概念引入訊號分析之中，並與Grossmann 定義連續小波轉換，也正式將之命名為「wavelet」。
- 1985年：Meyer 建立Meyer wavelet，具正交、易衰減等特性的平滑小波，其傅利葉轉換為緊湊支撐(compact support)，且屬於連續函數空間。
- 1985年：Tchamitchian 提出雙正交小波。
- 1986年：Battle、Lemarié提出Battle-Lemarié wavelet，為將B-spline 正交後所得，對時頻域的支撐為infinity，具有指數衰減的效應。
- 1987年：Mallat 證實可透過多重解析度分析來產生小波基底。
- 1988年：Daubechies 提出Daubechies wavelet，為有限支撐，且具有良好的時頻局部定位性。
- 1989年：Coifman 提出一組包含vanishing moment 的正規化尺度函數及小波函數建造方式。
- 1991年：Wickerhauser 提出小波包(wavelet packet)算法。
- 1993年：Goodman 提出多重小波(multi-wavelet)的概念，並與Lee、Tang 等拓展Mallat 的正交小波，而得出多重小波。
- 1996年：Chui-Lian 研究orthonormal 多重小波。

1998年：Jiang 提出時頻解析度最佳化的正交多重小波。

3.4 小波分析基本理論

小波分析基本理論在於以數學上的函數轉換的方式，將資料分成幾個頻率分量。若傳統的頻譜分析，是以傅利葉轉換(Fourier transform)為數學上的理論依據。

函數 $f(t)$ 的傅利葉轉換定義為：

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt \quad \text{式(3.1)}$$

上式 3.1 是表示將時間函數 $f(t)$ 經過基底函數 $e^{-j\omega t}$ 投影至各頻率的分量。以尤拉公式(Euler Formula)展開 $e^{-j\omega t}$ 可得：

$$e^{-j\omega t} = \cos(\omega t) - j \sin(\omega t)$$

所以 $F(\omega)$ 表示的是 $f(t)$ 在整個頻域上的分佈量。而且式 3.1 的時間分佈包函正負無窮大，因此傅利葉轉換缺乏局部時間的解析度。

如果訊號的性質並未隨時間變化而發生很大變動，則稱這種訊號屬於固定性(stationary)的訊號。然而大多數的情形下，訊號中屬於固定性的分量並非真正希望量測到的，反而是非固定性的分量才是要研究的部分，例如一些不定時產生的突波(spike)，或是不規則的漂移。這個時候傅利葉分析在應用上就顯得不夠充分。那是因為若以傅利葉分析轉換至頻域，將會失去時域之相關訊息。

D. Gabor 為了修正這個缺點，提出將訊號切割成許多小段，逐段作傅利葉分析以得到較多時域上的訊息，這就是所謂的短時間傅利葉轉換(Short-Time Fourier Transform, STFT)。然而這個方法也同樣有

其限制，就是切段的長度究竟該如何選擇。如果訊號的頻率很低，則訊號長度必須拉長；如果要分析的頻率很高的訊號，當然必須選擇較短的訊號視窗長度(window length)。但是如果事先無法得知要分析的訊號究竟屬於那個頻帶，這種方法也無法使用。為了解決這個問題小波分析就被推導出來了。

所謂小波，是指一個定義為有限長度且平均值為零的波形。藉著拉伸(stretch)或壓縮(compress)小波函數之後，逐段平移(shifting)，展開被分析的訊號，如圖 3.1 所示。

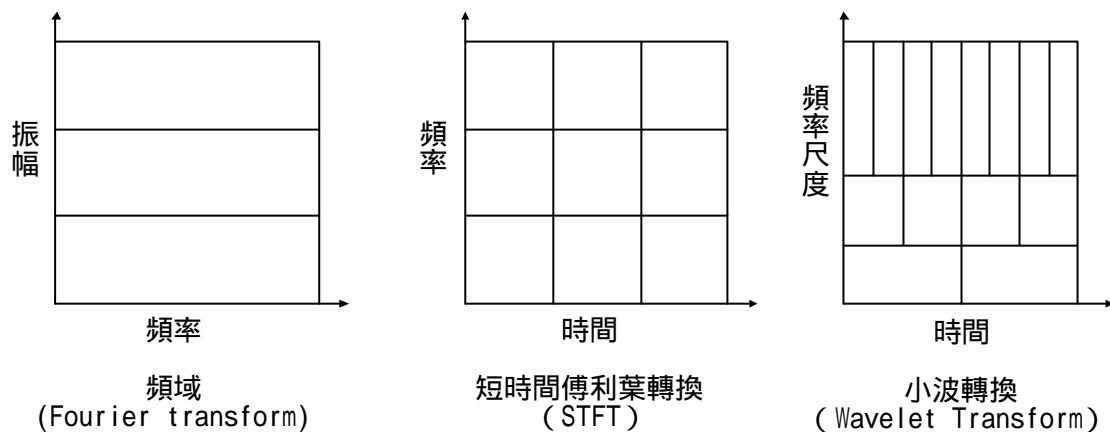


圖 3.1 傅利葉轉換(Fourier transform)、短時間傅利葉轉換 (STFT) 與小波轉換 (Wavelet Transform) 在時間、頻率及尺度構成平面的分割

小波轉換的原理和傅利葉轉換很相似，差異在於小波轉換只是以小波函數經過一定比例之拉伸或壓縮之後，所形成的一系列函數作為展開之基底函數。可以數學式表示如下：

$$W_y f(b, a) \equiv \int_{-\infty}^{\infty} f(t) \overline{y_{b,a}(t)} dt$$

其中， $y(t)$ 為小波函數，

$$y_{b,a}(t) = a^{-1/2} y\left(\frac{t-b}{a}\right) \quad a > 0$$

$\overline{y_{b,a}(t)}$ 為 $y_{b,a}(t)$ 的共軛複數

其中 a 表示對 $y(t)$ 之壓縮或拉伸的比例， b 表示對整段訊號作平移。一般數學式之推導中， a 可以由零至無窮大連續地改變。但是電腦所運算分析的訊號其取樣頻率為有限的，所以 a 值通常是二的次冪，可表示為： $a = 2^n$ ， n 為整數。

本研究所使用的小波分解 Haar 方法之運算程序如下：

1. 根據小波 Haar 函數分別推導出小波分解之高通 低通濾波器
2. 首先將訊號分別通過小波分解之高通與低通濾波器，便可以分別得到高頻(cd1)與低頻(ca1)的兩個分量訊號。再對 cd1 與 ca1 做 2 倍降頻(down-sampling)的處理，此處之 cd1 與 ca1 稱為小波係數。
3. 將 ca1 及 cd1 通過小波分解之高通與低通濾波器，再經過一次降頻之後可得第二層的小波係數 cd1_cd2、cd1_ca2 與 ca1_cd2、ca1_ca2。

本研究是以不整脈心電圖資料庫(MIT/BIH arrhythmia database)第二肢導極(Lead II)之 QRS 波為研究資料來源。QRS 波之取得是以 R 波為中心，共取 64 點為一個研究用之 QRS 波信號，記錄時間約 30 分 5 秒，取樣頻率為 360Hz，故每通道包含資料數為 650000 筆資料，如式 3.2 所示：

$$360 \frac{\text{筆}}{\text{秒}} \times \left(30 \text{分} \times 60 \frac{\text{秒}}{\text{分}} + 5 \text{秒} \right) = 649800 \text{筆} \approx 650000 \text{筆} \quad \text{式 (3.2)}$$

要正確地分離出 ECG 所需要分解的次數，則視實驗所取樣的結果而定。例如實驗中以 64 點取樣，則需要做兩層三次的小波分解，分離出 ECG 頻帶之小波係數。如圖 3.2 所示。

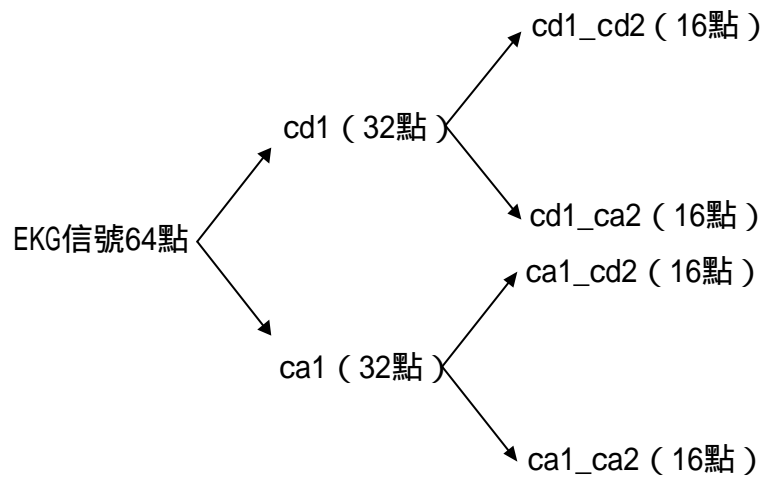


圖 3.2 ECG 訊號的小波分解程序

因為小波分解包含一次降頻的程序，所以各分頻中小波係數的長度會越來越短，在上圖中，cd1 與 ca1 的長度是 ECG 訊號的一半，cd1_cd2 與 cd1_ca2 的長度又是 cd1 的一半，依此類推。如圖 3.3 所示。

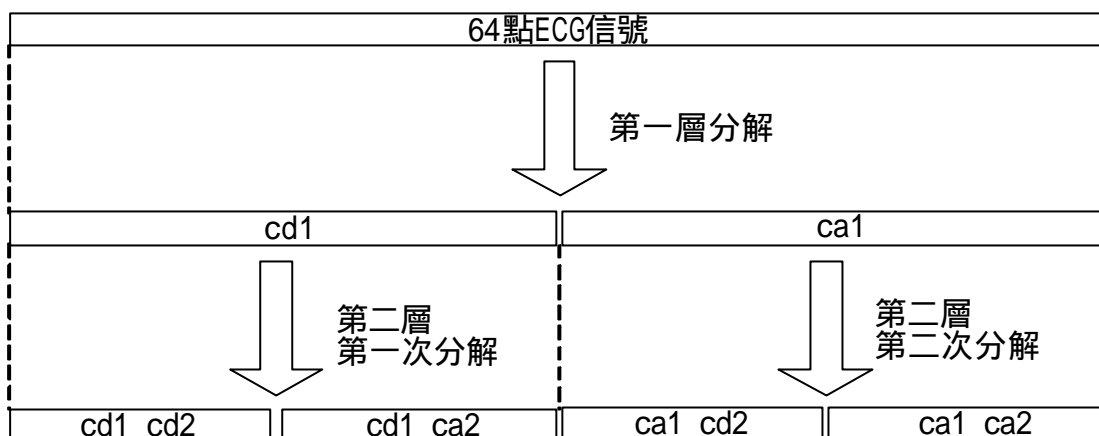


圖 3.3 三次小波分解後的係數長度

我們知道在這一個週期的心電圖時間內，共取了 64 個點，如圖 3.4。如果將這段訊號依照圖 3.3 的方式進行小波分解，所得到的係數分佈如圖 3.5 至 3.7 所示。其中可以發現：ca1、cd1 的長度是原訊號的一半，32 點，如圖 3.5；cd1_cd2 與 cd1_ca2 是由 cd1 在分解而得到，所以長度是 cd1 的一半，16 點，如圖 3.6；ca1_cd2 與 ca1_ca2 是由 ca1 在分解而得到，所以長度是 ca1 的一半，16 點，如圖 3.7。

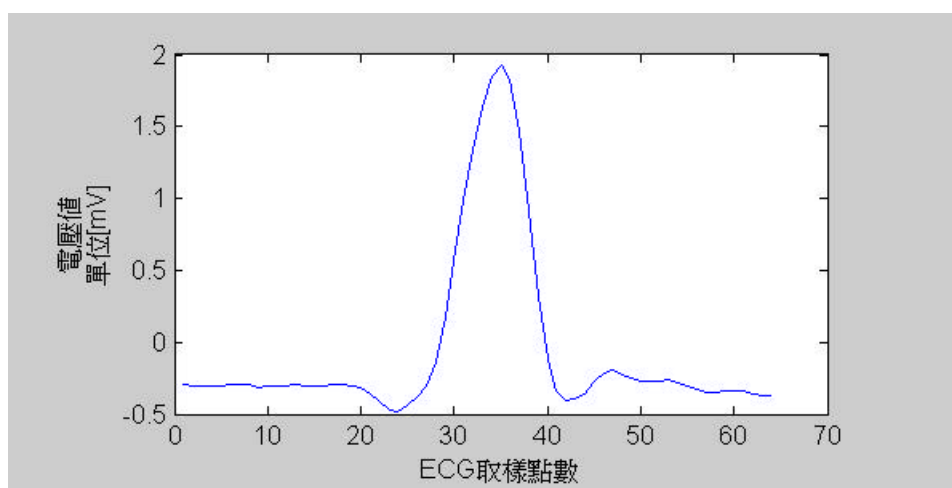


圖 3.4 一組 ECG 取樣圖

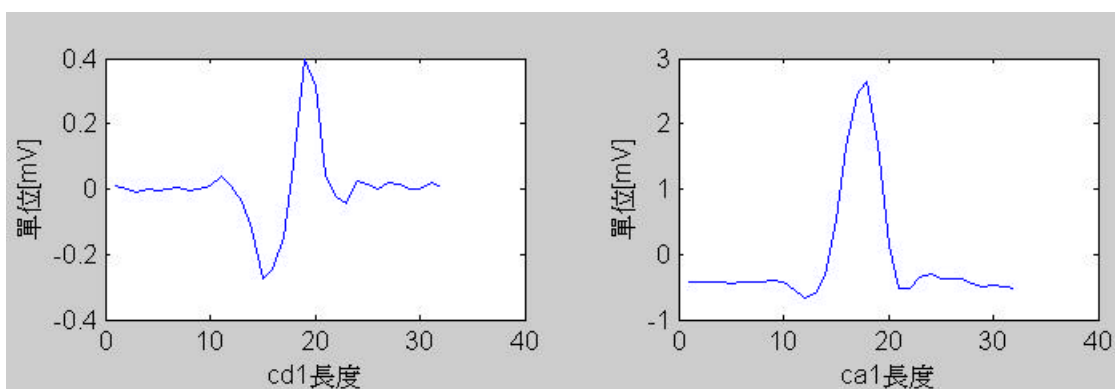


圖 3.5 第一層小波分解 (含 $ca1$ 、 $cd1$)

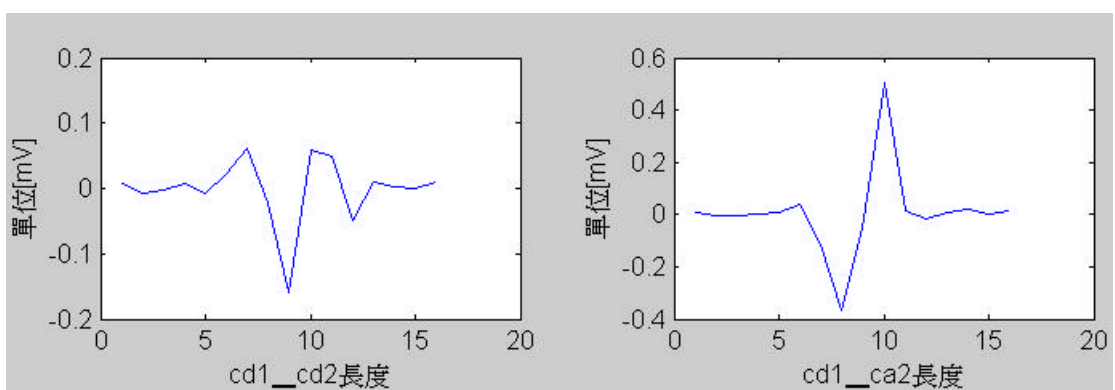


圖 3.6 $cd1$ 之分解 (含 $cd1_cd2$ 、 $cd1_ca_2$)

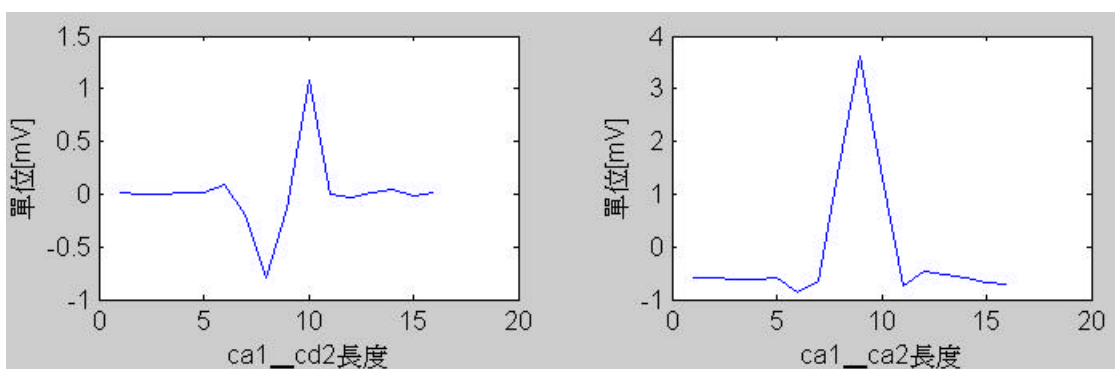


圖 3.7 $ca1$ 之分解 (含 $ca1_cd2$ 、 $ca1_ca_2$)

3.5 特徵值之產生

本研究使用的特徵值共十組[19]，其意義分別為：

- 1、 平均值 \bar{x}_j ：設病患資料檔為 j 筆資料，每位病患所有 ECG 信號總數為 n ，其 ECG 信號之平均值為 \bar{x}_j ，則：

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_i$$

- 2、 變異數 s_x^2 ：設病患資料檔為 j 筆資料，每位病患所有 ECG 信號總數為 n ，其 ECG 信號之變異數為 s_x^2 ：

$$s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_j)^2$$

- 3、 $\min(\text{cd1}) / \max(\text{cd1})$ ：第一層分解之高頻(cd1)分量訊號之最小值與最大值之比值。
- 4、 $\min(\text{cd1_cd1}) / \max(\text{cd1_cd1})$ ：第二層分解的小波係數 cd1_cd1 之最小值與最大值之比值。
- 5、 $(\text{cd1 之 autocorrelation})^2$ ：為第一層分解之高頻(cd1)分量訊號之自相關 (autocorrelation) 函數之平方為的是取正值。
- 6、 $(\text{cd1_cd1 之 autocorrelation})^2$ ：為第二層分解的小波係數 cd1_cd1 自相關 (autocorrelation) 函數之平方為的是取正值；
- 7、 cd1^2 ：為第一層分解之高頻(cd1)分量訊號之能量。
- 8、 $(\text{cd1_cd1})^2$ ：為第二層分解的小波係數 cd1_cd1 之能量。
- 9、 $\min(\text{ca1}) / \max(\text{ca1})$ ：第一層分解之低頻(ca1)分量訊號之最小值與最大值之比值。
- 10、 $\min(\text{ca1_cd1}) / \max(\text{ca1_cd1})$ ：第二層分解的小波係數 ca1_cd1 之最小值與最大值之比值。

第四章 分類法

本章將介紹如何利用小波分析方法將 ECG 分類。本研究以小波分析理論為基礎，將資料庫所取得的 ECG 藉由小波分析得到每個 ECG 之分解，並以歐氏距離（Euclidean Distance）、線性區別分析（Linear Discriminant analysis）、倒傳遞網路（Backpropagation）及徑向基網路（Radial Basis Networks）等四種方式做實驗分類。

4.1 最近鄰法（Nearest Neighbor）

要將資料作分類，最直接想到的方式是找出測試資料和已知分類的資料類別中最近的那一類。例如如何定義二人之間體型的"距離"？最自然想到的是歐氏距離（Euclidean Distance）。如國中男生班上新轉入之一位學生的身高 170 公分，體重 60 公斤，此座標點為（170,60），而班上身高及體重分類為已發育及未發育兩種群集座標，如圖 4.1 所示分別為（172,68）及（160,56），計算此學生與已發育群集座標歐氏距離 d_1 為：

$$d_1 = \sqrt{(170 - 172)^2 + (60 - 68)^2} = \sqrt{68} = 8.2462$$

此學生與未發育群集座標歐氏距離 d_2 為：

$$d_2 = \sqrt{(170 - 160)^2 + (60 - 56)^2} = \sqrt{116} = 10.7703$$

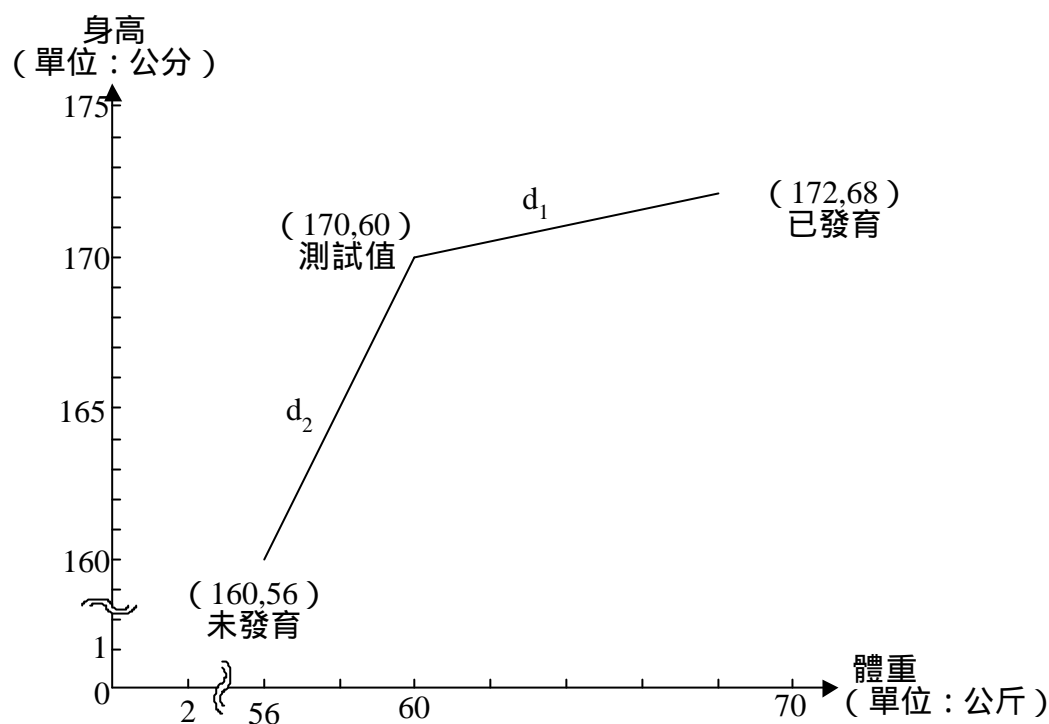


圖 4.1 測試值與已知群集之歐氏距離

由於歐氏距離有根號計算之麻煩，因此通常以歐氏距離平方（Squared Euclidean Distanc）為依據做分析。用前例數據求此學生與已發育群集座標歐氏距離平方為：

$$(170 - 172)^2 + (60 - 68)^2 = 68 = d_1^2$$

此學生與未發育群集座標歐氏距離平方為：

$$(170 - 160)^2 + (60 - 56)^2 = 116 = d_2^2$$

由以上兩種方式皆可得到此新轉入學生與已發育學生之群集距離較未發育學生之群集距離近，因此分類為已發育。

若原本即為距離的資料可以直接用來計算，距離基本尚可為一般的分數。但若原本為座標類分數（即多個分數同為一群（cluster）），

則可經由解析幾何的座標運算公式算成一般距離。不過在各個群（cluster）間的運算時，所取決的距離也就是上述的「一般距離」可以是解析幾何學距離運算公式的直接值或平方值。

4.2 線性區別分析（Linear Discriminant Analysis）

區別分析是將不同群體的個體分離以及將新的東西歸類到前面所說的群體中的多變量技巧[15]。而區別分析的目的一般有：

1. 確定在兩個以上事先界定之群體的一組變數上的平均分數間是否有統計上的顯差異存在。
2. 確定那些預測變數最能解釋兩個或以上群體之平均分數的差異。
3. 建立將統計單位依據他們在一組預變數上的分數劃歸到不同群體的程序。
4. 建立由一組預測變數所形成之群體間的區別構面的數目和組合。

我們可以利用區別函數將不同群體的個體盡可能清楚的分開，並根據區別函數對新的東西進行分類或歸類。舉個例子來說，醫院在病人的記錄中，使用有關之變數（如體重、身高、年齡、性別、血壓、血糖含量、膽固醇含量），以便判斷哪些變數最能區別將來會"無法康復"（第一群）、"部分康復"（第二群）以及"完全康復"（第三群）。

本研究在線性區別分析部份中，所使用的方法是線性區別分析中的馬氏距離法。兩個平均數修正之重心 x 與 y 之間的馬氏距離定義為：

$$(x - y)'s^{-1}(x - y)$$

其中 s 為資料的共變異矩陣。

線性區別分析計算方式，以 3.1 節的例子做說明。假定已發育

及未發育學生數分別為 n_1 及 n_2 ，並設定區別變數 x 表身高， y 表體重，首先算出已發育學生身高、體重的平均數向量及共變異數矩陣分別為

$\begin{bmatrix} \bar{x}_1 \\ \bar{y}_1 \end{bmatrix}$ 及 S_1 ，而未發育學生身高、體重的平均數向量及共變異數矩陣分

別為 $\begin{bmatrix} \bar{x}_2 \\ \bar{y}_2 \end{bmatrix}$ 及 S_2 。然後算出"綜合樣本共變異矩陣"為：

$$S = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{(n_1 + n_2 - 2)} \quad (\text{式 4.1})$$

線性區別規則 (Discriminant Rule) 如下：

令

$$d_1 = [\bar{x}_1, \bar{y}_1] S^{-1} \begin{bmatrix} x \\ y \end{bmatrix} - \frac{1}{2} [\bar{x}_1, \bar{y}_1] S^{-1} \begin{bmatrix} \bar{x}_1 \\ \bar{y}_1 \end{bmatrix} \quad (\text{式 4.2})$$

$$d_2 = [\bar{x}_2, \bar{y}_2] S^{-1} \begin{bmatrix} x \\ y \end{bmatrix} - \frac{1}{2} [\bar{x}_2, \bar{y}_2] S^{-1} \begin{bmatrix} \bar{x}_2 \\ \bar{y}_2 \end{bmatrix} \quad (\text{式 4.3})$$

而當新轉入之一位學生的身高 x 公分，體重 y 公斤，帶入式 4.2 及式 4.3 的 $\begin{bmatrix} x \\ y \end{bmatrix}$ 中，如果 $d_1 > d_2$ ，就判他屬於第一個群體（已發育），若是 $d_2 > d_1$ ，就判他屬於第二個群體（未發育）。

4.3 倒傳遞網路

本節將簡單地介紹倒傳遞網路 (Back-propagation network ; BPN)，以做為分析之基礎。

4.3.1 倒傳遞網路架構

介紹倒傳遞網路前，先要了解其神經元模型，如圖 4.2 所示。神

經元模型，為一個具有 R 個輸入的基本神經元，每個輸入都用一個適當的權重值 w 來加權。此加權後輸入和偏權值 b 總和後形成轉移函數 f （即 $f(Wp + b)$ ）來產生神經元的輸出。

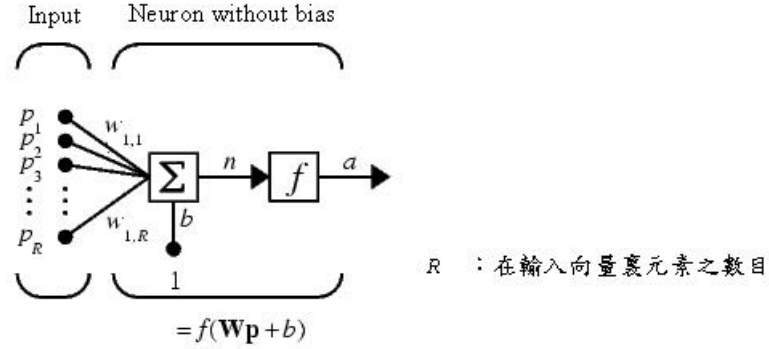


圖 4.2 倒傳遞網路神經元模型

倒傳遞網路（Back-Propagation Network，BP）是類神經網路中應用最普遍的模式，標準的倒傳遞網路演算法是一種梯度下降（Gradient Descent）演算法，網路權重值是沿著性能函數的負的梯度方向移動，而將誤差函數予以最小化。最常見到採用倒傳遞演算法的網路架構是多層前饋網路，如圖 4.3 所示，本研究便是以此為架構。其網路架構，其中包括：

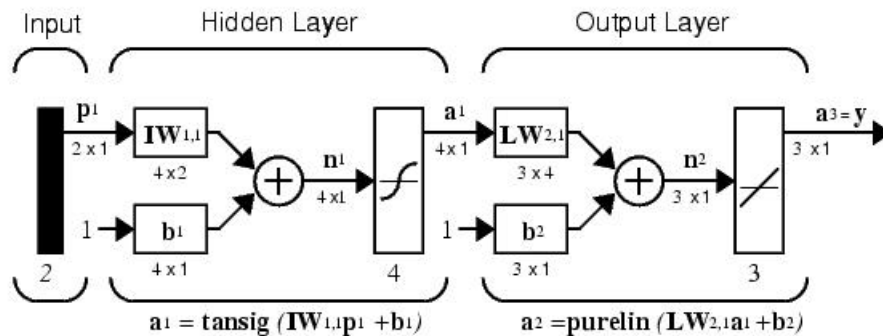


圖 4.3 多層前饋網路

1. 輸入層：用以表現網路的輸入變數，其處理單元依問題而定。使用非線性轉換函數。本研究使用 10 個特徵值作分類，故輸

入變數有 10 個。

2. 隱藏層：用以表現輸入處理單元間的交互影響，其處理單元數目並無標準方法可以決定，經常需以試驗方式決定其最佳數目。使用非線性轉換函數。網路可以不只一層隱藏層，也可以沒有隱藏層。本研究經目前實驗結果，以使用一層隱藏層的辨識率最佳，隱藏層為 30 個神經元。

多層前饋網路隱藏層常用的兩種轉移函數，一是對數雙彎曲轉移函數，如圖 4.4 所示，另一種是正切雙彎曲轉移函數，如圖 4.5 所示。對數雙彎曲轉移函數產生介於 0 和 1 之間的輸出，而正切雙彎曲轉移函數則產生介於 -1 和 1 之間的輸出，本研究採用正切雙彎曲轉移函數創建網路。

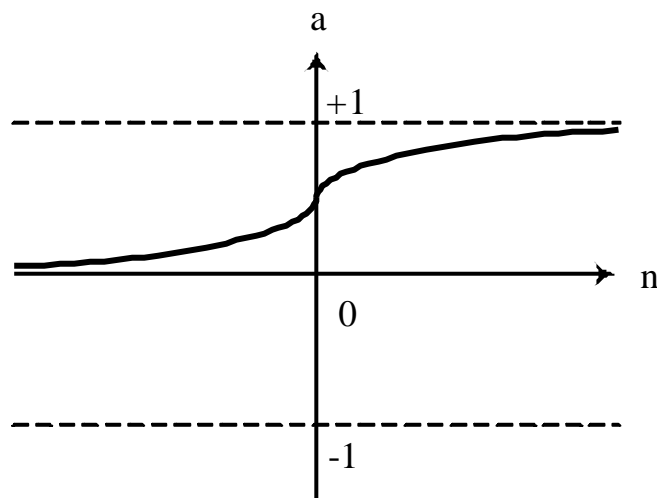


圖 4.4 對數雙彎曲轉移函數

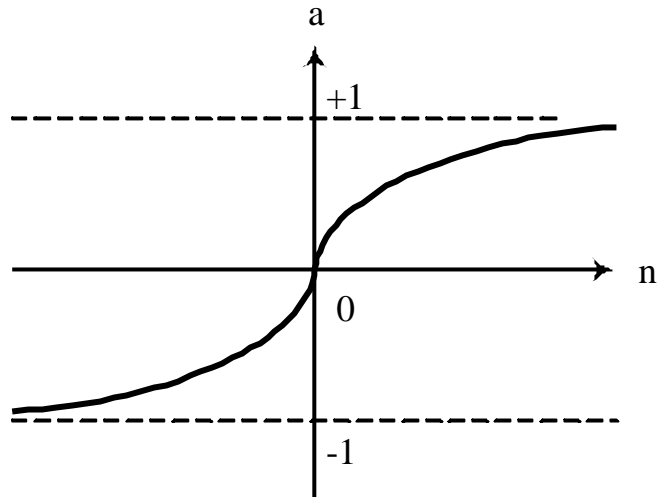


圖 4.5 正切雙彎曲轉移函數

3. 輸出層：用以表現網路的輸出變數，其處理單元數目依問題而定，本研究因分類為六種病症，故輸出層有 6 個神經元。一般經常用線性轉換函數於輸出層內，如圖 4.6 所示。其目的是輸出不被限制在很小的範圍內，可使網路產生範圍在-1 到+1 以外的數據。故本研究也採用具線性轉換函數的神經元。

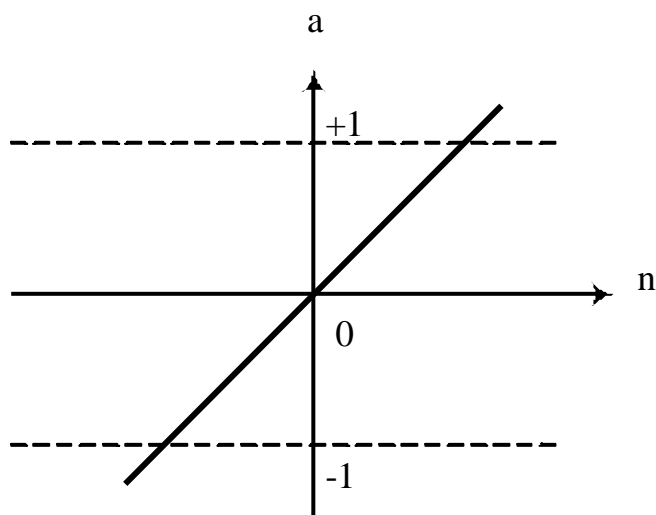


圖 4.6 線性轉換函數

「倒傳遞」此是針對非線性多層網路內計算梯度的方法。基本的倒傳遞演算法在其他不同標準的最佳化技巧上，而有許多種變化的演算方法，就像是牛頓法和共軛梯度法（Conjugate Gradient）。經過適當訓練後的倒傳遞網路，當給定此網路新的輸入值（向量）時，網路就能給出合理的答案。基本上這個新的輸入值所引導出的值是利用網路廣義化的特性。廣義化的特性使得只須用具代表性的輸入/目標對數據集來訓練的網路，就能夠對於不是訓練數據集內的輸入也可以得到滿意的輸出結果。然而從曲線擬合的觀點來看，倒傳遞網路的這種函數逼近（即非線性迴歸）能力，是只對數據內插有效，而對外插就不一定準確了。因此倒傳遞網路之輸入/輸出的關係，可以看做是一種非線性的輸入/輸出映射關係，而其廣義化特性就可以看做是一種非線性內插。

4.3.2 倒傳遞演算法

對於多層網路，每一層的輸出將成為下一層的輸入，因此可以下式表示：

$$\mathbf{a}^{m+1} = \mathbf{f}^{m+1}(\mathbf{W}^{m+1}\mathbf{a}^m + \mathbf{b}^{m+1}) \quad m = 0, 2, \Lambda, M-1 \quad (\text{式 4.4})$$

M 為網路中層的數目。

第一層神經元所接受到外部的輸入如下：

$$\mathbf{a}^0 = \mathbf{p} \quad (\text{式 4.5})$$

式 4.5 為式 4.4 的起始點。最後一層神經元的輸出為此網路的輸出如下：

$$\mathbf{a} = \mathbf{a}^M \quad (\text{式 4.6})$$

1. 性能指標

用於多層網路的倒傳遞演算法 (BP) 是一種廣義化的最小均方演算法 (LMS), 而且 BP 演算法和 LMS 演算法, 這兩種演算法都是使用均方誤差 (Mean Square Error) 當作性能指標。

當每一個輸入施於網路, 就比較其網路輸出與目標。演算法應該要來調整網路參數已達到最小均方誤差 $F(x)$:

$$F(\mathbf{x}) = E[e^2] = E[(t - a)^2] \quad (\text{式 4.7})$$

x 為網路權重的權重值和偏權向量

如果網路具有多個輸出, 則可將式 4.7 廣義化, 如下式:

$$F(\mathbf{x}) = E[\mathbf{e}^T \mathbf{e}] = E[(\mathbf{t} - \mathbf{a})^T (\mathbf{t} - \mathbf{a})] \quad (\text{式 4.8})$$

如同使用 LMS 演算法, 將式 4.8 近似化, 如下式:

$$\hat{F}(\mathbf{x}) = (\mathbf{t}(k) - \mathbf{a}(k))^T (\mathbf{t}(k) - \mathbf{a}(k)) = \mathbf{e}^T(k) \mathbf{e}(k) \quad (\text{式 4.9})$$

其中平方誤差的期望值已經由第 k 次疊代的平方誤差所取代。

對於近似誤差的均方誤差其最陡坡降 (Steepest Descent) 演算法如下式:

$$W_{i,j}^m(k+1) = W_{i,j}^m(k) - \mathbf{a} \frac{\partial \hat{F}}{\partial W_{i,j}^m} \quad (\text{式 4.10})$$

$$b_i^m(k+1) = b_i^m(k) - \mathbf{a} \frac{\partial \hat{F}}{\partial b_i^m} \quad (\text{式 4.11})$$

a 表學習速率

2. 連鎖律

使用微積分中的連鎖律, 可以找出在式 4.10 及式 4.11 兩式中的微分, 如下:

$$\frac{\partial \hat{F}}{\partial W_{i,j}^m} = \frac{\partial \hat{F}}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial W_{i,j}^m} \quad (\text{式 4.12})$$

$$\frac{\partial \hat{F}}{\partial b_i^m} = \frac{\partial \hat{F}}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial b_i^m} \quad (\text{式 4.13})$$

層 m 的淨輸入 n 是層 m 中權重值和偏權值的顯函數, 如下式:

$$n_i^m = \sum_{j=1}^{S^{m-1}} W_{i,j}^m a_j^{m-1} + b_i^m \quad (\text{式 4.14})$$

因此

$$\frac{\partial n_i^m}{\partial W_{i,j}^m} = a_j^{m-1}, \frac{\partial n_i^m}{\partial b_i^m} = 1 \quad (\text{式 4.15})$$

接下定義下式:

$$S_i^m \equiv \frac{\partial \hat{F}}{\partial n_i^m} \quad (\text{式 4.16})$$

\hat{F} 的靈敏度 (Sensitivity) 就是在層 m 的淨輸入的第 I 個元素的變化量。接著, 可以將式 4.12 及式 4.13 簡化成如下:

$$\frac{\partial \hat{F}}{\partial W_{i,j}^m} = s_i^m a_j^{m-1} \quad (\text{式 4.17})$$

$$\frac{\partial \hat{F}}{\partial b_i^m} = s_i^m \quad (\text{式 4.18})$$

近似化的最陡坡降演算法式 4.10 及式 4.11 可以寫成如下：

$$W_{i,j}^m(k+1) = W_{i,j}^m(k) - \alpha s_i^m a_j^{m-1} \quad (\text{式 4.19})$$

$$b_i^m(k+1) = b_i^m(k) - \alpha s_i^m \quad (\text{式 4.20})$$

上兩式若以矩陣型式來表示，則可以寫成如下：

$$\mathbf{W}^m(k+1) = \mathbf{W}^m(k) - \alpha \mathbf{s}^m (\mathbf{a}^{m-1})^T \quad (\text{式 4.21})$$

$$\mathbf{b}^m(k+1) = \mathbf{b}^m(k) - \alpha \mathbf{s}^m \quad (\text{式 4.22})$$

其中

$$\mathbf{s}^m \equiv \frac{\partial \hat{F}}{\partial \mathbf{n}^m} = \begin{bmatrix} \frac{\partial \hat{F}}{\partial n_1^m} \\ \frac{\partial \hat{F}}{\partial n_2^m} \\ \vdots \\ \frac{\partial \hat{F}}{\partial n_{s^m}^m} \end{bmatrix} \quad (\text{式 4.23})$$

3. 靈敏度

靈敏度 s^m 的計算過程是一種遞迴關係，就是從第 $m+1$ 層的靈敏度來計算第 m 層的靈敏度。我們使用 jacobian 矩陣來推導靈敏度的遞迴關係：

$$\frac{\partial \mathbf{n}^{m+1}}{\partial \mathbf{n}^m} = \begin{bmatrix} \frac{\partial n_1^{m+1}}{\partial n_1^m} & \frac{\partial n_1^{m+1}}{\partial n_2^m} & \Lambda & \frac{\partial n_1^{m+1}}{\partial n_{s^m}^m} \\ \frac{\partial n_2^{m+1}}{\partial n_1^m} & \frac{\partial n_2^{m+1}}{\partial n_2^m} & \Lambda & \frac{\partial n_2^{m+1}}{\partial n_{s^m}^m} \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ \frac{\partial n_{s^{m+1}}^{m+1}}{\partial n_1^m} & \frac{\partial n_{s^{m+1}}^{m+1}}{\partial n_2^m} & \Lambda & \frac{\partial n_{s^{m+1}}^{m+1}}{\partial n_{s^m}^m} \end{bmatrix} \quad (\text{式 4.24})$$

若考慮此矩陣的第 i,j 元素，此矩陣表示式如下：

$$\begin{aligned} \frac{\partial n_i^{m+1}}{\partial n_j^m} &= \frac{\partial \left(\sum_{l=1}^{s^m} w_{i,l}^{m+1} a_l^m + b_i^{m+1} \right)}{\partial n_j^m} \\ &= w_{i,l}^{m+1} \frac{\partial a_j^m}{\partial n_j^m} = w_{i,l}^{m+1} \frac{\partial f^m(n_j^m)}{\partial n_j^m} = w_{i,l}^{m+1} f^m(n_j^m) \end{aligned} \quad (\text{式 4.25})$$

$$\text{其中 } f^m(n_j^m) = \frac{\partial f^m(n_j^m)}{\partial n_j^m} \quad (\text{式 4.26})$$

因此，jacobian 矩陣可以寫成：

$$\frac{\partial \mathbf{n}_j^{m+1}}{\partial \mathbf{n}_j^m} = \mathbf{W}^{m+1} \mathbf{F}^m(\mathbf{n}^m) \quad (\text{式 4.27})$$

$$\mathbf{F}^m(\mathbf{n}^m) = \begin{bmatrix} f^m(n_1^m) & 0 & \Lambda & 0 \\ 0 & f^m(n_2^m) & \Lambda & 0 \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ 0 & 0 & \Lambda & f^m(n_{s^m}^m) \end{bmatrix} \quad (\text{式 4.28})$$

藉由使用在矩陣型式中的鏈鎖律寫出靈敏度的遞迴關係：

$$\begin{aligned}
 \mathbf{s}^m &= \frac{\partial \hat{F}}{\partial \mathbf{n}^m} = \left(\frac{\partial \mathbf{n}^{m+1}}{\partial \mathbf{n}^m} \right)^T \frac{\partial \hat{F}}{\partial \mathbf{n}^{m+1}} \\
 &= \mathbf{F}^m(\mathbf{n}^m) (\mathbf{W}^{m+1})^T \frac{\partial \hat{F}}{\partial \mathbf{n}^{m+1}} \\
 &= \mathbf{F}^m(\mathbf{n}^m) (\mathbf{W}^{m+1})^T \mathbf{s}^{m+1} \quad (\text{式 4.29})
 \end{aligned}$$

靈敏度從最後一層到第一層是被倒著傳遞穿過網路，如圖 4.7，這就是為何被稱為倒傳演算的原因了。



圖 4.7 倒傳演算之靈敏度演算方式

4.4 徑向基網路 (Radial Basis Networks)

本節對徑向基網路做簡單介紹，以及說明本研究所採用的方式。

甲、 1 徑向基網路神經元模型

介紹徑向基網路前，先要了解其神經元模型，如圖 4.8 所示。

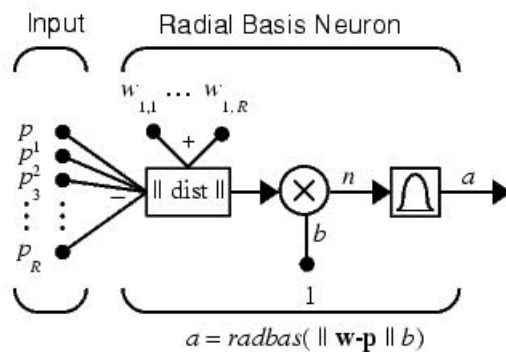


圖 4.8 徑向基神經元模型圖

圖 4.8 中神經元模型為一個具有 R 個輸入元素的徑向基網路。徑向基轉移函數 $radbas$ 的輸入 n 是權重值向量 w 和輸入向量 p 之間的歐幾里得距離（Euclidean Distance）乘上偏權值 b ，如下式：

$$n = \|w - p\|b$$

圖 4.8 中的 $\|dist\|$ 就是 w 和 p 的歐幾里得距離。

徑向基神經元的轉移函數為高斯函數，如下式：

$$a = radbas(n) = e^{-n^2}$$

圖 4.9 為其轉移函數。

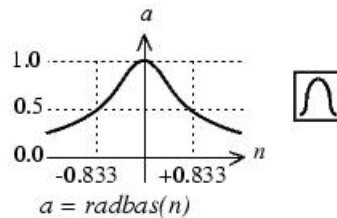


圖 4.9 徑向基轉移函數

當轉移函數的輸入是 0 時，徑向基轉移函數的輸出 a 有最大值 1。當 w 和 p 之間的距離降低時，輸出 a 就增加了。因此徑向基神經元的運作就像是一個偵測器，所以一旦輸入 p 和權重值向量 w 相等時，它就會產生 1。

4.4.2 徑向基網路架構

徑向基網路的架構它是由兩層所組成，分別為具有 s^1 個神經元的

隱藏（Radial Basis 徑向基）層與 S^2 個神經元的輸出（Linear：線性）層，如圖 4.10 所示。

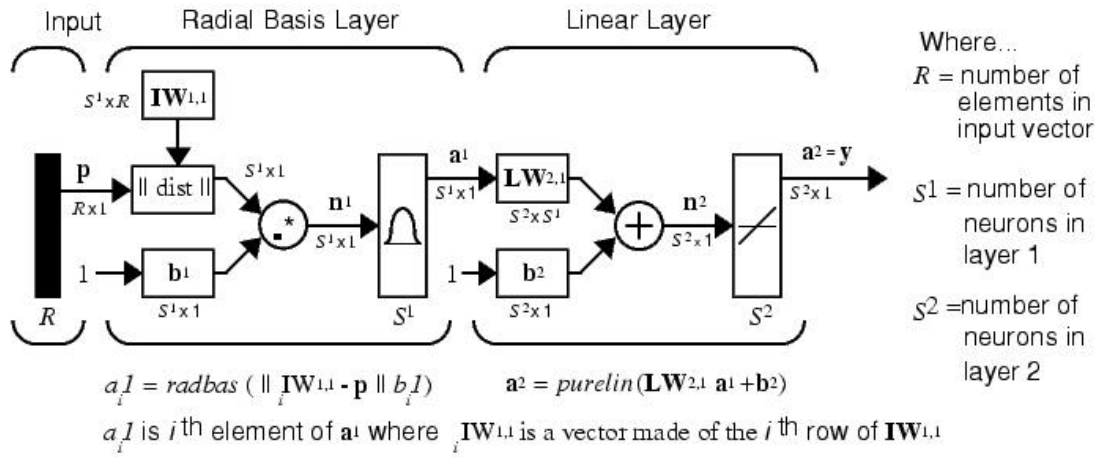


圖 4.10 徑向基網路的架構圖

圖 4.10 中的 $\| \text{dist} \|$ 是輸入權重值矩陣 $IW^{1,1}$ 和輸入向量 p 之間的歐幾里得距離。因此 $\| \text{dist} \|$ 的輸出是具有 S^1 個元素的向量，而其每一個元素即為向量 $IW^{1,1}$ 和輸入向量 p 之間的歐幾里得距離 Z ：

$$Z = \sqrt{\sum_i^n (W_i - p_i)^2}$$

其中

$W: S \times R$ 的權重值矩陣

$P: R \times Q$ 的輸入向量矩陣

$Z: S \times Q$ 的向量距離的矩陣

因此，當 w 和 p 之間的歐幾里得距離大時，徑向基神經元的輸出將接近 0，其輸出對線性（輸出）層的影響幾乎可以忽略。當 w 和 p 之間的歐幾里得距離小時，徑向基神經元的輸出將接近 1，如果為 1 時，則它在第二層中的輸出權重值將傳遞其值到第二層中的線性神

經元。

4.4.3 機率類神經網路架構

徑向基網路和倒傳遞網路一樣可以近似任何連續非線性函數，主要分別於各自使用不同的轉移函數。倒傳遞網路內的隱藏層神經元（節點）使用雙彎曲（Sigmoid）函數，因此函數值在輸入向量空間中無限大的範圍是非零值；而徑向基網路內的轉移函數則是局部的，即輸入向量在越靠近徑向基函數中心，其輸出值越大。

徑向基網路有不同的類型，可分為機率類神經網路（Probabilistic Neural Network; PNN）及廣義迴歸類神經網路（Generalized Regression Neural Network : GRNN）兩種。機率類神經網路一般可用於分類（Classification）問題，而廣義迴歸類神經網路特性是常用於函數逼近（Function Approximation）[16]，因此本研究選擇機率類神經網路。圖 4.11 為機率類神經網路的架構圖。

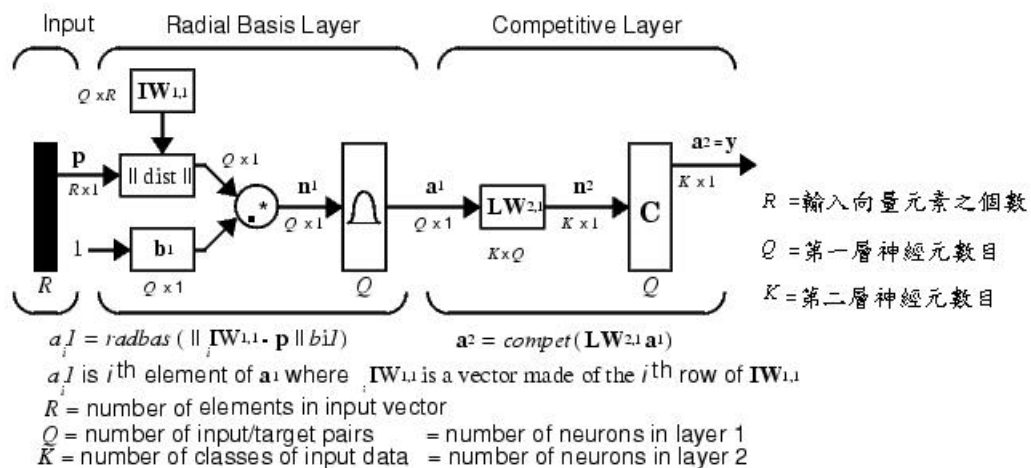


圖 4.11 徑向基網路之機率類神經架構圖

假定 PNN 中有 Q 個輸入向量/目標向量對，每個目標向量有 K

個元素，這些元素其中一個是 1，其餘則為 0。因此，每個輸入向量與 K 種類之中的一個相連結，第一層的輸入權重值 $IW^{1,1}$ 設定為矩陣的轉置 P^T ，及 Q 個訓練用數據集 P 所形成的矩陣之轉置。

當提供一個輸入時， $\|dist\|$ 方塊盒產生一個向量，此向量的元素意指此輸入是多靠近訓練用數據集。這些元素乘以偏權值並且送進徑向基轉移函數。

一個靠近訓練用向量的輸入向量，藉由靠近 1 的數目表示在輸出向量 a^1 之中。如果一個輸入是靠近單一種類的幾個訓練用向量，則它藉由靠近 1 的 a^1 之中的幾個元素所表示。

第二層權重值 $LW^{1,2}$ 設定為目標向量的矩陣 T 。其第一行元素代表輸入的特定種類，第三行均用 1 來表示。由於 K 個輸入種類的每一種乘積 Ta^1 加總了 a^1 的元素。最後第二層競爭 (Compete) 轉移函數產生一個 1 對應於 n^2 內最大的元素，其餘均為 0。因此，因為正確的種類有最大的機率，所以此網路已將輸入向量分類到 K 個種類中特定的一種之內。

第五章 實驗設計

5.1 心電圖資料庫

本研究所購置之不整脈心電圖資料庫(MIT/BIH arrhythmia database)，是該資料 1997 年 5 月出版，CD-ROM 第三版。由麻省理工學院(MIT)與波士頓的以色列貝色醫院(Beth Israel Hospital)所共同研究開發出來的原始資料。資料庫內包含取自 47 位病人之 48 組心電圖，每組資料包含一個標頭檔(Head file，.hea)，一個說明檔(Annotation file，.atr)，一個資料檔(Data file，.dat)，茲分別敘述如下：

- 1、標頭檔：記載著每組資料名稱、ECG 資料長度、取樣頻率，資料所屬導極等，固定為 512 位元組。每個資料檔儲存兩個導極的心電圖資料，一般是第二肢導極(Lead II)及第一胸導極(V1)，取樣頻率為 360Hz，解析度 11bits，電壓值介於 -5mv 至 5mv 之間，記錄時間約 30 分 5 秒。
- 2、說明檔：註解資料檔中每一心電圖之屬性，其 R 波與前一 R 波在資料檔中之間距(intervals)及其所屬病症，資料結構如圖 5.1。較高之六位元是病症代碼，較低之十位元是 R 波與前一 R 波之間距，本研究使用到之病症代碼如表 5.1，每個波形信號之症狀註解至少經由兩位心臟科醫生判定後再相互核對，如果判定後彼此註解不同，則經過研究後再做最後決定。
- 3、資料檔：每通道包含 650000 筆資料，其讀取時之資料結構如圖 5.2，每次讀取三個位元組資料，此三位元組表示二個值，第一個位元組(最低位元組)及第二位元組之低四位元，如圖 5.3，組成第一個通道(通常是 Lead II)心電圖的一個點之值，第二位元組之高四位元及第三位元組如圖 5.4，組成第二通道(通常是 chest V1)心

電圖的一個點之值。

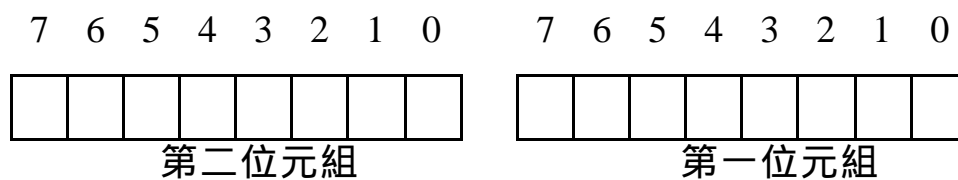


圖 5.1 說明檔(.atr)之格式 (MIT format)

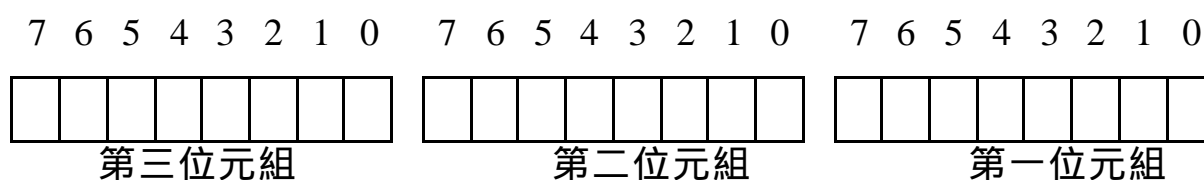


圖 5.2 資料檔(.dat)格式 (format 212)

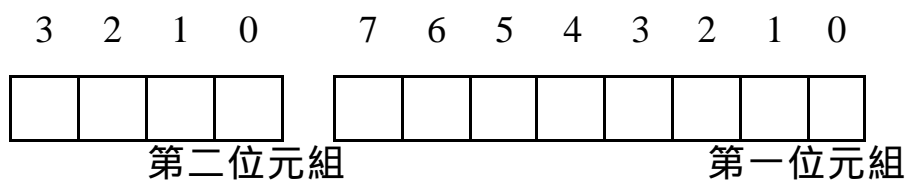


圖 5.3 組成第一個通道之值(通常是 Lead)

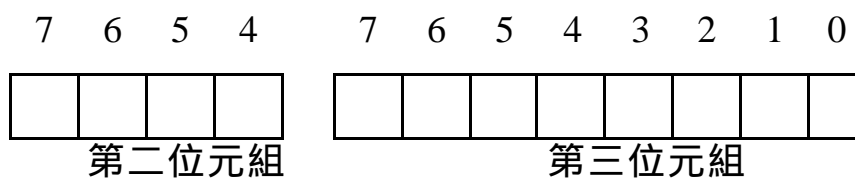


圖 5.4 組成第二通道之值(通常是 chest V1)

表 5.1 本研究使用到之病症代碼

說明檔病症代碼	英文代碼	病 症
1	N	正常 NORM
2	L	左束分支阻塞 LBBB
3	R	右束分支阻塞 RBBB
5	V	心室過早收縮 PVC
8	A	心房過早收縮 APB
12	/	Paced Beat

5.2 心電圖樣本之取得

本研究所取之資料名稱數量如表 5.2，包含六種症狀：

1. 正常 QRS 複合波(Normal QRS)
2. 左束分支阻塞(Left Bundle Branch Block Beat , LBBB)
3. 右束分支阻塞(Right Bundle Branch Block Beat , LBBB)
4. 心室過早收縮(Premature Ventricular Contraction , PVC)
5. 心房過早收縮(Atrial Premature Beat)
6. Paced Beat

本研究是以不整脈心電圖資料庫(MIT/BIH arrhythmia database) 第二肢導極(Lead II) 之 QRS 波為研究資料來源。QRS 波之取得是以 R 波為中心，共取 64 點為一個研究用之 QRS 波信號，取得之詳細過程如圖 5.5 之流程圖所示。

表 5.2 從資料檔 (.dat) 讀取 QRS 複合波數量表

症狀	檔名	測試樣本數量
NORM 正常	103	600
	113	600
	115	600
	123	600
	220	600
	234	600
LBBB 左束分支阻塞	109	600
	111	600
	207	600
	214	600
RBBB 右束分支阻塞	118	600
	124	600
	212	600
	231	600
PVC 心室過早收縮	119	200
	221	150
	233	400
	200	400
APB 心房過早收縮	209	150
	222	100
	232	600
Paced Beat	102	600
	107	600
	217	600
	104	600

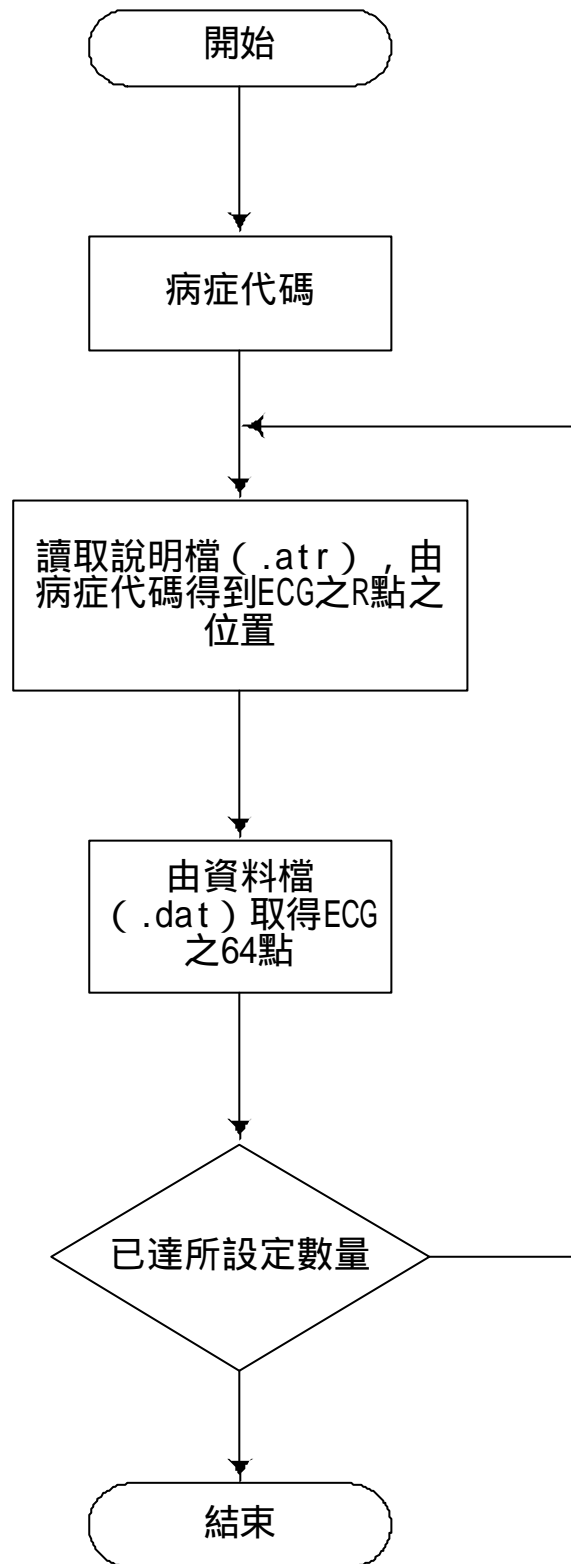


圖 5.5 讀取 EKG 波之流程

5.3 實驗設備

本研究實驗設備為華碩 A7N266 主機板 CPU 為 AMD K7-1.8GHz 及硬碟 20G Bytes 及 Hp 5000 印表機 資料來源為 MIT/BIH arrhythmia database 第三版[17]。研究所用軟體為 MathWorks 公司之 Matlab [18]、Microsoft office 2000 及 Microsoft visio 。

5.4 評估方式

以辨識率來評估系統效能 (performance) , 分為病症辨識及總 ECG 測試信號辨識兩種。若辨識出之病症類別與已知之病症類別相同, 則代表辨識正確, 反之則否, 其辨識率定義如下:

$$\text{病症辨識率} = \frac{\text{正確辨識病症數目}}{\text{辨識病症總數}} \times 100\%$$

若辨識出之信號類別與已知之病症類別相同, 則代表辨識正確, 反之則否, 其辨識率定義如下:

$$\text{信號辨識率} = \frac{\text{正確辨識信號數目}}{\text{辨識信號總數}} \times 100\%$$

5.5 實驗流程圖

本研究之系統流程如圖 5.6 所示, 首先先取得實驗所需之心電圖, 將心電圖作小波分解, 再擷取其特徵值, 然後以四種辨識方式做分類、病症辨識。

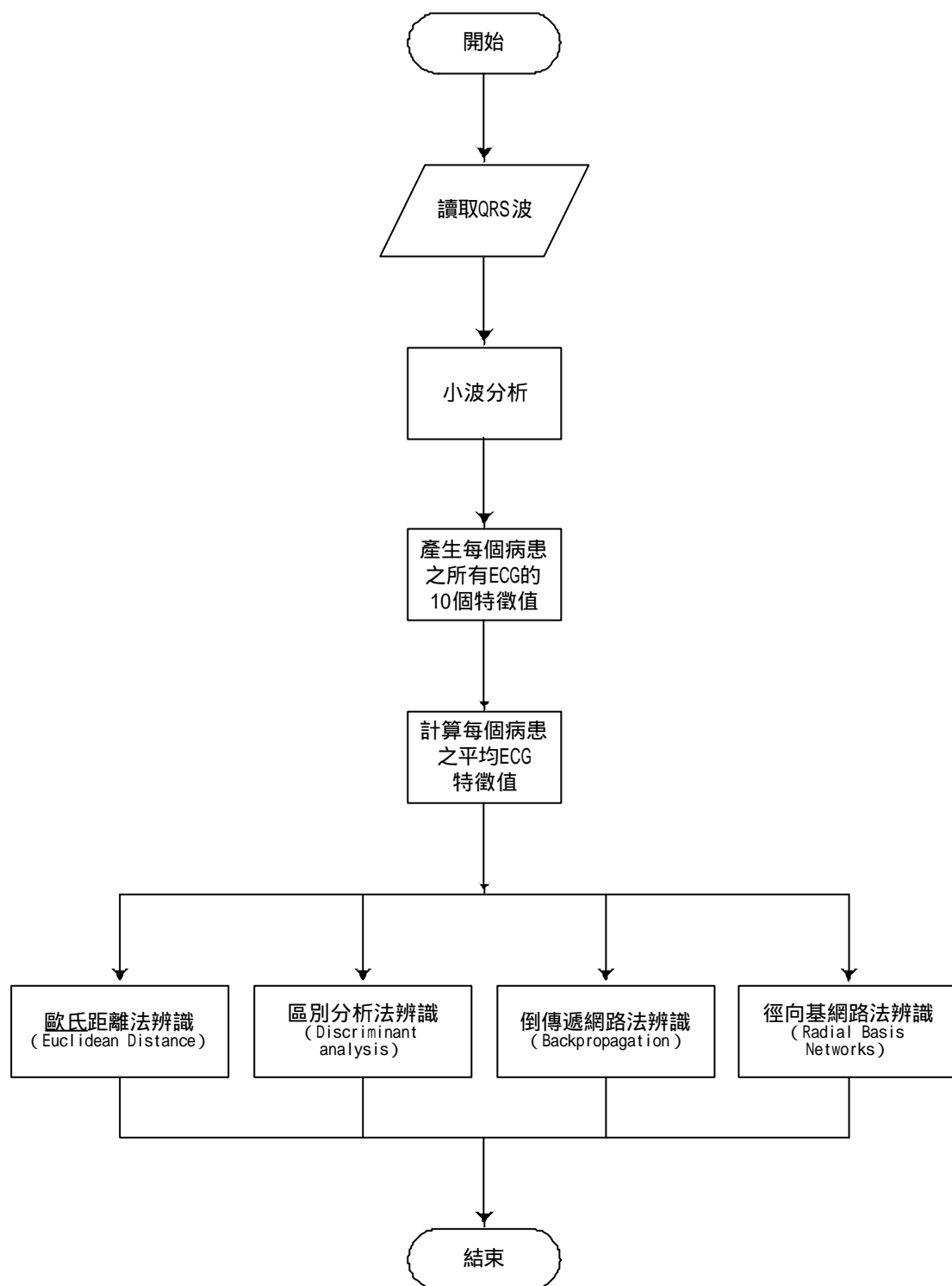


圖 5.6 病症辨識流程

而四種辨識、分類方法流程圖如下：

1. 歐氏距離法辨識流程圖一，分為分類組及測試組，以病症為比較對象，如圖 5.7。

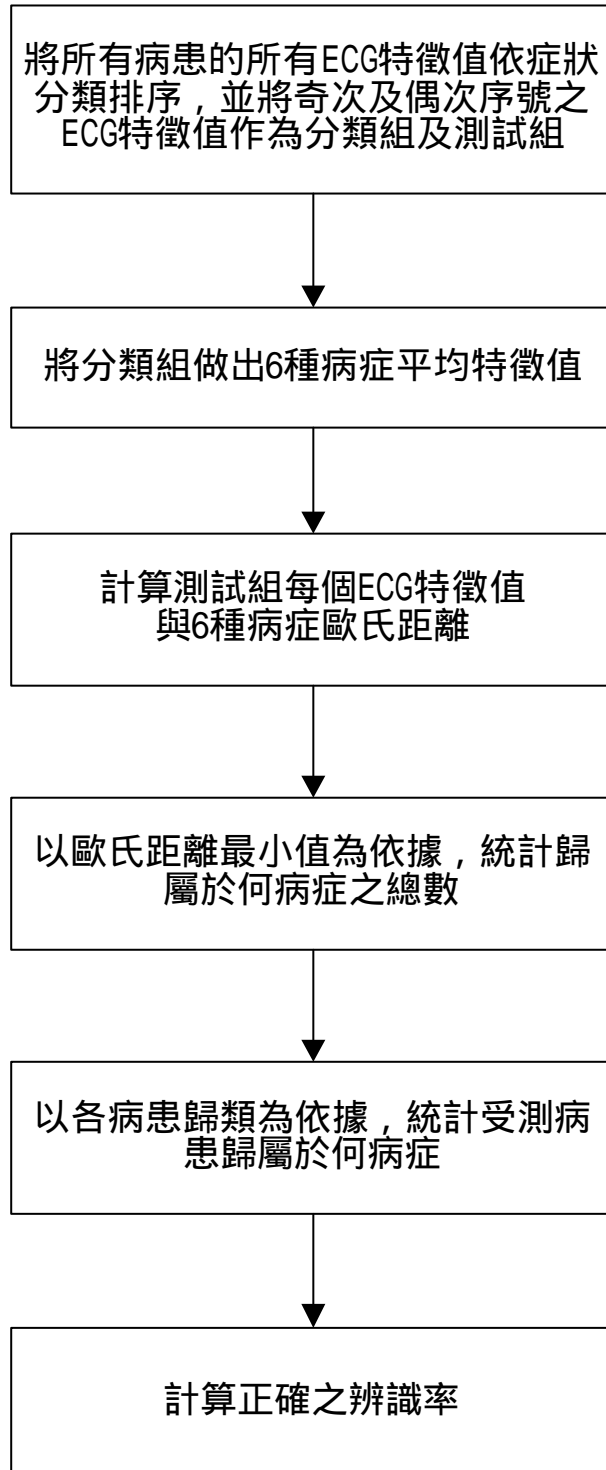


圖 5.7 歐氏距離法辨識流程圖一

2. 歐氏距離法辨識流程圖二，分為分類組及測試組，以病患為比較對象，如圖 5.8。

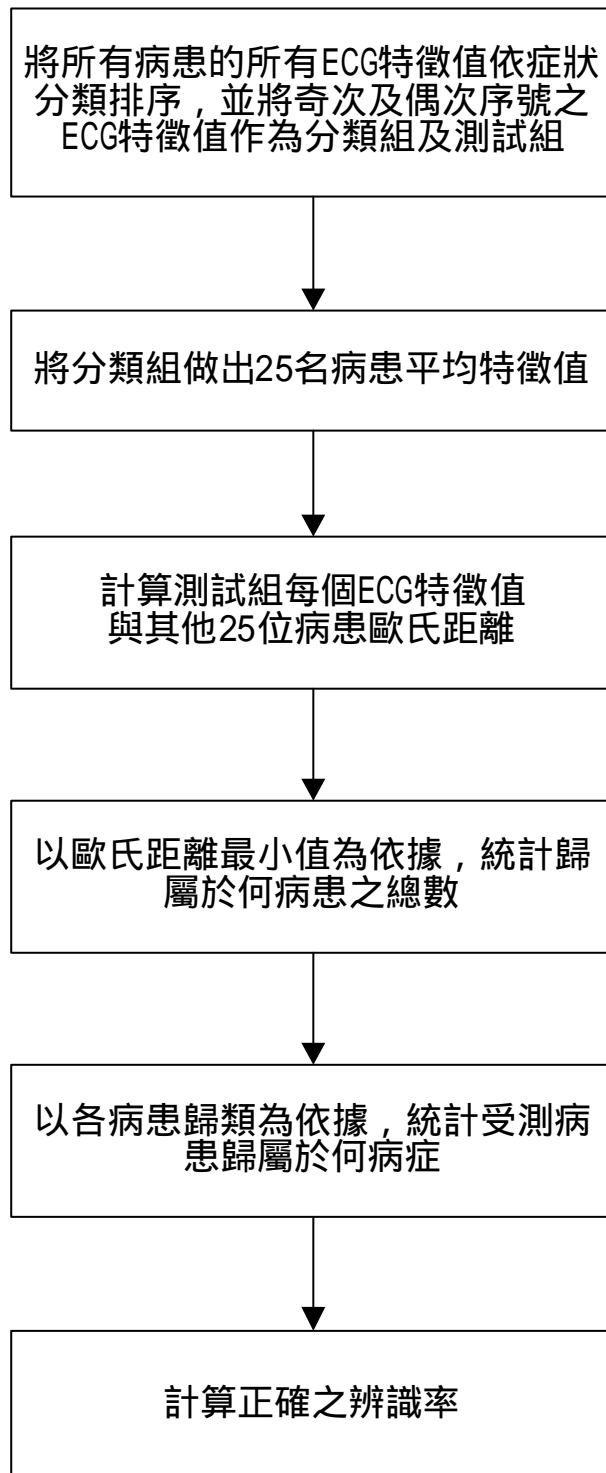


圖 5.8 歐氏距離法辨識流程圖二

3. 歐氏距離法辨識流程圖三，以病症為比較對象。一人為測試者，其餘為分類組，測試者以其平均 ECG 特徵值辨識，如圖 5.9。

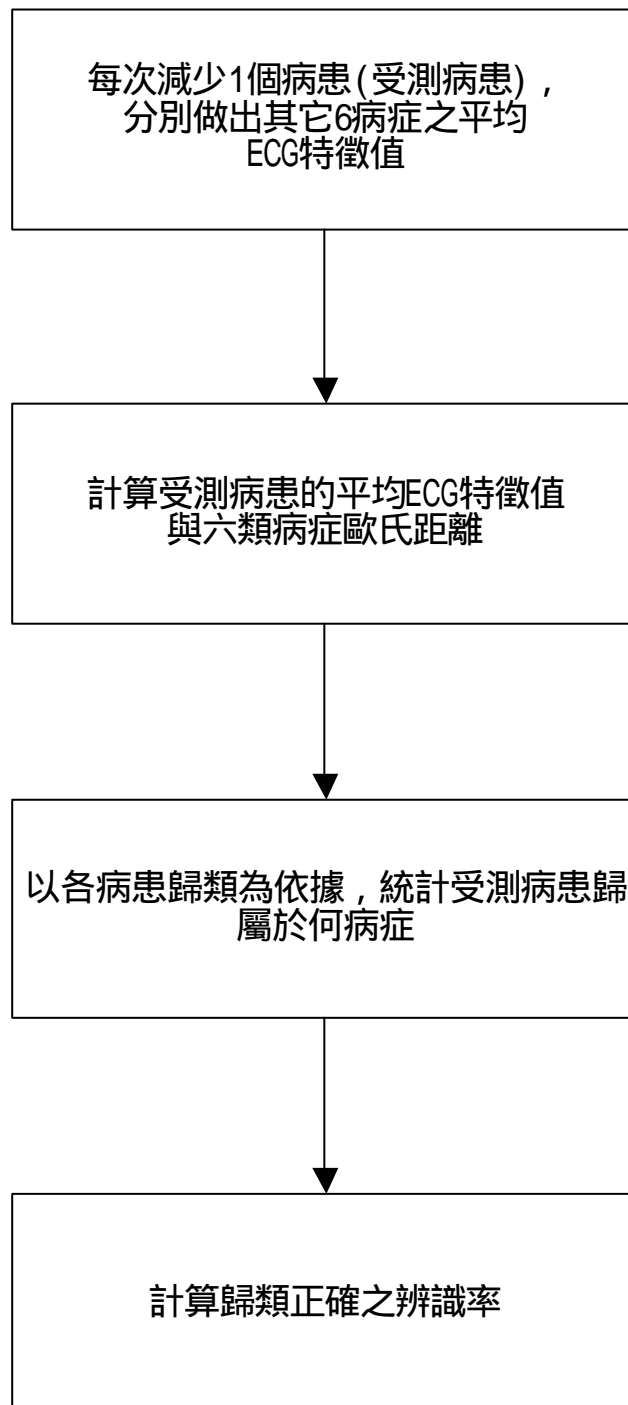


圖 5.9 歐氏距離法辨識流程圖三

4. 歐氏距離法辨識流程圖四，以病患為比較對象。一人為測試者，其餘為分類組，測試者以其平均 ECG 特徵值辨識，如圖 5.10。

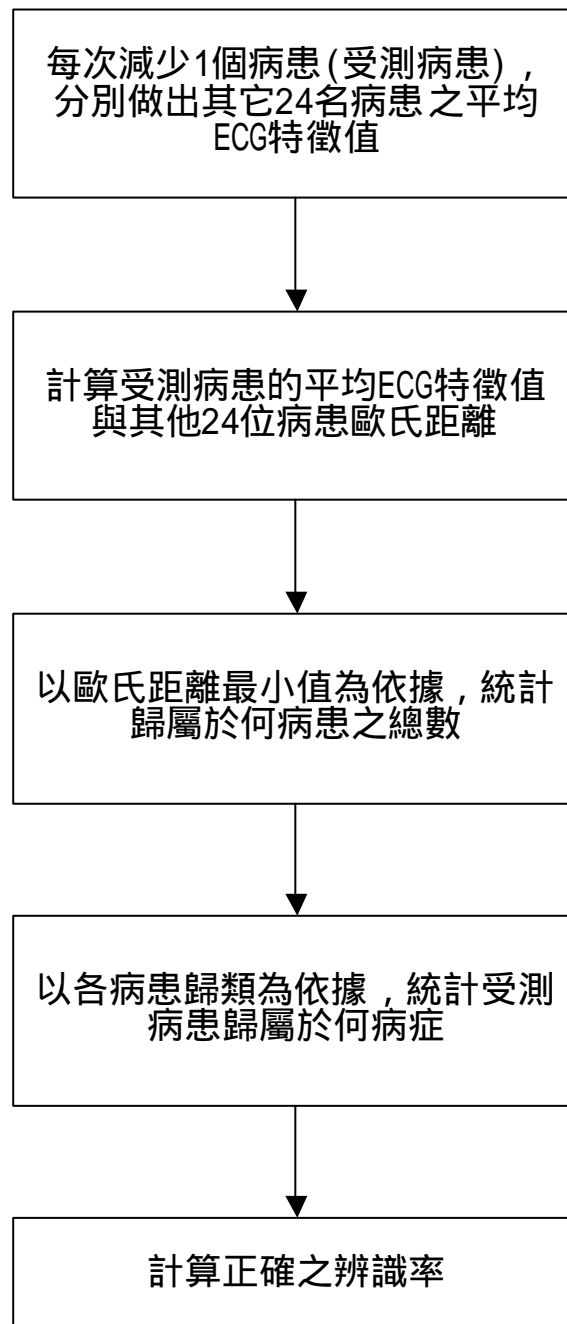


圖 5.10 歐氏距離法辨識流程圖四

5. 歐氏距離法辨識流程圖五，以病症為比較對象。一人為測試者，其餘為分類組，測試者以其所有 ECG 特徵值辨識，如圖 5.11。

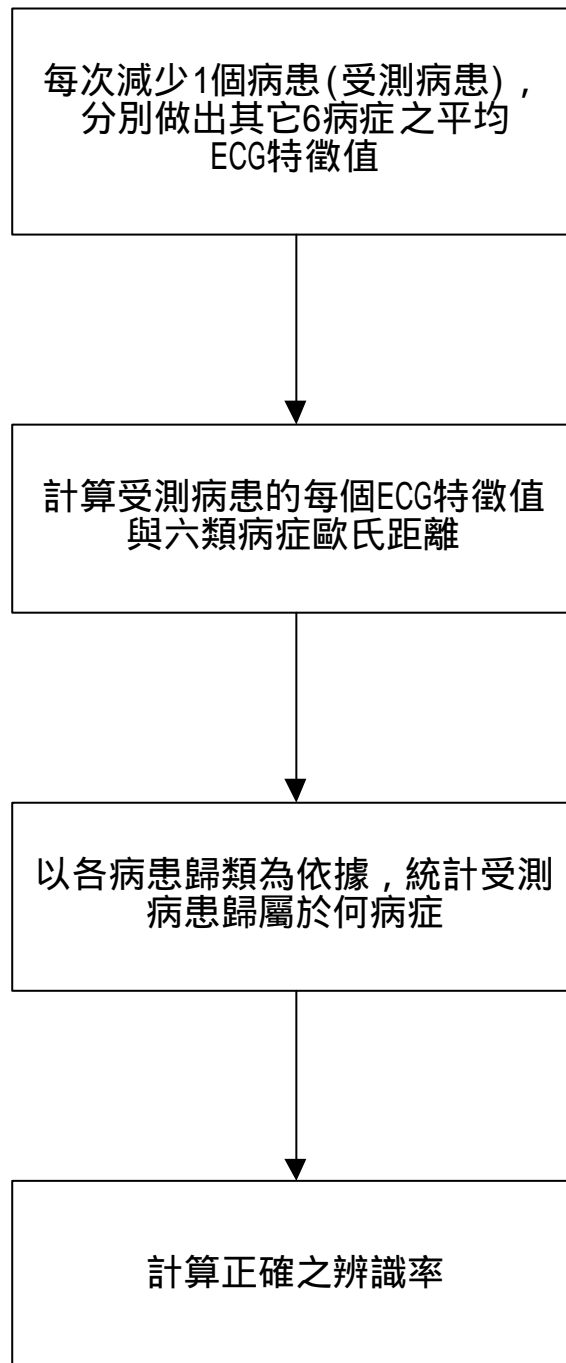


圖 5.11 歐氏距離法辨識流程圖五

6. 歐氏距離法辨識流程圖六，以病患為比較對象。一人為測試者，其餘為分類組，測試者以其所有 ECG 特徵值辨識，如圖 5.12。

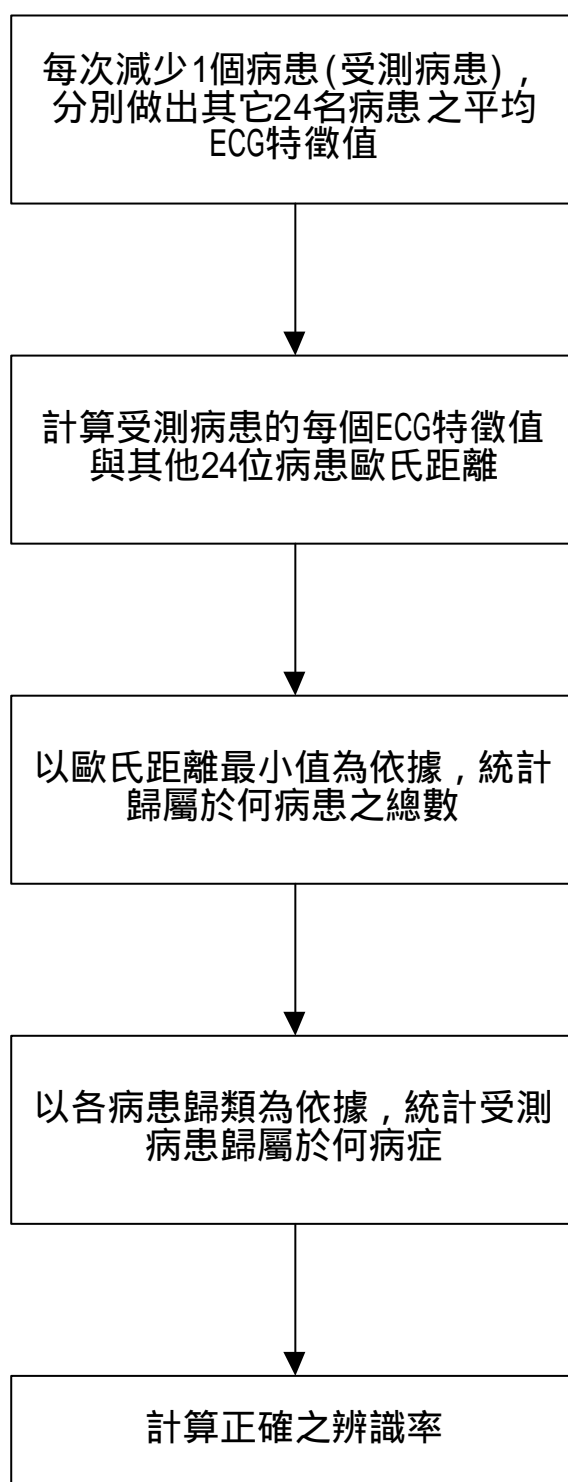


圖 5.12 歐氏距離法辨識流程圖六

7. 區別分析法辨識流程圖一，分為分類組及測試組。如圖 5.13。

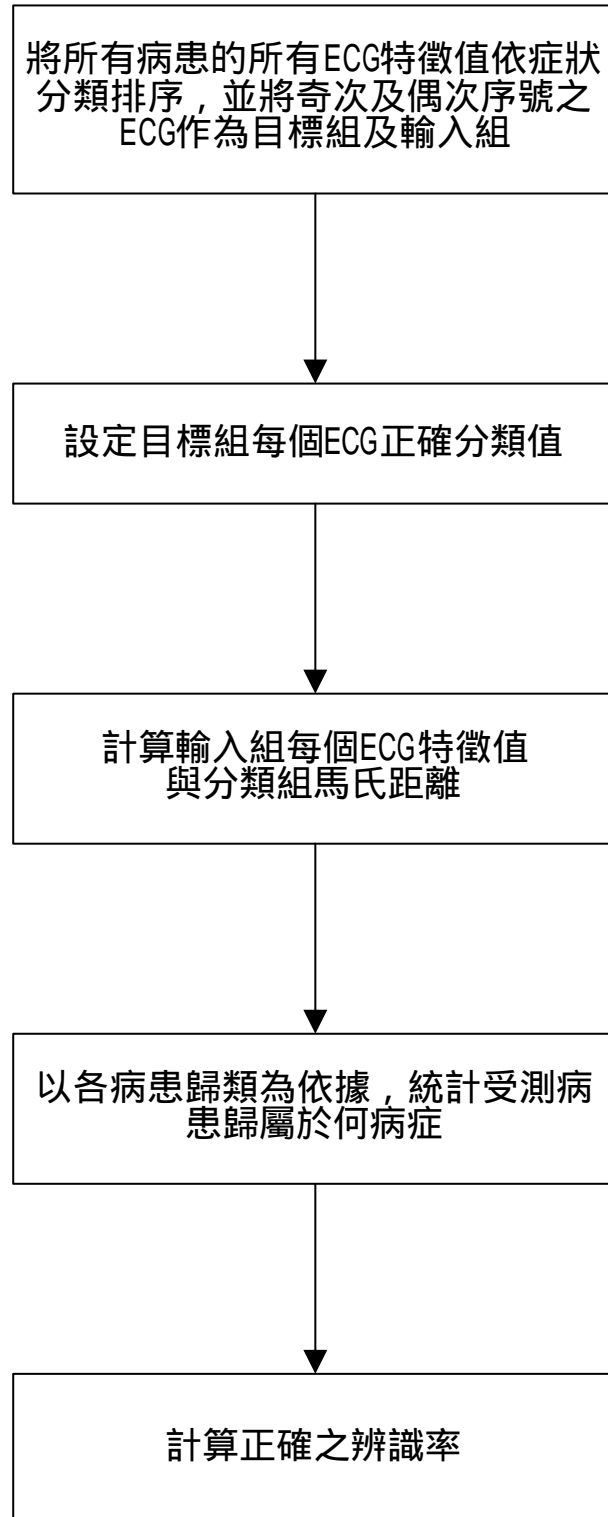


圖 5.13 區別分析法辨識流程圖一

8. 區別分析法辨識流程圖二，一人為測試者其餘為分類組，以測試者其平均 ECG 特徵值分類。如圖 5.14。

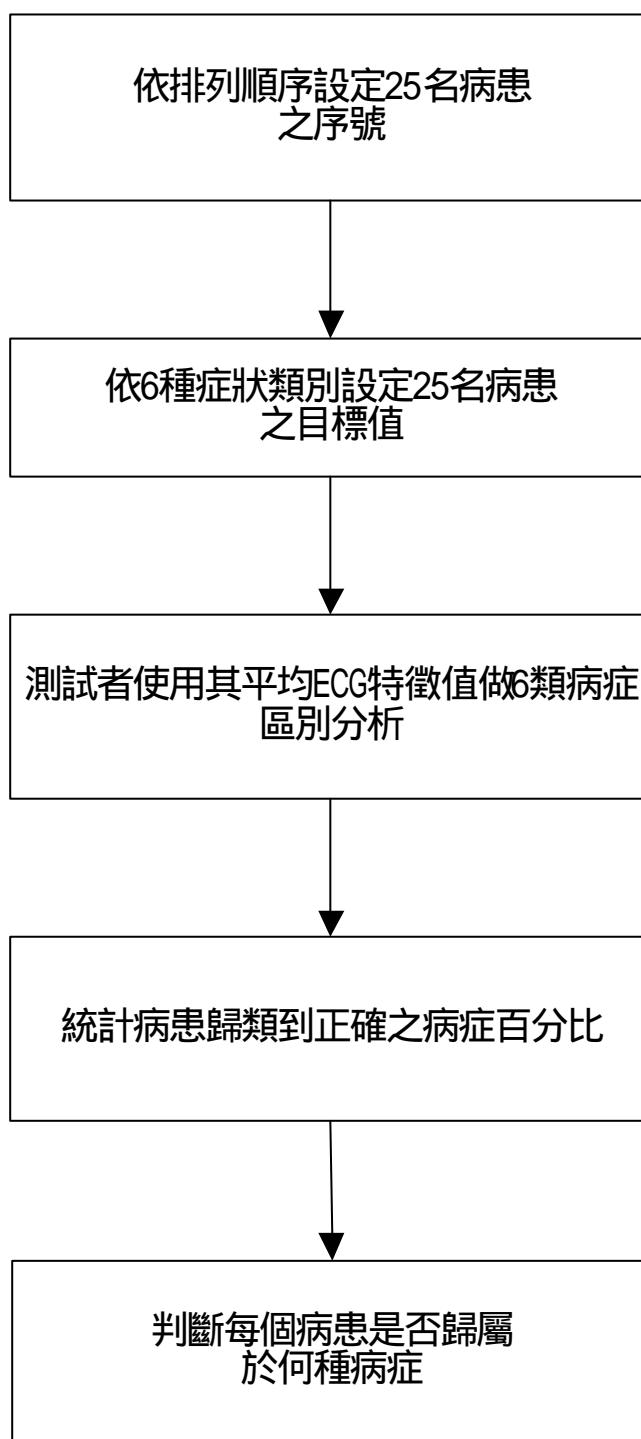


圖 5.14 區別分析法辨識流程圖二

9. 區別分析法辨識流程圖三，一人為測試者其餘為分類組，以測試者其所有 ECG 特徵值分類。如圖 5.15。

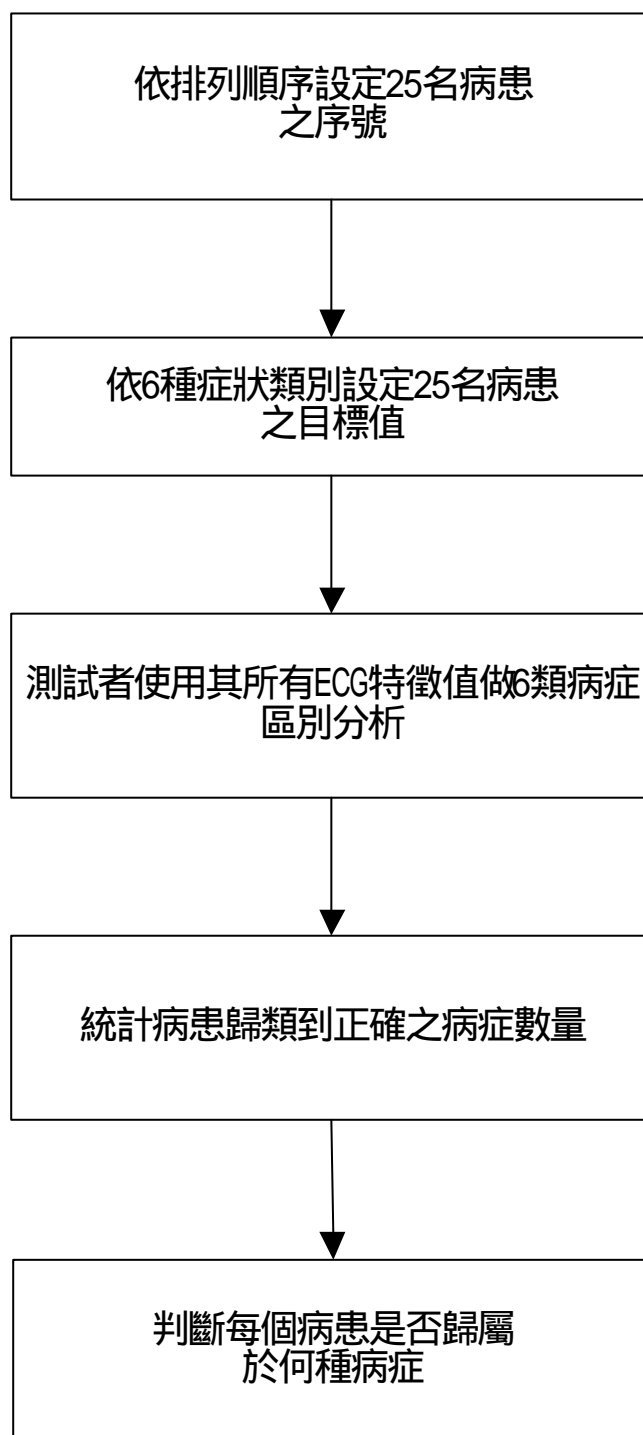


圖 5.15 區別分析法辨識流程圖三

10. 倒傳遞網路法辨識流程圖一，分為分類組及測試組。如圖 5.16。

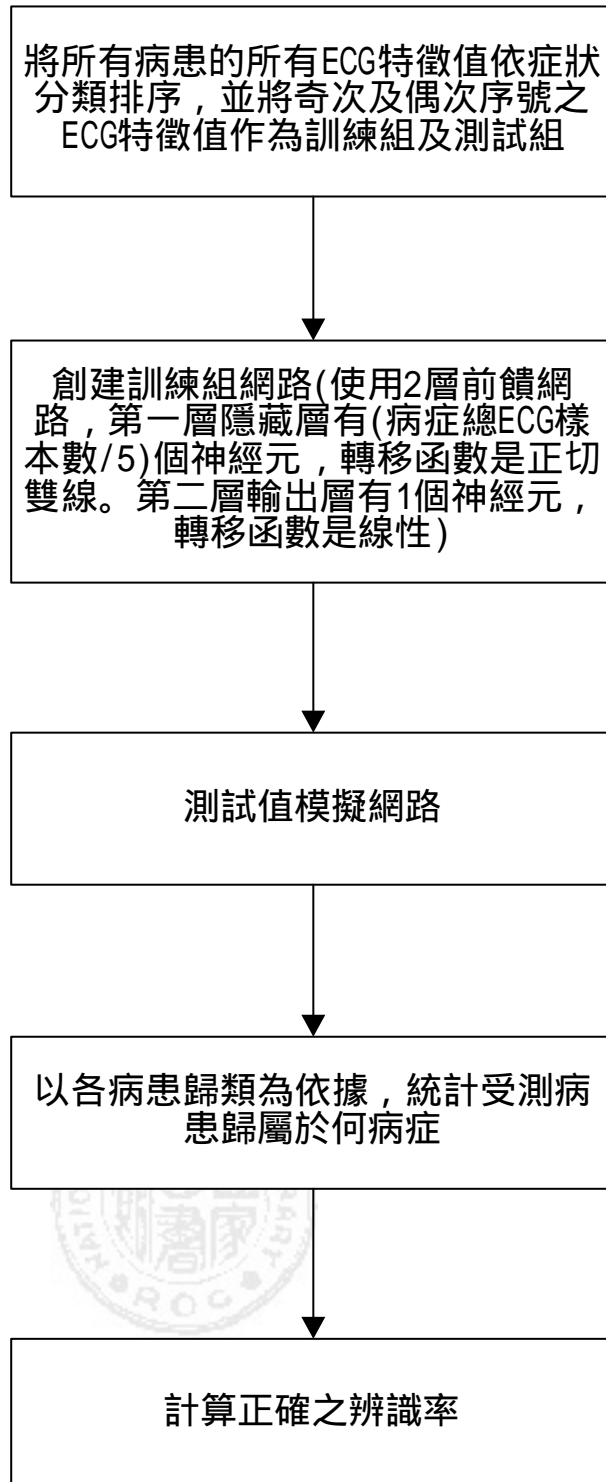


圖 5.16 倒傳遞網路法辨識流程圖一

11. 倒傳遞網路法辨識流程圖二，以每位病患其平均 ECG 特徵值做辨識，如圖 5.17 所示。

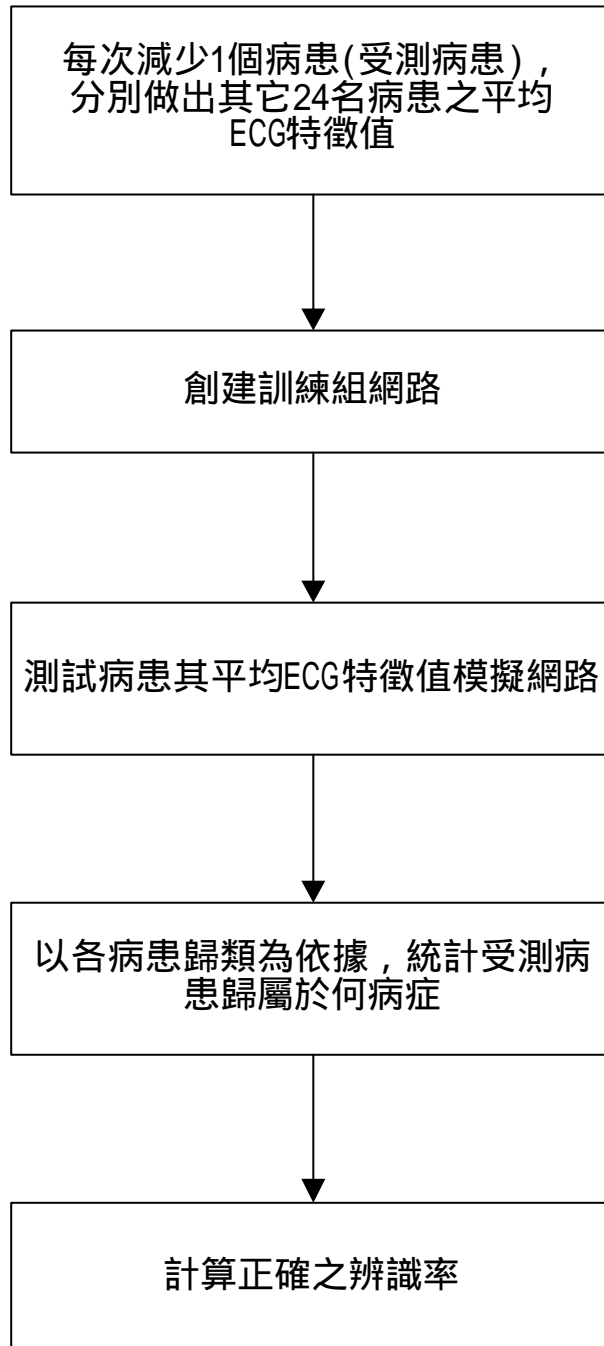


圖 5.17 倒傳遞網路法辨識流程圖二

12. 倒傳遞網路法辨識流程圖三，將每位病患的所有 ECG 特徵值逐一模擬歸類，再取歸類後最大值為歸類病症，如圖 5.18 所示。

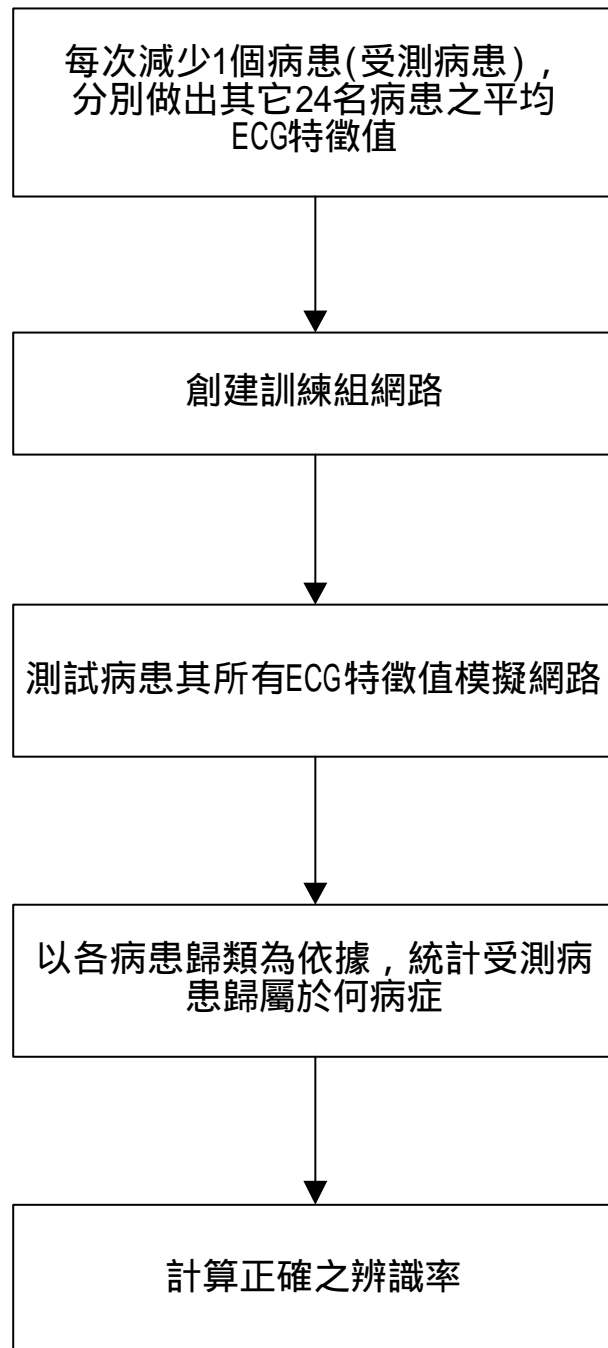


圖 5.18 倒傳遞網路法辨識流程圖三

13. 徑向基網路法辨識流程圖一，分為分類組及測試組。如圖 5.19。

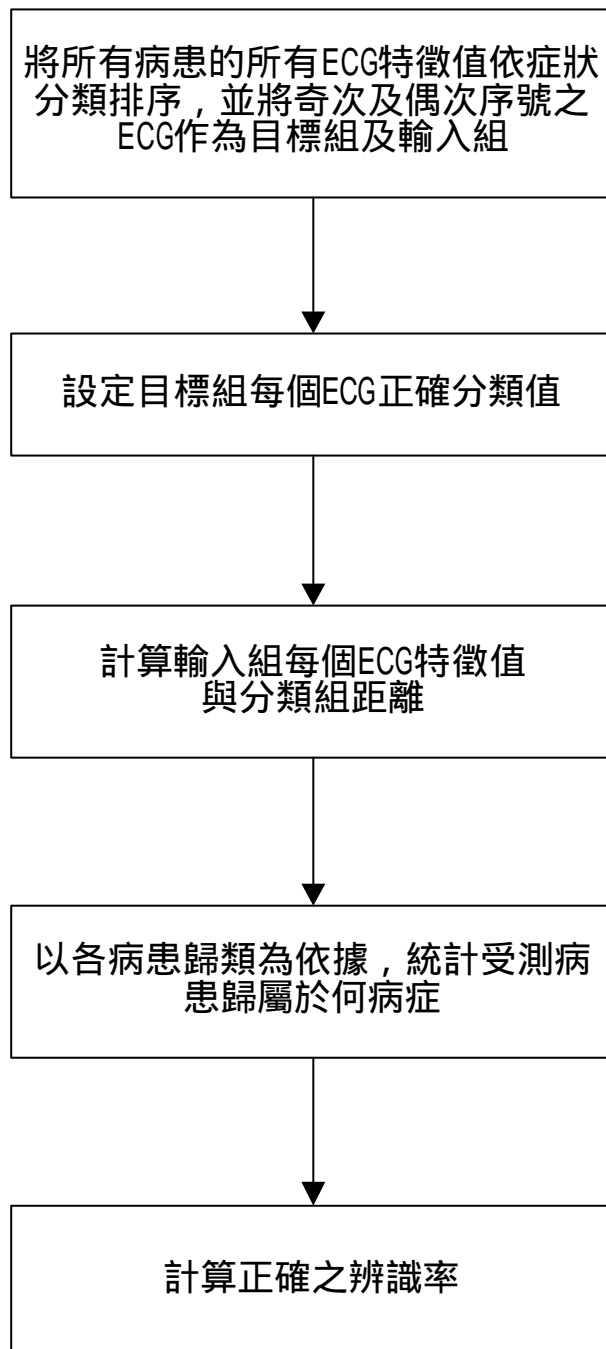


圖 5.19 徑向基網路法辨識流程圖一

14. 徑向基網路法辨識流程圖二，以每位病患其平均 ECG 特徵值做辨識，如圖 5.20 所示。

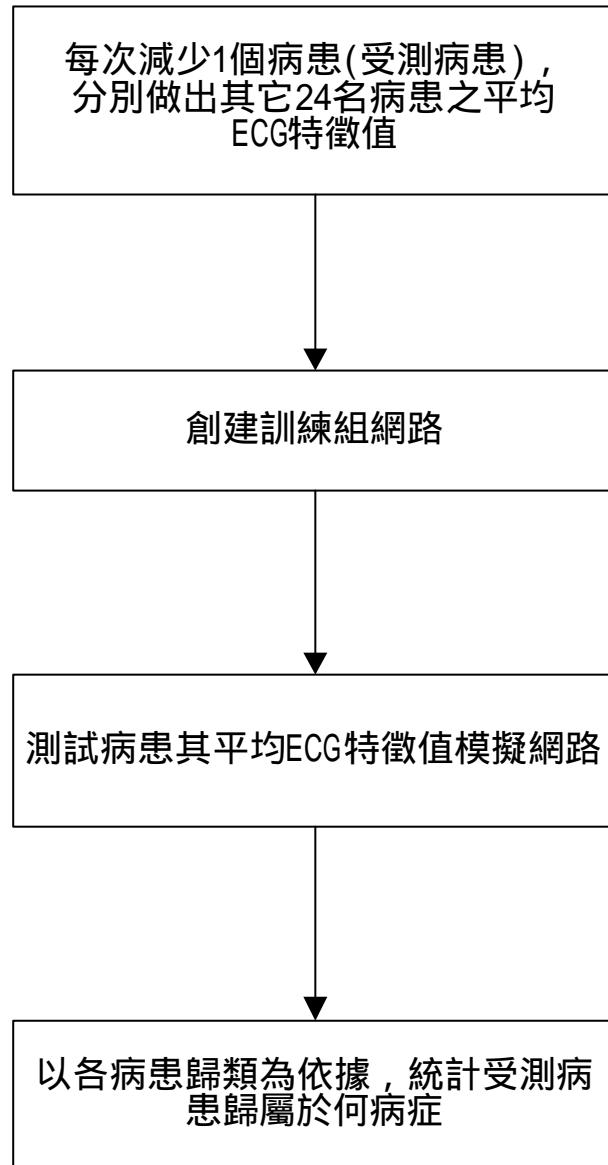


圖 5.20 徑向基網路法辨識流程圖二

15. 徑向基網路法辨識流程圖三，以每位病患其平均 ECG 特徵值做辨識，如圖 5.21 所示。

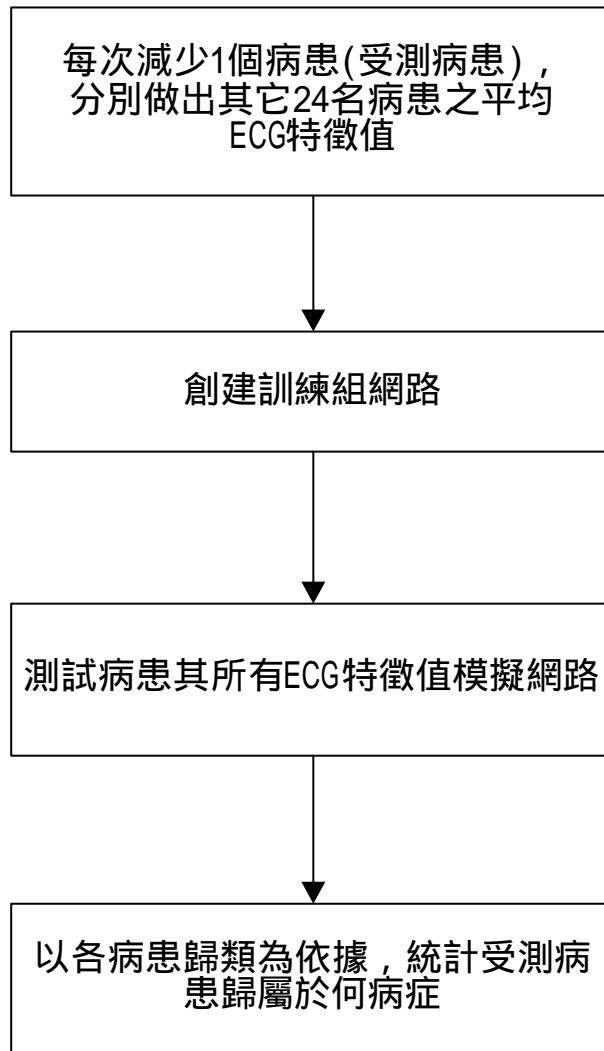


圖 5.21 徑向基網路法辨識流程圖三

第六章 研究結果

本章依不同之辨識方式將研究數據列表，以供比對。本章各表內符號『』，表示辨識正確。

6.1 歐氏距離法辨識結果

歐氏距離法有六種實驗方式，其結果分別如表 6.1 至表 6.6 所示。

表 6.1 歐氏距離識別法一
(分為分類組及測試組，以病症為比較對象)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	178	0	82	0	40	0	NORM	59.3%	
	113	87	0	17	0	196	0	APB	29.0%	
	115	283	0	0	17	0	0	NORM	94.3%	
	123	240	0	0	60	0	0	NORM	80.0%	
	220	271	0	0	26	3	0	NORM	90.3%	
	234	1	1	115	0	183	0	APB	0.3%	
LBBB	109	0	300	0	0	0	0	LBBB	100.0%	
	111	4	73	142	1	71	9	RBBB	24.3%	
	207	9	242	2	38	9	0	LBBB	80.7%	
	214	83	2	151	2	62	0	RBBB	0.7%	
RBBB	118	100	1	22	138	38	1	PVC	7.3%	
	124	3	170	116	2	0	9	LBBB	38.7%	
	212	1	0	0	0	299	0	APB	0.0%	
	231	0	0	185	0	115	0	RBBB	61.7%	
PVC	119	0	15	0	85	0	0	PVC	85.0%	
	221	4	12	22	3	34	0	APB	4.0%	
	233	8	81	1	110	0	0	PVC	55.0%	
	200	18	30	2	150	0	0	PVC	75.0%	
APB	209	30	0	0	2	43	0	APB	57.3%	
	222	0	2	0	0	42	6	APB	84.0%	
	232	57	3	0	22	218	0	APB	72.7%	
Paced Beat	102	0	0	0	1	86	213	Paced Beat	71.0%	
	107	0	0	0	0	0	300	Paced Beat	100.0%	
	217	0	2	6	0	1	291	Paced Beat	97.0%	
	104	0	2	1	0	4	293	Paced Beat	97.7%	

表 6.2 歐氏距離識別法二

(分為分類組及測試組，以病患為比較對象)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	121	167	0	0	12	0	LBBB	40.3%	
	113	112	15	172	0	1	0	RBBB	37.3%	
	115	284	4	12	0	0	0	NORM	94.7%	
	123	293	3	3	1	0	0	NORM	97.7%	
	220	292	0	6	0	2	0	NORM	97.3%	
	234	295	4	0	0	1	0	NORM	98.3%	
LBBB	109	0	3	0	297	0	0	PVC	1.0%	
	111	26	100	117	10	20	27	RBBB	33.3%	
	207	18	14	5	256	7	0	PVC	4.7%	
	214	147	8	91	0	54	0	NORM	2.7%	
RBBB	118	184	11	43	5	56	1	NORM	14.3%	
	124	240	26	21	13	0	0	NORM	7.0%	
	212	117	1	7	0	175	0	APB	2.3%	
	231	60	1	41	0	198	0	APB	13.7%	
PVC	119	0	98	1	1	0	0	LBBB	1.0%	
	221	7	13	19	1	35	0	APB	1.3%	
	233	52	20	39	89	0	0	PVC	44.5%	
	200	28	126	22	24	0	0	LBBB	12.0%	
APB	209	41	0	21	0	13	0	NORM	17.3%	
	222	1	2	3	40	4	0	PVC	8.0%	
	232	96	3	78	0	123	0	APB	41.0%	
Paced Beat	102	1	1	3	15	67	213	Paced Beat	71.0%	
	107	0	0	0	0	0	300	Paced Beat	100.0%	
	217	8	26	5	5	45	211	Paced Beat	70.3%	
	104	1	3	2	1	30	263	Paced Beat	87.7%	

表 6.3 歐氏距離識別法三

(以病症為比較對象。一人為測試者，其餘為分類組，測試者以其平均 ECG 特徵值辨識)

病症	資料檔編號	辨識結果	備註
NORM	103	NORM	
	113	APB	
	115	RBBB	
	123	PVC	
	220	RBBB	
	234	APB	
LBBB	109	PVC	
	111	RBBB	
	207	PVC	
	214	RBBB	
RBBB	118	NORM	
	124	LBBB	
	212	APB	
	231	APB	
PVC	119	LBBB	
	221	APB	
	233	LBBB	
	200	LBBB	
APB	209	NORM	
	222	RBBB	
	232	NORM	
Paced Beat	102	Paced Beat	
	107	Paced Beat	
	217	Paced Beat	
	104	APB	

表 6.4 歐氏距離識別法四

(以病患為比較對象。一人為測試者，其餘為分類組，測試者以其平均 ECG 特徵值辨識)

病症	資料檔編號	辨識結果	備註
NORM	103	LBBB	
	113	RBBB	
	115	NORM	
	123	NORM	
	220	NORM	
	234	NORM	
LBBB	109	RBBB	
	111	RBBB	
	207	PVC	
	214	NORM	
RBBB	118	RBBB	
	124	LBBB	
	212	PVC	
	231	APB	
PVC	119	LBBB	
	221	APB	
	233	PVC	
	200	LBBB	
APB	209	RBBB	
	222	RBBB	
	232	APB	
Paced Beat	102	Paced Beat	
	107	Paced Beat	
	217	Paced Beat	
	104	Paced Beat	

表 6.5 歐氏距離識別法五

(以病症為比較對象。一人為測試者，其餘為分類組，測試者以其所有 ECG 特徵值辨識)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	407	0	93	0	100	0	NORM	67.8%	
	113	180	0	26	0	393	1	APB	30.0%	
	115	547	0	1	52	0	0	NORM	91.2%	
	123	455	0	0	145	0	0	NORM	75.8%	
	220	516	0	0	78	6	0	NORM	86.0%	
	234	2	3	217	0	378	0	APB	0.3%	
LBBB	109	0	600	0	0	0	0	LBBB	100.0%	
	111	6	173	275	0	128	18	RBBB	28.8%	
	207	18	332	4	223	23	0	LBBB	55.3%	
	214	184	16	278	3	119	0	RBBB	2.7%	
RBBB	118	196	2	37	273	91	1	PVC	6.2%	
	124	5	490	88	0	0	17	LBBB	14.7%	
	212	6	0	1	0	593	0	APB	0.2%	
	231	0	0	411	0	189	0	RBBB	68.5%	
PVC	119	0	41	0	159	0	0	PVC	79.5%	
	221	13	24	37	8	68	0	APB	5.3%	
	233	20	162	2	216	0	0	PVC	54.0%	
	200	31	29	7	331	2	0	PVC	82.8%	
APB	209	69	0	2	4	75	0	APB	50.0%	
	222	0	2	1	0	87	10	APB	87.0%	
	232	132	2	1	43	422	0	APB	70.3%	
Paced Beat	102	0	0	0	1	173	426	Paced Beat	71.0%	
	107	0	0	0	0	0	600	Paced Beat	100.0%	
	217	0	4	11	0	1	584	Paced Beat	97.3%	
	104	0	3	4	0	8	585	Paced Beat	97.5%	

表 6.6 歐氏距離識別法六

(以病患為比較對象。一人為測試者，其餘為分類組，測試者以其所有 ECG 特徵值辨識)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	234	346	1	0	19	0	LBBB	39.0%	
	113	263	30	300	0	6	1	RBBB	43.8%	
	115	561	11	26	0	2	0	NORM	93.5%	
	123	581	6	12	1	0	0	NORM	96.8%	
	220	580	2	8	0	10	0	NORM	96.7%	
	234	585	12	0	0	3	0	NORM	97.5%	
LBBB	109	0	6	1	593	0	0	PVC	1.0%	
	111	44	202	257	11	37	49	RBBB	33.7%	
	207	44	21	13	507	15	0	PVC	3.5%	
	214	311	33	154	0	102	0	NORM	5.5%	
RBBB	118	369	10	93	15	112	1	NORM	15.5%	
	124	273	276	19	32	0	0	LBBB	3.2%	
	212	221	0	17	0	362	0	APB	2.8%	
	231	143	11	78	0	368	0	APB	13.0%	
PVC	119	0	186	0	14	0	0	LBBB	7.0%	
	221	20	23	33	3	71	0	APB	2.0%	
	233	28	10	38	74	0	0	PVC	49.3%	
	200	44	267	48	38	3	0	LBBB	9.5%	
APB	209	93	1	29	0	27	0	NORM	18.0 %	
	222	3	11	2	78	5	1	PVC	5.0%	
	232	207	4	169	0	220	0	APB	36.7%	
Paced Beat	102	2	42	4	3	138	411	Paced Beat	68.5%	
	107	0	0	0	0	0	600	Paced Beat	100.0%	
	217	22	37	10	30	64	437	Paced Beat	72.8%	
	104	6	4	3	0	61	526	Paced Beat	87.7%	

6.2 區別分析法辨識結果

區別分析法辨識結果如表 6.7、6.8 及 6.9 所示。

表 6.7 區別分析法辨識流程圖一
(分為分類組及測試組)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	293	0	7	0	0	0	NORM	97.7%	
	113	295	0	5	0	0	0	NORM	98.3%	
	115	300	0	0	0	0	0	NORM	100.0%	
	123	294	0	6	0	0	0	NORM	98.0%	
	220	297	0	3	0	0	0	NORM	99.0%	
	234	1	0	20	0	279	0	APB	0.3%	
LBBB	109	0	299	0	0	1	0	LBBB	99.7%	
	111	1	199	0	0	97	3	LBBB	66.3%	
	207	0	245	2	3	34	16	LBBB	81.7%	
	214	0	29	121	107	42	1	RBBB	9.7%	
RBBB	118	1	0	147	141	10	1	RBBB	49.0%	
	124	0	84	216	0	0	0	RBBB	72.0%	
	212	1	1	232	0	66	0	RBBB	77.3%	
	231	8	0	291	0	1	0	RBBB	97.0%	
PVC	119	0	0	0	100	0	0	PVC	100.0%	
	221	0	51	0	4	20	0	LBBB	5.3%	
	233	0	8	1	191	0	0	PVC	95.5%	
	200	0	30	1	163	4	2	PVC	81.5%	
APB	209	26	0	19	0	30	0	APB	40.0%	
	222	0	10	10	0	16	14	APB	32.0%	
	232	0	3	0	0	297	0	APB	99.0%	
Paced Beat	102	0	0	4	0	0	296	Paced Beat	98.7%	
	107	0	0	0	0	0	300	Paced Beat	100.0%	
	217	0	2	0	1	0	297	Paced Beat	99.0%	
	104	1	0	1	0	2	296	Paced Beat	98.7%	

表 6.8 區別分析法辨識流程圖二

(一人為測試者其餘為分類組，以測試者其平均 ECG 特徵值分類)

病症	資料檔編號	辨識結果	備註
NORM	103	NORM	
	113	NORM	
	115	NORM	
	123	NORM	
	220	NORM	
	234	APB	
LBBB	109	APB	
	111	APB	
	207	PVC	
	214	RBBB	
RBBB	118	PVC	
	124	LBBB	
	212	APB	
	231	NORM	
PVC	119	PVC	
	221	APB	
	233	NORM	
	200	LBBB	
APB	209	NORM	
	222	LBBB	
	232	LBBB	
Paced Beat	102	Paced Beat	
	107	Paced Beat	
	217	Paced Beat	
	104	Paced Beat	

表 6.9 區別分析法辨識流程圖三

(一人為測試者其餘為分類組，以測試者其所有 ECG 特徵值分類)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	445	0	3	150	2	0	NORM	74.2%	
	113	467	1	7	122	3	0	NORM	77.8%	
	115	417	0	0	183	0	0	NORM	69.5%	
	123	332	0	42	226	0	0	NORM	55.3%	
	220	286	0	171	142	0	1	NORM	47.7%	
	234	0	154	1	215	229	1	APB	0%	
LBBB	109	162	4	1	131	302	0	APB	0.7%	
	111	2	290	0	96	69	143	LBBB	48.3%	
	207	16	156	40	305	11	72	PVC	26.0%	
	214	36	3	245	300	2	14	PVC	0.5%	
RBBB	118	95	1	207	284	12	1	PVC	34.5%	
	124	86	117	133	130	0	134	Paced Beat	22.2%	
	212	130	2	75	246	133	14	PVC	12.5%	
	231	328	0	0	140	132	0	NORM	0.0%	
PVC	119	0	0	64	100	0	36	PVC	50.0%	
	221	3	31	28	56	13	19	PVC	37.3%	
	233	155	1	0	5	0	239	Paced Beat	1.3%	
	200	0	282	2	115	0	1	LBBB	28.8%	
APB	209	76	0	5	49	12	8	NORM	8.0%	
	222	0	5	12	1	29	53	Paced Beat	29.0%	
	232	28	146	112	234	79	1	PVC	13.2%	
Paced Beat	102	0	41	2	0	2	555	Paced Beat	92.5%	
	107	34	0	112	8	0	446	Paced Beat	74.3%	
	217	18	22	199	170	1	190	RBBB	31.7%	
	104	2	10	159	38	46	345	Paced Beat	57.5%	

6.3 倒傳遞網路法辨識結果

倒傳遞網路法如表 6.10、 6.11 及 6.12 所示。

表 6.10 倒傳遞網路法辨識結果一
(分為分類組及測試組)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	300	0	0	0	0	0	LBBB	100.0%	
	113	297	2	1	0	0	0	NORM	99.0%	
	115	298	1	0	0	0	0	APB	99.7%	
	123	298	2	0	0	0	0	LBBB	99.3%	
	220	298	1	1	0	0	0	NORM	99.3%	
	234	295	3	1	0	0	0	LBBB	98.7%	
LBBB	109	0	295	4	0	0	0	RBBB	98.7%	
	111	11	269	16	2	0	1	RBBB	90.0%	
	207	7	276	10	2	0	0	PVC	93.6%	
	214	0	289	10	0	1	0	RBBB	96.3%	
RBBB	118	1	2	290	7	0	0	NORM	96.7%	
	124	0	5	293	1	0	0	NORM	98.0%	
	212	0	8	282	7	3	0	NORM	94.0%	
	231	0	4	295	0	1	0	LBBB	98.3%	
PVC	119	0	0	0	98	2	0	NORM	98.0%	
	221	2	3	3	67	0	0	APB	89.3%	
	233	0	0	6	193	0	0	RBBB	97.0%	
	200	1	0	16	175	7	1	NORM	87.5%	
APB	209	0	0	1	7	67	0	RBBB	89.3%	
	222	0	0	2	6	26	14	RBBB	54.2%	
	232	0	1	2	2	292	3	LBBB	97.3%	
Paced Beat	102	0	1	0	1	3	295	LBBB	98.3%	
	107	0	0	0	0	1	299	APB	99.7%	
	217	0	0	0	1	6	291	Paced Beat	97.7%	
	104	0	0	0	1	5	290	PVC	98.0%	

表 6.11 倒傳遞網路法辨識結果二
(以每位病患其平均 ECG 特徵值做辨識)

資料檔編號		辨識結果	備註
NORM	103	NORM	
	113	NORM	
	115	NORM	
	123	NORM	
	220	LBBB	
	234	RBBB	
LBBB	109	LBBB	
	111	APB	
	207	APB	
	214	LBBB	
RBBB	118	LBBB	
	124	NORM	
	212	PVC	
	231	LBBB	
PVC	119	APB	
	221	NORM	
	233	LBBB	
	200	RBBB	
APB	209	LBBB	
	222	Paced Beat	
	232	Paced Beat	
Paced Beat	102	RBBB	
	107		
	217	PVC	
	104		

表 6.12 倒傳遞網路法辨識結果三

(將每位病患的所有 ECG 特徵值逐一模擬歸類，再取歸類後最大值
為歸類病症)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正 確率	備 註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	238	61	4	0	0	0	NORM	39.7%	
	113	139	146	106	81	71	43	LBBB	23.2%	
	115	48	39	135	70	19	4	RBBB	8.0%	
	123	296	126	146	27	5	0	NORM	49.3%	
	220	213	282	42	5	0	0	LBBB	35.5%	
	234	55	81	72	104	51	23	PVC	9.2%	
LBBB	109	96	92	0	0	0	0	NORM	15.3%	
	111	40	94	66	73	58	42	LBBB	15.7%	
	207	31	22	27	31	44	125	Paced Beat	3.7%	
	214	218	193	93	44	25	9	NORM	32.2%	
RBBB	118	88	191	198	96	3	0	RBBB	33.0%	
	124	166	126	170	48	2	0	RBBB	28.3%	
	212	117	40	18	50	133	87	APB	3.0%	
	231	67	70	56	63	57	62	LBBB	9.3%	
PVC	119	19	7	20	18	28	97	Paced Beat	9.0%	
	221	33	14	17	4	1	0	NORM	2.7%	
	233	52	30	31	89	38	7	PVC	22.3%	
	200	71	128	82	75	23	8	LBBB	18.8%	
APB	209	18	67	39	13	5	6	LBBB	3.3%	
	222	4	11	2	5	7	11	LBBB	7.0%	
	232	24	71	189	117	29	2	RBBB	4.8%	
Paced Beat	102	16	23	50	146	219	143	PVC	23.8%	
	107	4	27	54	99	64	272	Paced Beat	45.3%	
	217	18	67	183	112	89	70	RBBB	11.7%	
	104	1	2	10	46	87	191	Paced Beat	31.8%	

6.4 徑向基網路法辨識結果

徑向基網路辨識結果如表 6.13、6.14 及 6.15.所示。

表 6.13 徑向基網路辨識結果一
(分為分類組及測試組)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	300	0	0	0	0	0	NORM	100.0%	
	113	296	1	3	0	0	0	NORM	98.7%	
	115	300	0	0	0	0	0	NORM	100.0%	
	123	299	0	1	0	0	0	NORM	99.7%	
	220	300	0	0	0	0	0	NORM	100.0%	
	234	298	0	1	0	1	0	NORM	99.3%	
LBBB	109	0	299	0	1	0	0	LBBB	99.7%	
	111	2	293	0	2	2	1	LBBB	97.7%	
	207	1	290	0	8	0	1	LBBB	96.7%	
	214	6	270	18	1	5	0	LBBB	90.0%	
RBBB	118	3	2	293	2	0	0	RBBB	97.7%	
	124	1	4	293	2	0	0	RBBB	97.7%	
	212	1	1	291	0	6	1	RBBB	97.0%	
	231	0	1	299	0	0	0	RBBB	99.7%	
PVC	119	0	0	0	100	0	0	PVC	100.0%	
	221	1	5	3	49	17	0	PVC	65.3%	
	233	0	0	0	200	0	0	PVC	100.0%	
	200	2	5	4	187	2	0	PVC	93.5%	
APB	209	5	1	12	1	56	0	APB	74.7%	
	222	1	5	0	0	41	3	APB	82.0%	
	232	0	3	0	1	296	0	APB	98.7%	
Paced Beat	102	0	1	0	0	1	298	Paced Beat	99.3%	
	107	0	0	0	0	0	300	Paced Beat	100.0%	
	217	0	1	0	0	0	299	Paced Beat	99.7%	
	104	0	3	1	0	0	296	Paced Beat	98.7%	

表 6.14 徑向基網路辨識結果二
(以每位病患其平均 ECG 特徵值做辨識)

資料檔編號		辨識結果	備註
NORM	103	LBBB	
	113	RBBB	
	115	NORM	
	123	NORM	
	220	NORM	
	234	NORM	
LBBB	109	APB	
	111	RBBB	
	207	APB	
	214	NORM	
RBBB	118	LBBB	
	124	LBBB	
	212	APB	
	231	APB	
PVC	119	LBBB	
	221	APB	
	233	APB	
	200	LBBB	
APB	209	RBBB	
	222	RBBB	
	232	APB	
Paced Beat	102	Paced Beat	
	107	Paced Beat	
	217	Paced Beat	
	104	Paced Beat	

表 6.15 徑向基網路辨識結果三
(以每位病患其平均 ECG 特徵值做辨識)

資料檔編號		統計歸類病症之總數						辨識結果	辨識正 確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	234	346	1	0	19	0	LBBB	39.0%	
	113	263	30	300	0	6	1	RBBB	43.8%	
	115	561	11	26	0	2	0	NORM	93.5%	
	123	581	6	12	1	0	0	NORM	96.8%	
	220	581	2	7	0	10	0	NORM	96.8%	
	234	586	11	0	0	3	0	NORM	97.7%	
LBBB	109	1	5	1	593	0	0	PVC	0.8%	
	111	86	160	257	11	37	49	RBBB	26.7%	
	207	93	15	13	464	15	0	PVC	2.5%	
	214	327	16	154	0	103	0	NORM	2.7%	
RBBB	118	379	69	24	15	112	1	NORM	4.0%	
	124	276	292	0	32	0	0	LBBB	0.0%	
	212	221	10	7	0	362	0	APB	1.2%	
	231	144	88	0	0	368	0	APB	0.0%	
PVC	119	0	186	0	14	0	0	LBBB	7.0%	
	221	28	27	21	3	71	0	APB	2.0%	
	233	89	43	106	162	0	0	PVC	40.5%	
	200	52	269	38	38	3	0	LBBB	9.5%	
APB	209	94	4	25	0	27	0	NORM	18.0%	
	222	5	9	80	0	5	1	RBBB	5.0%	
	232	208	68	104	6	214	0	APB	35.7%	
Paced Beat	102	2	42	7	1	137	411	Paced Beat	68.5%	
	107	3	0	0	0	0	597	Paced Beat	99.5%	
	217	31	28	40	0	64	437	Paced Beat	72.8%	
	104	7	3	3	0	61	526	Paced Beat	87.7%	

第七章 討 論

7.1 研究資料之分析

本研究對心電圖辨識可分為兩部份，一為利用小波分析擷取特徵值，另一部分是嘗試用不同的方式進行分類，可針對正常(Norm)，左束分支阻塞(LBBB)，右束分支阻塞(RBBB)，心室過早收縮(PVC)，心房過早收縮(APB)，及 Pace Beat 等六種病症做病症辨識。

本研究是將測試病患 ECG 特徵值有否加入訓練組分類裡做了比較性的實驗研究。若是採用將病患所有 ECG 特徵值分為兩部份，一作為輸入，另一作為訓練目標，便可得到較高的辨識率，如表 7.1 所示。

表 7.1 將測試病患 ECG 特徵值加入訓練分類組辨識方式結果

分類識別方式	辨識率	時間
歐氏距離識別法(方式 1)	58.56%	333sec
歐氏距離識別法(方式 2)	44.93%	338sec
區別分析識別法(方式 2)	80.13%	331sec
倒傳遞網路識別法(方式 2)	96.38%	4931sec
徑向基網路識別法(方式 2)	97.49%	335sec

方式 1：所有 ECG 特徵值對 6 種病症為比較

方式 2：所有 ECG 特徵值對 25 位病患為比較

表 7.1 所使用信號的皆是以所有病患之所有 ECG 特徵值做出的分類識別，而非使用病患平均 ECG 特徵值，這是為了更精準測試避免誤差。

若是採用的分類方式是將測試病患單一獨立出來，以不影響其他病患做訓練分類，其結果做整理及比較，如表 7.2 所示。

由表 7.1 及表 7.2 可以明顯觀察出來，若有測試者信號加入分類訓練，則辨識率將會提升的有歐氏距離識別法(所有 ECG 特徵值對 25

表 7.2 測試病患獨立後各分類辨識方式結果

分類識別方式		識別率	辨識所需時間
歐氏距離識別法	方法 1	16%	331sec
	方式 2	44%	331sec
	方式 3	68%	336sec
	方法 4	40%	360sec
區別分析識別法	方法 1	40%	332sec
	方法 4	44%	383sec
倒傳遞網路識別法	方法 1	24%	2447sec
	方法 4	21%	2243sec
徑向基網路識別法	方法 1	36%	337sec
	方法 4	40%	417sec

方式 1：平均 ECG 特徵值對 6 種病症為比較

方式 2：平均 ECG 特徵值對 25 位病患為比較

方式 3：所有 ECG 特徵值對 6 種病症為比較

方法 4：所有 ECG 特徵值對 25 位病患為比較

位病患為比較)、區別分析識別法(所有 ECG 特徵值對 25 位病患為比較)、倒傳遞網路識別法(所有 ECG 特徵值對 25 位病患為比較)及徑向基網路識別法(所有 ECG 特徵值對 25 位病患為比較)。徑向基網路識別法有 97.5%的識別正確率,倒傳遞網路識別法則有 96.38%的識別正確率。

由此可知,使用徑向基網路識別法及倒傳遞網路識別法,若將原先信號加入分類組信號內,影響辨識率的程度將會很大,可能原因是提供訓練之 ECG 信號多達樣本之一半數量,徑向基網路識別法及倒傳遞網路識別法屬於非線性網路創建其方程式,故在創建兩種網路過程中加入了屬於該病患的特徵值,模擬時便會有較高的識別能力。

至於倒傳遞網路識別法在表 7.2 中,發覺辨識及分類能力降低到 21%及 24%,未能發現其有較高的辨識率,應屬於創見網路時,其神經元數量及隱藏層層數決定多少之結果。但本研究過程中,相同程式

在執行時，其執行結果每次都不相同，即使使用了自訂參數、相同神經元數量或是隱藏層數目，其每次結果仍是無法穩定相同。因此判斷倒傳遞網路識別法辨識率差的原因應是與參數設定、神經元數量及隱藏層數目等等因素有關，並非僅能有最高的 32% 的辨識率。

7.2 與已發表之相關文獻比較

ECG 之研究可分兩部分一為特徵值擷取，另一為辨識，本研究是以小波轉換擷取特徵值，歐氏距離法、區別分析法、倒傳遞網路法以及徑向基網路法四種方式做辨識。由 7.1 節之研究分析得知將測試病患 ECG 特徵值加入訓練分類組辨識方式時，以徑向基網路識別法有最佳辨識率，25 位病患測試所有 ECG 信號其辨識率為 97.49%。表 7.3 是 Osowski 等學者及楊朝樑先生之文獻與本研究之比較。

表 7.3 相關文獻與本研究之比較

	Fhyb-HOSA[8]	PCA[3]	Wavelet Transform (本研究)	
資料來源	MIT/HIB	MIT/HIB	MIT/HIB	
使用 Neural Network	SOM + MLP	PCA	Radial Basis	
使用 Record 數	15	22	25	22
辨識之病症數	7	6	6	6
辨識率	96.06%	98%	97.49%	86.98%
訓練及辨識時間	-	大於 10 小時	335 秒	295 秒

表 7.3 中列出與 PCA 使用相同基礎之 Record 數做比較，得到相對較差的辨識率 86.98%。雖然辨識率較差，但若依照各病患 ECG 特徵值之辨識正確總數做病症分類時，依然能完全正確辨識出病患屬於何種病症，如表 7.4 所示。

表 7.4 依照各病患 ECG 特徵值之辨識正確總數做病症分類

資料檔編號		統計歸類病症之總數						辨識結果	辨識正確率	備註
		NORM	LBBB	RBBB	PVC	APB	Paced Beat			
NORM	103	300	0	0	0	0	0	NORM	100.0%	
	113	296	1	3	0	0	0	NORM	98.7%	
	115	300	0	0	0	0	0	NORM	100.0%	
	123	299	0	1	0	0	0	NORM	99.7%	
	220	300	0	0	0	0	0	NORM	100.0%	
	234	298	0	1	0	1	0	NORM	99.3%	
LBBB	109	0	299	0	1	0	0	LBBB	99.7%	
	111	2	293	0	2	2	1	LBBB	97.7%	
	207	1	293	0	5	1	0	LBBB	97.7%	
	214	6	270	18	1	5	0	LBBB	90.0%	
RBBB	118	3	2	295	0	0	0	RBBB	98.3%	
	124	1	6	293	0	0	0	RBBB	97.7%	
	212	1	1	291	0	7	0	RBBB	97.0%	
	231	0	1	299	0	0	0	RBBB	99.7%	
PVC	119	0	0	0	100	0	0	PVC	100.0%	
	221	1	5	3	49	17	0	PVC	65.3%	
APB	209	5	1	12	1	56	0	APB	74.7%	
	222	1	5	0	0	44	0	APB	88.0%	
	232	0	3	0	1	296	0	APB	98.7%	
Paced Beat	102	0	1	1	0	1	297	Paced Beat	99.0%	
	107	0	0	0	0	0	300	Paced Beat	100.0%	
	217	0	1	0	0	0	299	Paced Beat	99.7%	

表 7.4 中特別值得注意的是測試樣本數較低的病患(小於 300 個樣本數) 其辨識正確率都較低，除了 119 資料檔編號之病患。因此徑向基網路識別法中，若提供較多的訓練數量來創建其網路，將會得到較高的辨識正確率。

第八章 未來方向

在研究心電圖辨識方法中，有使用主要元素分析神經網路[3]、奇異值分解法[20]、模糊類神經網路方法[21]、疊代分類法[22]等等研究文獻，同時也有使用小波轉換為基礎在使用倒傳遞類神經網路及其模糊版本方式[23]、利用小波與類神經網路進行心電圖特徵擷取與病症分類[24]等等相類似之方式，在此我們同樣使用小波轉換擷取十個特徵值[19]，來做辨識之分析。由這些實驗方法及結果，我們提出幾點建議，以作為將來繼續研究的方向：

- 1、取不同的特徵值，不論是減少不相關之特徵值或尋找更相關之特徵值，可以再次試驗是否能提高辨識率。因為在本研究中，小波轉換後，嘗試了十個特徵值，而這些特徵值並無特定限制規定，因此應該可以再多嘗試其他特徵值，藉以改善辨識率。
- 2、除了以 QRS 波辨識外，應該也可以嘗試其它要素，如圖 8.1 標準心電圖波形所示，各個點間段距離、各點高度差或是兩個相鄰兩 ECG 之 R 波距離等等。是否與所取的訊號有關，值得繼續研究、探討。
- 3、在使用線性與非線性方式做多變量分析時，理論上非線性方式應會有較佳的結果。在本研究中線性分析方式有歐氏距離法及區別分析法兩種，非線性分析方式有倒傳遞網路法及徑向基網路法兩種。分析結果是以非線性方式的徑向基網路分類方法較佳，辨識正確率達 97.5%。但是依相同條件分類時，同樣非線性方式的倒傳遞網路分類方式，卻是只有最高的 32% 的辨識率。會有如此結果，是否與某些相關的特徵值尚未被發現，或是其他未知因素有關也是值得再繼續研究重點。

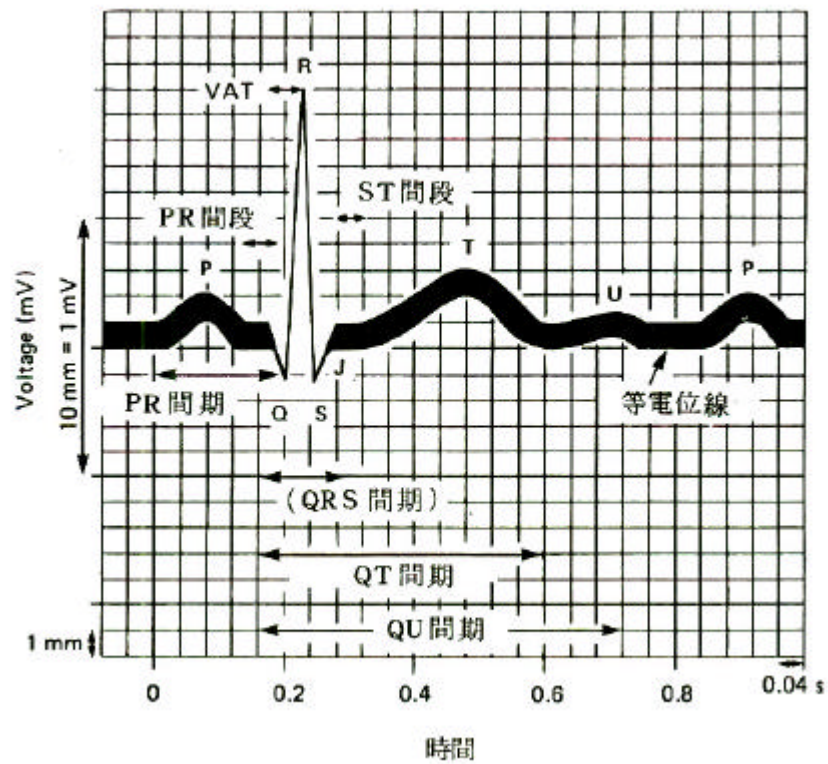


圖 8.1 標準心電圖波形

- 4、本研究是針對 ECG 波形來辨識六種病症，將來是否以此方式擴展到其他相關種類之辨識，應該也能成為另一研究方向。

參考文獻

- [1] 行政院衛生署衛生統計資訊網，<http://203.65.100.16>
- [2] 呂嘉陞編著，”心電圖學必備”第三版，合記圖書出版社，2001.
- [3] 楊朝樑，”使用主要元素分析神經網路的心電圖病症辨識系統”，中正大學碩士論文，2002.
- [4] J.Pan and W.Tompkins，”A Real-Time QRS Detection Algorithm”，IEEE Trans. Biomed. Eng.，Vol. BME-32，pp.230-236，1985.
- [5] 童淑芬，”類神經網路在心電圖分類之應用“，中原大學碩士論文，1990.
- [6] 林俊榮，”自動心電圖特徵值擷取與病症分類之研究”，中原大學碩士論文，1995.
- [7] Kei-ichiro Minami， Hiroshi Nakajima and Takeshi Toyoshima，”Real-time Discrimination of Ventricular Tachyarrhythmia with Fourier-Transform Neural Network”，IEEE Trans.on Biomed. Eng. Vol. 46. NO.2，February 1999.
- [8] S. Osowski and T.H. Linh，”ECG Beat Recognition Using Fuzzy Hybrid Neural Network“，IEEE Trans.Biomed. Eng.Vol.48.NO.11，November 2001.
- [9] 天星英漢百科醫學辭典，<http://www.geocities.com/>
- [10] 余政展等著，”醫學生理學”，合記圖書公司，1995.
- [11] John G.Webster，”Medical Instrumentation – Application and Design”，Third edition，John Wiley & Sons，Inc.1998.
- [12] <http://pthcwww.pthc.chc.edu.tw/info/輔導室/行職業資訊研發成果/001/002/a025/12.htm>

- [13] 林志龍, "生醫訊號小波壓縮之品質控制與加密演算法", 中原大學碩士論文, 2002.
- [14] 溫國棟, "智慧化臨床資訊自動篩選與分析技術研究", 中原大學碩士論文, 2002.
- [15] 陳順宇著, "多變量分析", 華泰書局, 2000.
- [16] 羅華強著, "類神經網路-MATLAB 的應用", 清蔚科技股份有限公司, 2001.
- [17] <http://www.physionet.org>.
- [18] 張智星, "Matlab 程式設計與應用", 清蔚科技, 2000.
- [19] Stefanos K. Goumas, Michael E Zervakis, Member
"Classification of Washing Machines Vibration Signals
Using Discrete Wavelet Analysis for Feature Extraction"
IEEE TRANSACTION ON INSTRUMENTATION AND MEASUREMENT, VOL.
51, NO. 3, JUNE 2002
- [20] 魏志中, "奇異值分解法圖樣分析於似週期信號之處理及其於
心電圖資料壓縮與雜訊消除之研究", 國立臺灣大學博士論
文, 1990
- [21] 林家陽, "以模糊類神經網路做多導程心電圖病症分類辨識",
中原大學碩士論文, 1988.
- [22] 石明家, "以疊代分類法分析生物醫學訊號之特徵及其應用",
中原大學碩士論文, 1988.
- [23] 簡明成, "以小波轉換為基礎之多導程心電圖資料壓縮與辨
識", 中原大學碩士論文, 1986.
- [24] 楊銘耀, "利用小波與類神經網路進行心電圖特徵擷取與病症
分類", 中原大學碩士論文, 1985.