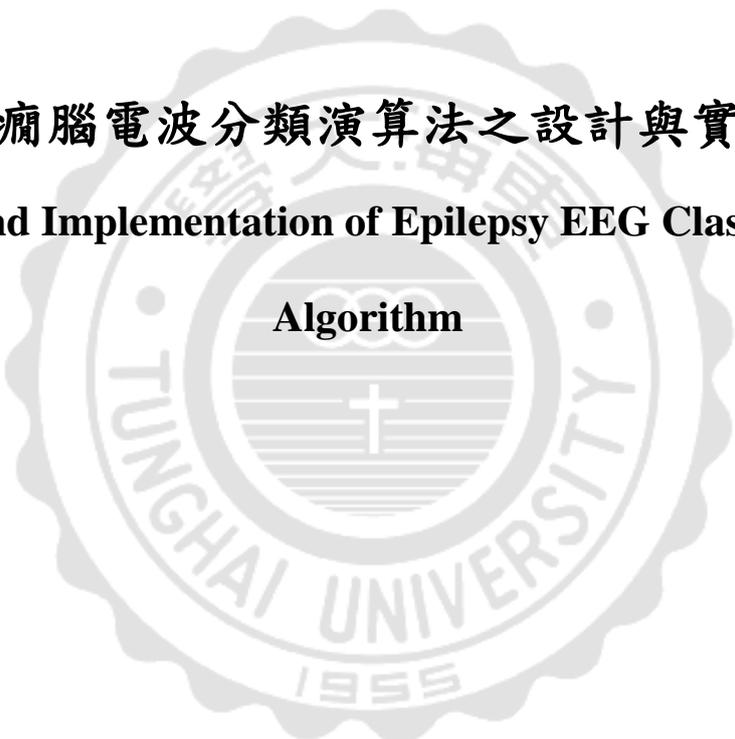


東海大學資訊管理研究所
碩士學位論文

癲癇腦電波分類演算法之設計與實作

Design and Implementation of Epilepsy EEG Classification

Algorithm

The seal of Tungshai University is a circular emblem with a scalloped outer edge. It features a central cross, a horizontal line, and the university's name in Chinese characters '東海大學' at the top and 'TUNGSHAI UNIVERSITY' in English around the inner circle. The year '1955' is inscribed at the bottom of the seal.

指導教授：林正偉 博士

研究生：黃侑晨 撰

中華民國 105 年 7 月

東海大學資訊管理學系碩士學位

考試委員審定書

資訊管理學系研究所_____黃侑晨_____君所提之論文

癲癇腦電波分類演算法之設計與實作

經本考試委員會審查，符合碩士資格標準。

學位考試委員會 召集人：

張瑞廷

(簽章)

委員：

林登昇

林正偉

吳中

高宏宇

中華民國 105 年 7 月 21 日

誌 謝

兩年的碩士學涯受到了許多人的照顧，在此感謝大家。尤其感謝我的指導教授林正偉老師，從大學時期教導我至今，老師諄諄教誨，不僅在學業上受益匪淺，也讓我在生活、未來規劃得到思考啟發的方向。在研究的最後階段，感謝口試委員張瑞益教授、高宏宇教授、林宣華教授、呂芳懌教授的指導，提供寶貴的建議，讓我的論文研究能更加完整。

感謝實驗室中的學長姐、同學以及學弟妹，讓我擁有兩年的充實生活與珍貴回憶。也特別感謝林祐陞、薛凱元學長在學業上的指導，還有承穎哥、允妍、炯文的陪伴，讓我的研究能順利完成。

最後感謝家人的支持，有你們的陪伴與鼓勵，是我求學路途上的最大動力，也因此，才能讓我無慮地走完全程。

論文名稱：癲癇腦電波分類演算法之實驗與實作

校所名稱：東海大學資訊管理學系研究所

畢業時間：105 年 7 月

研究生：黃侑晨

指導教授：林正偉

摘要：癲癇(epilepsy)是大腦皮質細胞不正常放電造成的一種症狀。腦電波儀(electroencephalography, 簡稱 EEG)是一種便宜且有效觀察癲癇發病的工具。在這個研究中，我們使用經驗模態分解(Empirical Mode Decomposition, 簡稱 EMD)將在癲癇患者觀察到的腦電波進行分解成數個本質模態函數(Intrinsic Mode Function, 簡稱 IMF)，從各個 IMF 抽取出能量(Energy)與標準差(SD)兩個特徵，配合基因演算法(Genetic Algorithm, 稱簡 GA)進行特徵選取，透過線性判別法(Linear Discriminant Analysis, 簡稱 LDA)進行分類。實驗結果顯示，我們的方法可以有效的區別出正常(Normal)、棘波(Spike)與癲癇(Seizure)三種不同類形的腦電波。結果顯示使用基因演算法配合線性判別法的分類，腦電波圖的正確率可以達到 Normal 97.4%、spike 84.6%、Seizure 95.6%。

關鍵詞 - 腦電圖、癲癇、基因演算法、經驗模態分解、線性判別法

Title of Thesis : Design and Implementation of Epilepsy EEG Classification

Algorithm

Name of Institute: Tunghai University, Graduate Institute of Information Management

Graduation Time : (07 / 2016)

Student Name : Yo-Chen Huang

Advisor Name : Jeng-Wei Lin

Abstract : Epilepsy is a cerebral cortical cells abnormal discharge caused by a symptom. EEG is a cheap and effective tool to observe epilepsy. In this study, we use the empirical mode decomposition to brain waves in epileptic patients was observed to be broken down into a number of nature mode function, extracted from various IMF out energy and standard deviation (SD) two features, with the genetic algorithm for feature selection, via a linear discriminant method to classify. Experimental results show that our method can effectively distinguish Normal, Spike and Seizure three different types of brain wave shape.

The results showed that the use of Genetic Algorithms with linear discriminant classification method, accuracy brainwave pattern can reach Normal 97.4%, Spike 84.6%, Seizure 95.6%.

Keywords : Electroencephalography · Epilepsy · Genetic Algorithms · Empirical

Mode Decomposition · Linear Discriminant Analysis

目錄

第一章 緒論	1
1.1 研究背景與動機	1
1.2 研究架構與流程	2
第二章 文獻探討	4
2.1 相關研究	4
2.2 特徵值抽取	4
2.3 癲癇	5
2.4 腦波圖	6
2.5 希爾伯特-黃轉換(HHT)	9
2.6 本質模態函數(IMF)	11
2.7 經驗模態分解(EMD)	13
2.8 基因演算法	14
2.9 線性判別分析(LDA)	16
第三章 研究方法	18
3.1 資料預處理	18
3.2 LDA 分類	21
3.3 特徵選擇	23
第四章 實驗	24
4.1 實驗資料	25
4.2 特徵值抽取與選取	26

4.3 結果	28
第五章 結論	30
參考文獻	31



圖次

圖 1-1 研究流程	3
圖 2-1 腦波頻率分佈.....	7
圖 2-2 國際 10-20 系統電極位置分布圖.....	8
圖 2-3 分類示意圖	8
圖 2-4 HHT 流程圖	10
圖 2-6 GA 流程圖	15
圖 2-7 SVM.....	17
圖 2-7 LDA.....	17
圖 3-1 單極	19
圖 3-2 雙極	19
圖 3-3 EMD with 400 data length.....	20
圖 3-4 Epileptic EEG classification	22
圖 3-5 GA 運作圖	23
圖 4-1 研究架構.....	24
圖 4-2 特徵值抽取流程圖	26
圖 4-3 Seizure	28
圖 4-4 spike	29

表次

表 4-1 腦波原始數據	25
表 4-2 數據 EMD	26
表 4-3 特徵值	27
表 4-3 實驗結果	29



第一章 緒論

1.1 研究背景與動機

癲癇(epilepsy)是大腦皮質細胞不正常放電造成的一種症狀，大約影響世界總人口的 1%。腦電波(Electroencephalography)訊號記錄了大腦皮質上不同電極間的電壓變化，在癲癇診斷上面扮演了十分重要的角色。文獻上，許多技術被提出用來偵測癲癇樣腦電波，如棘波(spike)、銳波(sharp wave)、與癲癇發作(seizure)等。他們使用了基於不同訊號處理技術的多種特徵，如使用小波轉換(wavelet transform, 簡稱 WT)、特徵分解(eigen-decomposition)、時頻分析(time-frequency analysis)、主成份分析(principal component analysis)、經驗模態分解(empirical mode decomposition)等等，採用不同的分析工具，如專家系統(expert systems)、樣版方法模式(template method pattern)、類神經網路(artificial neural network, 簡稱 ANN)、小波分析(wavelet analysis)、支持向量機(support vector machine)、卡爾曼濾波器(Kalman filter)、獨立成分分析(independent component analysis)、模糊分群法(fuzzy c-mean clustering)等等。此外，也有一些研究致力於腦電波(Electroencephalography)訊號的視覺化與音響化。

癲癇病患在進行腦電波檢驗時會持續數個小時，產生大量的資料需要許多人力來判讀，因此癲癇樣腦電波的自動分類技術在臨床上有很大的應用。在學術上，目前已知的技術仍可能無法有效的指出癲癇發作的部位，本研究提出一個系統化流程，透過基因演算法不斷尋找特徵值再運用線性判斷分析加以分類，其中我們針對三種腦電波進行分析，分別為正常(Normal)、棘波(spike)、發作(seizure)透過判別出這三類狀況，用以正確的判斷出病人腦波的狀況。

1.2 研究架構與流程

本研究如圖 1-1 所示，詳細如下：

一、緒論-研究主題與目的

本研究提出一個系統化流程，透過尋找數據中的共通點，再透過分類法加以分類。

二、文獻探討

此章節探討對於選擇演算法以及分類演算法的研究，透過這些演算法設計出一個流程。

三、研究方法

說明整個研究的流程從一開始的資料處理到最後的分類。

四、實驗結果

經由上述提出的研究方法，提出實驗的設計與模組。

五、結論

分析本研究的實驗結果做出結論，並藉此提出建議和未來方向。

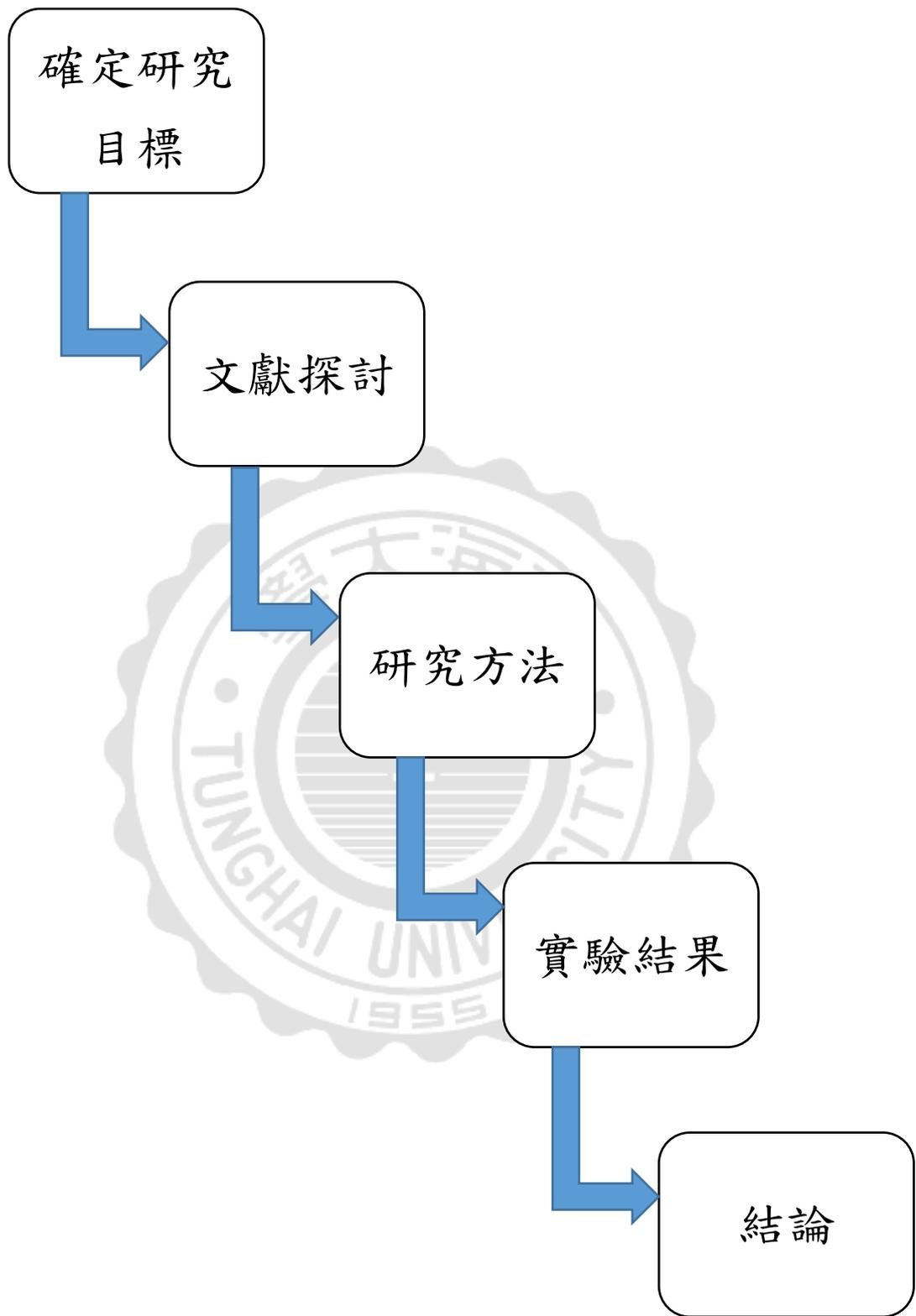


圖 1-1 研究流程

第二章 文獻探討

本章節分為兩個部份來探討，第一部分為實驗前資料的處理，第二部分為實驗的時候使用的演算法部分。

2.1 相關研究

在腦波辨識研究中，根據蕭宇純(2011)的研究我們可以發現主要有兩種研究方向：包含特徵計算與分類系統的設計。蒐集不同的腦波訊號，找尋有效的特徵，再搭配適合的分類系統，讓腦波辨識能在這兩種架構的組合下得到最好的辨識結果。

所以我們根據不同的特徵值抽取及分類組合試著找出最正確的方法，根據林偉仁 (2010)的研究提出的所有方法中發現基因演算法(Genetic Algorithm 簡稱 GA) 搭配支持向量機(Support Vector Machine 簡稱SVM)的正確率是最高的，而陳維(2013)的研究中是一種簡單的搜尋演算法搭配線性判斷分析 (Linear Discriminant Analysis 簡稱LDA)正確率也是很高的，所以本研究使用GA搭配LDA希望透過此方法能得到更好的結果。

2.2 特徵值抽取

在對於腦電波轉換這塊有非常多的人提出，例如傅立葉變換、小波轉換、希爾伯特-黃轉換等等，所以在本文中我們使用希爾伯特-黃轉換，希爾伯特-黃轉換是一種腦波轉換機制，經過經驗模態分解(Empirical Mode Decomposition 簡稱EMD)得出本質模態函數(Intrinsic Mode Function 簡稱 IMF)，再經由 IMF 抽取出特徵值，而經過希爾伯特-黃轉換之後我們能清楚的知道腦波的順時速率、震幅變化等等。

2.3 癲癇

癲癇是一種長期性神經系統疾患，以抽搐發作為特徵。這些抽搐發作形式很多，Chang (2003) 研究中指出有些可以是非常短暫且幾乎無法察覺，也可以是長時間的劇烈抽動。所以癲癇的抽搐往往反覆發作而並無直接起因，但有時候我們也會有些不正常抽搐的現象由一些，而這些因為特殊原因引起的抽搐則不被認為是癲癇。

雖然有些癲癇病例是由於腦部外傷、中風、腦腫瘤、服毒或酗酒等病因導致，但大多數癲癇發病並無明確病因。癲癇抽搐發作是由於大腦皮質神經細胞活動過度或不正常的結果。診斷癲癇通常需要排除其他能引起類似症狀的疾患（如暈厥），並查找是否有直接發病因素存在。癲癇通常可以以腦電圖(EEG) 確診。

癲癇不能根治，但 70%癲癇病例的抽搐發作都可以透過藥物進行控制。對於藥物無法控制的抽搐發作，可以考慮手術、神經刺激療法或改變飲食。並非所有癲癇症狀都是終身性，很多病人都可以得到緩解並達到無需服藥的水平。

全世界約有 1%的人口（六千五百萬人）患有癲癇，近 80%的病例都發生在發展中國家。癲癇發病率隨年齡增長而增高。在已開發國家最病例最常見於嬰幼兒和老人，在發展中國家則最常見於年齡較大的兒童或年輕成人，這是因為發病病因的比例不同。Epilepsia. (2008) 書中指出約 5 - 10%的病例在 80 歲前會發生一次無明顯誘因的抽搐，發生第二次抽搐的幾率在 40%到 50%之間。全世界許多地區都會對癲癇患者駕車做出限制或禁止，但大多數情況下他們在無抽搐生活一段時間後可以重新開始駕車。

2.4 腦波圖

在 1929 年，德國耶那(Jena)大學的精神科教授 Berger 博士經過多年的研究，終於在人類的頭蓋骨上測得電位變化(Berger 1929)，而將此腦電位活動命名為腦波圖(Electroencephalogram) 如圖 2-1，並在之後陸續發表有關腦波研究的論文，因此對腦波研究上有極大的貢獻所以被稱為「人類腦波圖之父」。由於腦波圖的量測方式安全、簡單，因此在醫學領域上常用來診斷癲癇、昏迷、睡眠障礙、腦外傷等神經科疾病，除此之外，在電機、心理等學科也利用腦波來發展腦機介面用於復建工程、遊戲、教育等領域。

通常腦波的產生主要是經由大腦皮質或頭皮上的神經細胞活動所產生的電位變化總合。腦波是屬於非穩態性(Nonstationary)的波形，但大致上可以從振幅、頻率、波形這三方面來分析，

1. 振幅：

腦波由大腦皮質傳到頭皮會受到組織、顱骨的阻隔，所以實際量測到的電訊號相當微弱，只有數 10μ 伏特大小的變化，且容易受到外界的雜訊或其他生理訊號干擾，目前量測腦波的機器都需要使用放大器將訊號放大百萬倍後再處理。

2. 頻率：

正常的腦波頻率範圍約在 0.5 - 50 Hz 之間，其中又可依不同頻帶化分為 δ 、 θ 、 α 、 β 四種基本背景波，再加上 γ 波，如圖 2-1。

3. 波形：

常見的腦波波形有棘波(Spike)、多重棘波、 μ 波(8 - 13Hz)、高幅快波(50μ V 以上)等。

通常量測腦波的方式都是使用國際公定的 10-20 系統(International 10-20system)，將大腦劃分成左半邊與右半邊，再用英文字母和數字給予每一個部位不同名稱，如圖 2-2，我們的數據來源是來自於台大醫院它們是使用 16 個電

位，分別為左邊 8 個電位 Fp1、F3、F7、C3、T3、P3、T5、O1，以及右邊 8 個電位 Fp2、F4、F8、C4、T4、P4、T6、O2，從 16 個電為抽取出腦波數據並加以運用。

本研究主要針對三種癲癇腦波進行分類分別，為正常(Normal)、棘波(spike)、發作(seizure)透過判別出這三類狀況，如圖 2-3。

δ	0.5 – 4 Hz
θ	4 – 8 Hz
α	8 – 13 Hz
β	13 – 30 Hz
γ	>30 Hz

圖 2-1 腦波頻率分佈

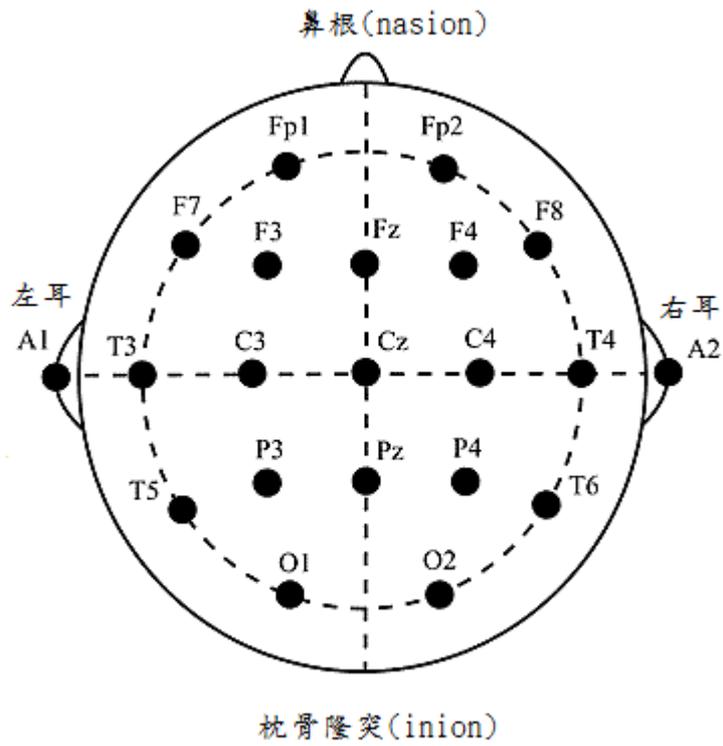


圖 2-2 國際 10-20 系統電極位置分布圖

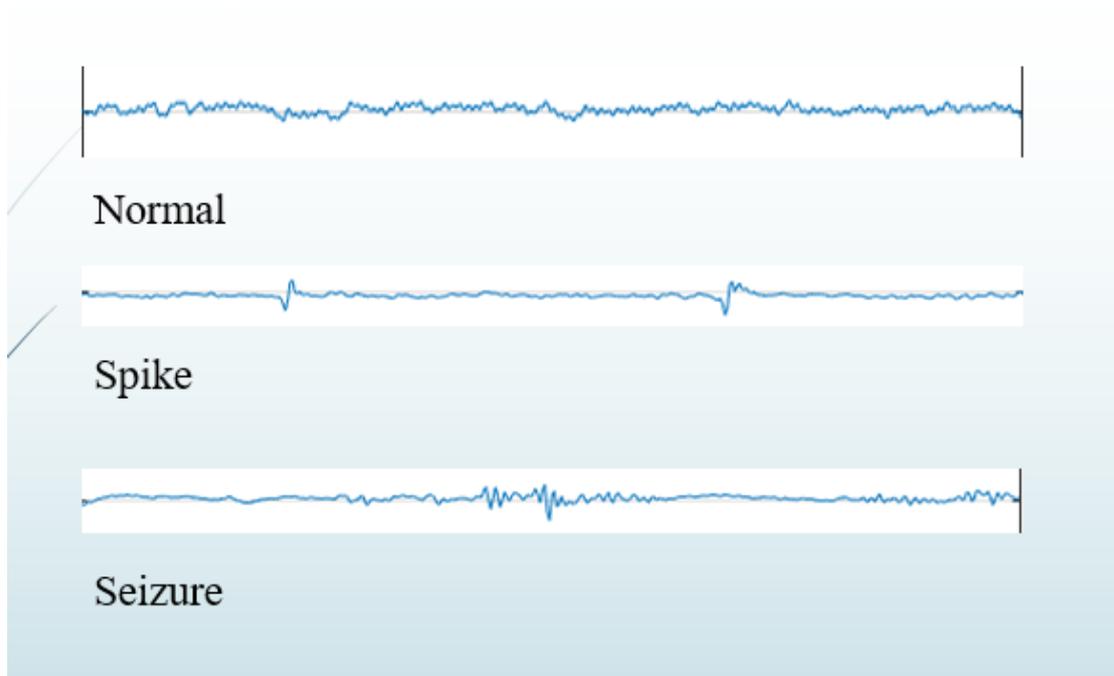


圖 2-3 分類示意圖

2.5 希爾伯特-黃轉換

在對於處理訊號轉換上有非常多的方法如使用小波轉換(wavelet transform)、特徵分解(eigen-decomposition)、時頻分析(time-frequency analysis)、主成份分析(principal component analysis)、經驗模態分解(empirical mode decomposition)等等，而本研究使用希爾伯特-黃轉換。

希爾伯特-黃轉換(Hilbert Huang Transform 簡稱 HHT)是 1998 年黃鐸博士於美國太空總署(National Aeronautics and Space Administration)發表(黃鐸 1998)，透過 EMD 利用資料變化的內部時間尺度來做能量的直接析出，將原來訊號資料展開成多個 IMF，把內建模態函數當作展開的基底，如此一來這些基底就可以完全顯示原來訊號的物理特性。因此被分析的訊號就可以為非線性或非平穩性，其因為此法是利用訊號本身的變化來做分解，再利用 HHT 來求出內建模態函數的瞬時頻率及振幅，為頻率加權的正規化同時結合時間、振幅、頻率之分佈，使訊號能表現瞬時變化的訊息與特性，稱為 HHT。HHT 換具有以下四點特性。所示為 HHT 架構之流程圖，如圖 2-4。

1. 完整性：確保展開之精確度。
2. 正交性：保持能量的正定及避免能量不保守。
3. 局部性：使訊號表現瞬時變化的特性。
4. 可適性：適應暫態、非線性及非穩定性之系統條件。

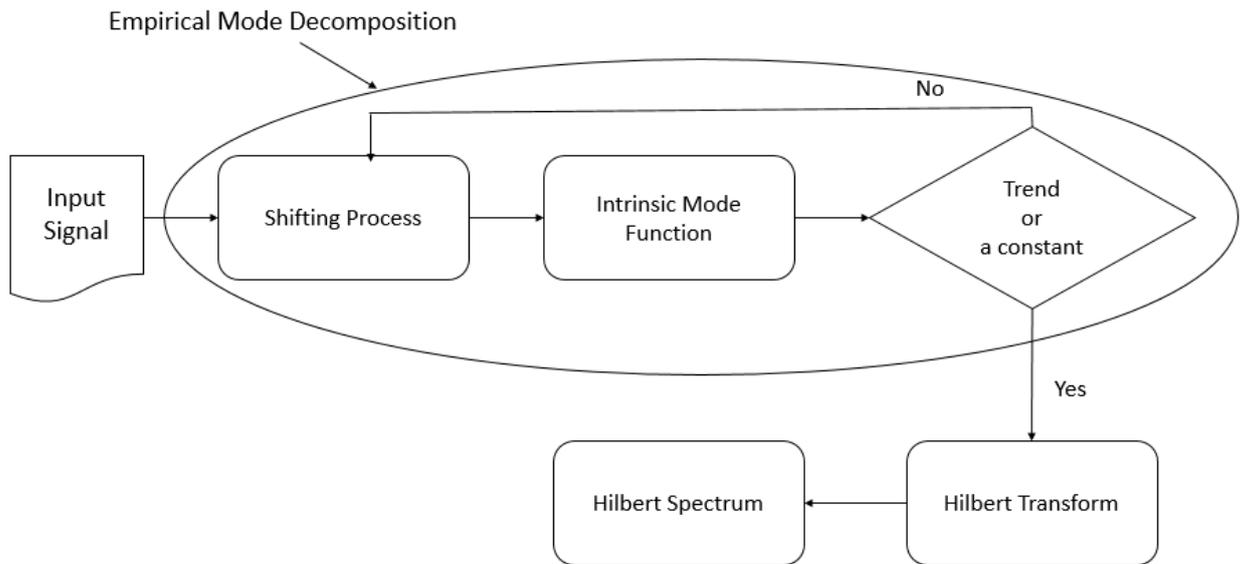


圖 2-4 HHT 流程圖



2.6 經驗模態分解(Empirical Mode Decomposition)

IMF 為 HHT 的前置條件，但大部分的輸入訊號都不是 IMF 這種形式，因為在任何給定的時間，可能包含了不只一個振動模態，因此必須分解輸入訊號成為多個 IMF 分量，要達到這樣的目的，則必須藉由 EMD，簡單的來說，EMD 就是將輸入訊號分解成多個帶有物理特性的 IMF，而 IMF 就如同傅立葉轉換中的正弦、餘弦函數為 HHT 的基底。

首先將訊號分解成 IMF 的組合，且每一個 IMF 都滿足上述所提之兩個條件。其實 EMD 是藉著不斷重複作用的篩選程序(Shifting Process)來逐步找出本質模態函數。以訊號 $x(t)$ 為例，將篩選程序的流程概述如下：

1. 首先找出 $x(t)$ 中的所有局部極大值及局部極小值，接著利用 cubic spline 將局部極大值串連成上包絡線，至於局部極小值也利用 cubic spline 連接構成下包絡線。

2. 求出上下包絡線之平均，而得到均值包絡線 $m_1(t)$ 。

3. 將原始訊號與均值包絡線相減，得到第一個分量 $h_1(t)$

$$h_1(t) = x(t) - m_1(t) \quad (1)$$

(1) 至此視為完成第一個篩選程序。接著檢查 $h_1(t)$ 是否符合 IMF 的條件，如果不符合，則回到步驟 1 並且將 $h_1(t)$ 當作原始訊號，進行第二次的篩選程序。

亦即

$$h_2(t) = h_1(t) - m_2(t) \quad (2)$$

(2) 重複篩選 k 次

$$h_k(t) = h_{k-1}(t) - m_k(t) \quad (3)$$

(3) 直到 $h_k(t)$ 符合 IMF 的條件，即得到第一個 IMF 分量 $c_1(t)$

，亦即

$$c_1(t) = h_k(t) \quad (4)$$

(3) 將原訊號 $x(t)$ 減去 $c_1(t)$ 可以得到殘餘量 $r_1(t)$ ，可表示如下式

$$r_1(t) = x(t) - c_1(t) \quad (5)$$

(5) 接著將 $r_1(t)$ 當作新的資料，重複經由篩選程序解析出第二個 IMF 分量 $c_2(t)$ ，並將其由中 $r_1(t)$ 減去，以得到新的殘餘量 $r_2(t)$ 。如此重複分解 n 次

$$r_2(t) = r_1(t) - c_2(t)$$

$$r_3(t) = r_2(t) - c_3(t)$$

$$r_4(t) = r_3(t) - c_4(t)$$

(6)

$$r_n(t) = r_{n-1}(t) - c_n(t)$$

(6) 當第 n 個殘餘量 $r_n(t)$ 已成為單調函數 (Monotonic Function)，無法再解析出 IMF 時，整個經 EMD 的分解過程即可視為完成。EMD 的基底是從原始訊號分解而來的，因此這個方法是直觀的、直接的、後知的且具有可適性的。

2.7 本質模態函數(Intrinsic Mode Function)

任何一個資料序列，滿足下列兩個條件即可稱作 IMF。

1. 在整個資料中，局部極大值(local maxima)及局部極小值(local minima)的數目之和必須與零交越點(zero crossing)的數目相等或是最多只能差 1。
2. 在任一時間點上，由局部極大值所定義的上包絡線(upper envelope)與局部極小值所定義的下包絡線(lower envelope)之平均值為零。



2.8 基因演算法(Genetic Algorithm)

GA是一種是計算數學中用於解決最佳化的搜索算法是根據達爾文於1859年提出了「物競天擇，適者生存」演化學說(達爾文1859)，而演化出來的。在對於特徵值的挑選上有些人使用費雪鑑別分析、貪婪演算法、逐次反向式搜尋法、逐次正向式搜尋法等。

費雪鑑別分析主要是用來觀察每一種特徵與每一類之間的關聯，而逐次反向式搜尋法、逐次正向式搜尋法是根據在文中提到這兩種搜尋法一種是逐次反向式搜尋法是將全部特徵個數依序減少一個特徵個數，直到設定的特徵個數為止，藉此觀察被保留的特徵的分類效果，而另一種則是逐次正向式搜尋法是將全部特徵個數依序增加一個特徵個數。

GA概念則首次出現於1975年John H. Holland大作Adaptation in Natural and Artificial Systems (Holland, 1975)，Holland在該書中認為，除了自然界生物基因學研究之外，經濟學、遊戲理論、模式辨識、控制與函數最佳化等領域當中，都有近似基因工程之現象。換句話說，不管是自然環境或是人造環境，都可以運用基因演算法描述一些現象，甚至可能預測某些未知現象之發生。Holland在該書所提示領域，幾乎指出後來GA所應用的範疇。基因演算法是運用電腦模擬適者生存，不適者淘汰的演化規則，經由複製、交配及突變運算，搜尋出最佳解。基因演算法經過多年的發展，現今已成功地應用在很多領域。工程方面如類神經網路、模糊理論、控制、水管路、太空軌道、航空等。物理方面如液晶、VLSI佈局、雷射技術。其中應用在最佳化設計方面十分的廣泛，例如：工程設計最佳化、幾何外型最佳化、機械設備規劃、生產線排程規劃、機器人動線規劃等(謝宗廷 2012)。

GA不同於一般傳統的搜尋法，它是隨機性的搜尋法，並不需要初始設計值，一開始會先由電腦亂數產生一群初始的設計值組，稱為初始族群，此設計值組是以類似於染色體的資料結構進行編碼，並使用一些模仿生物演化規則的電腦運算

方式，如交配、突變對大量的染色體進行運算。運算後所產生的子代除了能保留親代中優良特徵之外，也可能因為交配和突變而產生出比親代更優良的特徵，所示為基因演算法之流程圖 2-5。

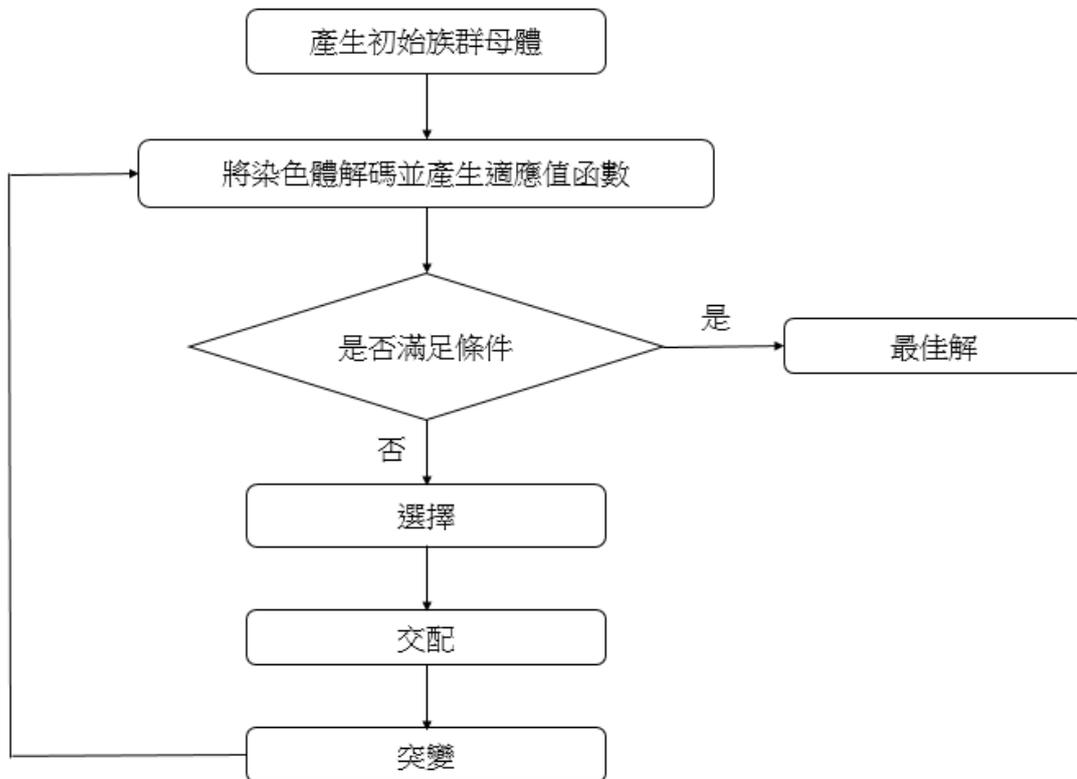


圖 2-5 GA 流程圖

2.9 線性判別分析 (Linear Discriminant Analysis)

線性判別分析 (Linear Discriminant Analysis)，簡稱判別分析，是統計學上的一種分析方法，用於在已知的分類之下遇到有新的樣本時，選定一個判別標準，以判定如何將新樣本放置於哪一個類別之中。

它不同於以往的分析法，大部分的分析法是以群體的相似度去進行分析例如 SVM，如圖 2-6、交叉驗證等，SVM 是 Corinna Cortes 和 Vapnik 等人於 1995 年首先提出的，它在解決小樣本、非線性及高維模式識別中表現出許多特有的優勢，並能夠推廣應用到函數擬合等其他機器學習問題中。

交叉驗證，有時亦稱循環估計是一種統計學上將數據樣本切割成較小子集的實用方法。於是可以先在一個子集上做分析，而其它子集則用來做後續對此分析的確認及驗證。一開始的子集被稱為訓練集。而其它的子集則被稱為驗證集或測試集(Kohavi, Ron.1995)。

在本研究我們使用線性判別分析，如圖 2-7，而線性判別分析是從已知的結果中尋找特別突出特徵當作判斷依據進而分類，而這種方法比較能準確的分類出所需的結果，此分類法主要應用於醫學的患者疾病分級，以及人臉識別、經濟學的市場定位、產品管理及市場研究等範疇。

關於線性判別分析的研究應追溯到 Fisher 在 1936 年發表的經典論文(Fisher R A. The use of multiple measurements in taxonomic problems) (Fisher 在 1936)，其基本思想是選擇使得 Fisher 準則函數達到極值的向量作為最佳投影方向，從而使得樣本在該方向上投影后，達到最大的類間離散度和最小的類內離散度。

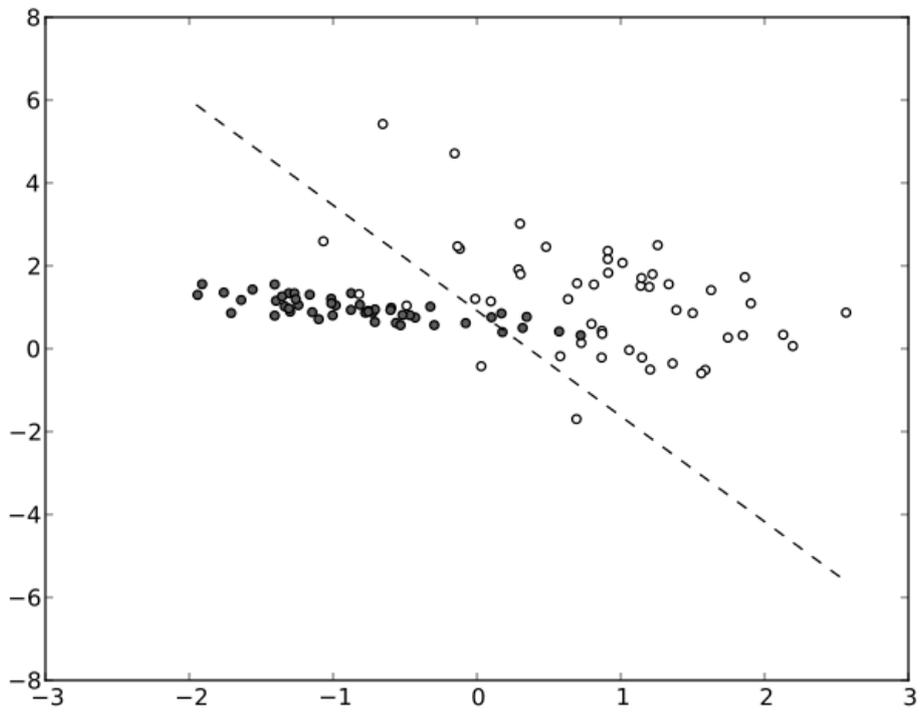


圖 2-6 SVM

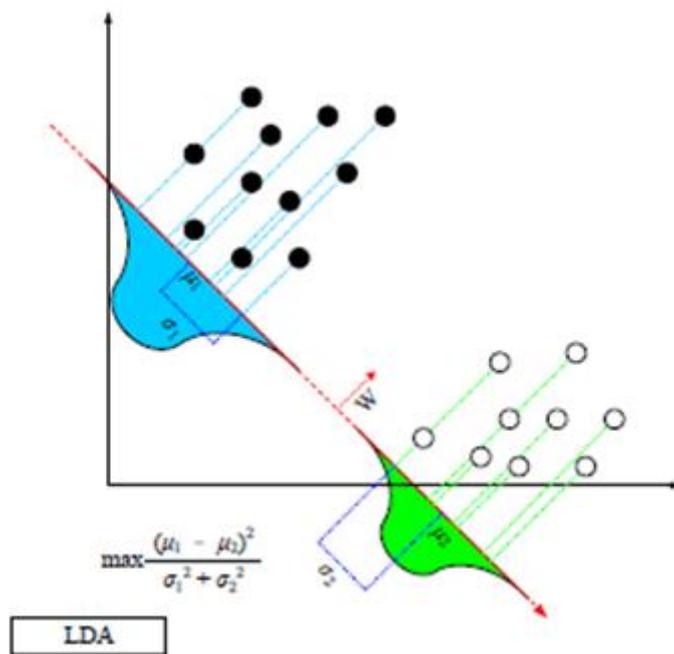


圖 2-7 LDA

第三章 研究方法

本研究可以分成三個部分，1. 資料預處理、2. 特徵選擇、3. LDA 分類

3.1 資料預處理

一開始我們把腦電波資料進行資料預處理，我們首先把腦電圖數據從 unipolar montage 如圖 3-1 轉換成 bipolar montage 如圖 3-2 一般而言，使用單極組合範式時，因能看到訊號真正的強度，可以比較清楚的看出最強或最弱病灶所在，但也極易受到污染。雙極組合範式雖然不能看出電極的真正電位，在某些情形下，較單極組合範位更容易看出病灶。。

每 2 秒分成一段位移距離為 0.05 秒，信號數據長度為 400，之後運用蝶形濾波器提取信號從 1Hz 到 60Hz，消除低頻電壓和高頻噪音。

在腦電波資料中總共有 16 個 channel，然後每一個 channel 經由 EMD 分解為 8 個 IMFs，如圖 3-3。

之後每個 IMF 都提取出兩個特徵分別為能量(power)和標準差(SD)，其結果會提取出 256 個特徵值(16 bipolar channel * 8 IMFs * 2 features)。

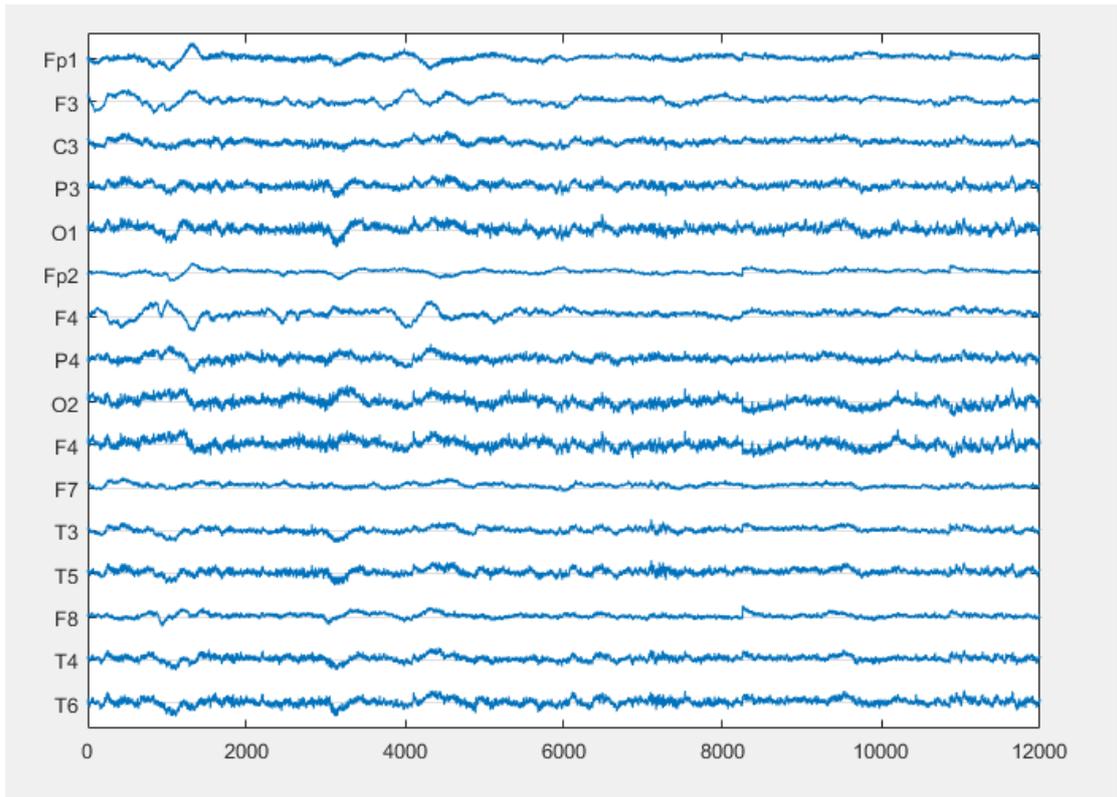


圖 3-1 單極

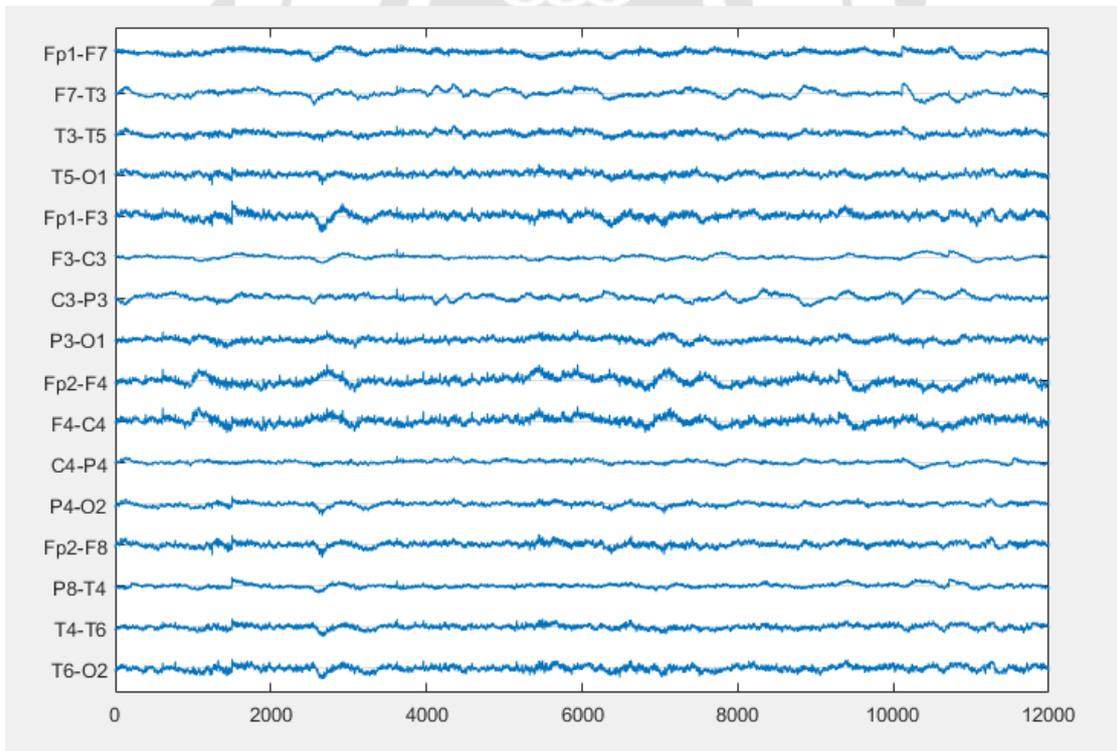


圖 3-2 雙極

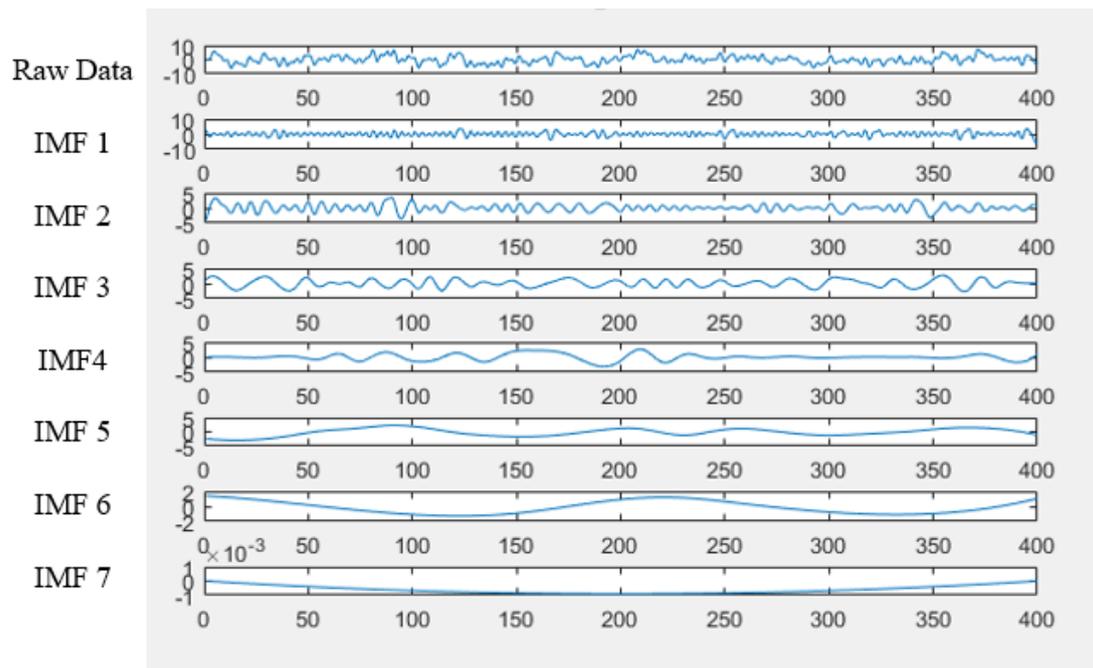


圖 3-3 EMD with 400 data length



3.2 LDA 分類

在這項研究中，我們會研究出一個簡單的結構來分類，如圖 3-4，因為通過一次 LDA 我們只能把資料分成兩類 LDA，但是本研究是針對三個分類進行探討，所以我們通過兩次 LDA 把資料分成了三類，對於分類的順序我們透過以往的研究發現 seizure 是最容易分辨的，我們會做兩次分類。一次分類為 seizure or non-seizure 另一次則是 spike or non-spike。

我們使用 LDA 分類時必須給定一組含有 2 個類別的訓練資料集，每筆訓練資料為包含特徵向量 x 與其類別標記 1 或 2，第 1 類與第 2 類的平均值與共變異數矩陣(covariance matrix)分別為 (μ_1, Σ_1) 與 (μ_2, Σ_2) ，透過 Bayes 最佳解，可由公式(1)來判別特徵向量 x 是否被分類到類別 1，其中 T 為一個經驗閾值(threshold)。

$$(x - u_1)^T \Sigma_1^{-1} (x - u_1) + \ln|\Sigma_1| - (x - u_2)^T \Sigma_2^{-1} (x - u_2) + \ln|\Sigma_2| < T \quad (7)$$

在 LDA 中，假定 $\Sigma_1 = \Sigma_2 = \Sigma$ 。預設閾值 T 為 0，則公式(1)可以簡化為公式(2)。

$$u_1^T \Sigma^{-1} x - (u_1^T \Sigma^{-1} u_1) > u_2^T \Sigma^{-1} x - (u_2^T \Sigma^{-1} u_2) \quad (8)$$

我們定義 $L_c(x)$ 為特徵向量 x 的類別判定指標， $c=1$ 或 2，如公式(3)所示。

$$L_c(x) = u_c^T \Sigma^{-1} x - \frac{1}{2} (u_c^T \Sigma^{-1} u_c) \quad (9)$$

亦即，當 $L_1(x) > L_2(x)$ 時，即 $L_1(x) - L_2(x) > 0$ 時，特徵向量 x 被分類到類別 1。

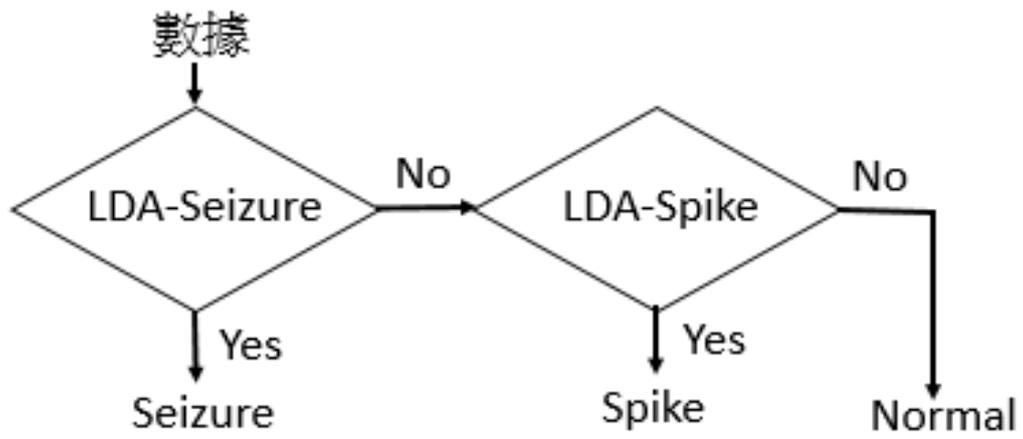


圖 3-4 Epileptic EEG classification



3.3 特徵選擇

在特徵抽取後，我們預計沿用之前研究所使用的基因演算法(GA)來進行特徵子集的選取。在分類的過程，不同的特徵子集會有不同的分類效果。在 GA 進行特徵子集演化的過程中，有時，某些對於分類沒有顯著效果的特徵會一直保留在某些精英特徵子集中。如圖 3-5。

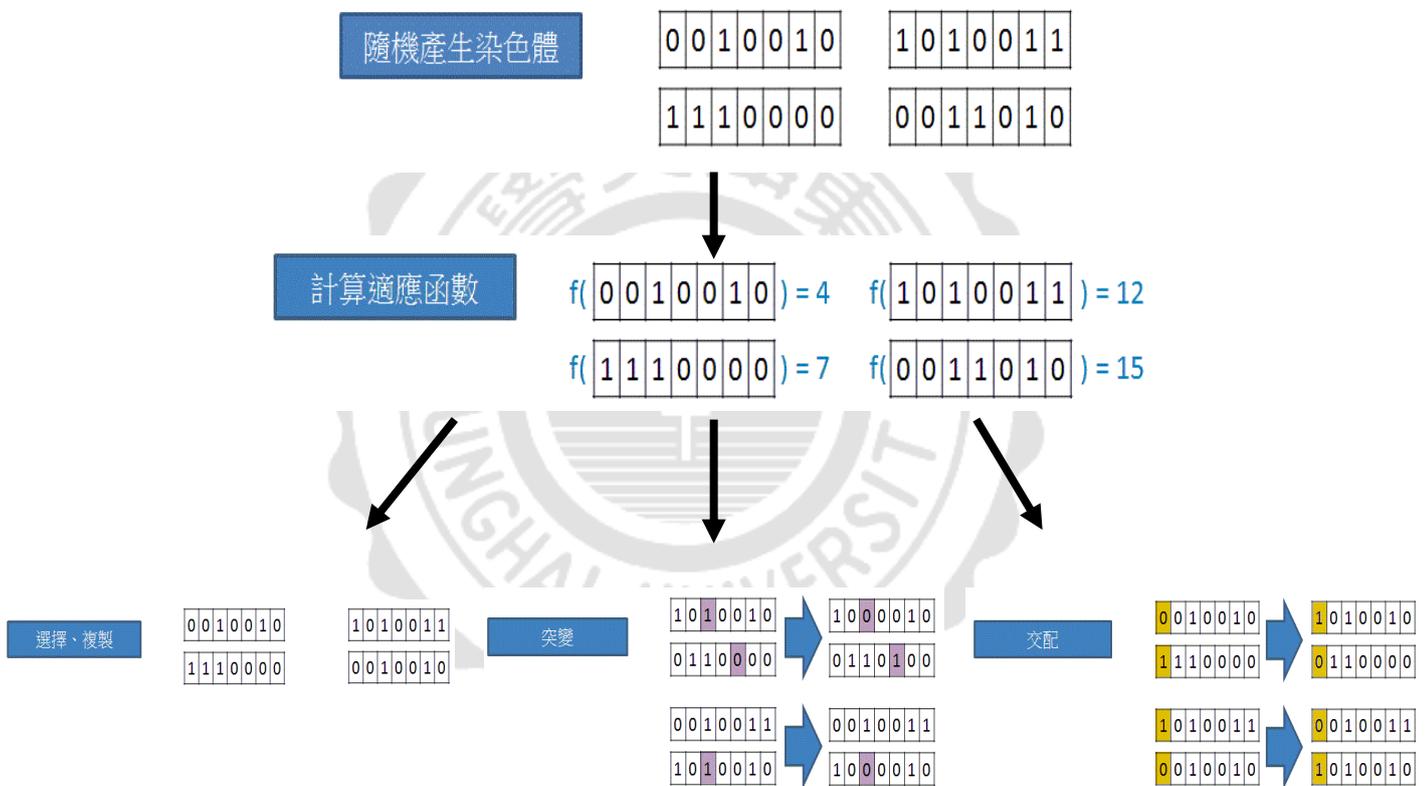


圖3-5 GA運作圖

第四章 實驗

根據研究方法本研究分為四個部分別為1. 資料預處理 2. 特徵值提取 3. 特徵值選擇 4. 分類，本章節會說明每個流程最後再根據實驗結果對比之前的研究來探討，如圖4-1。

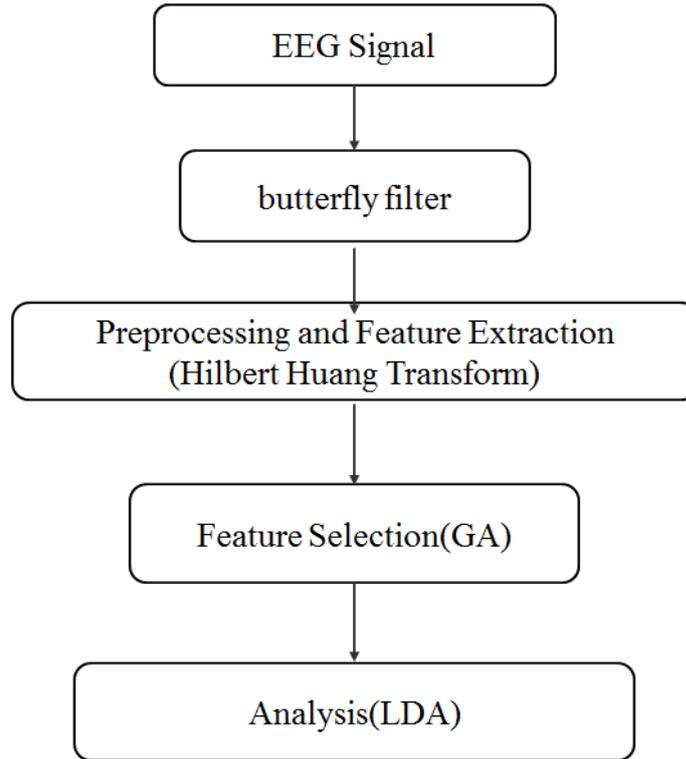


圖4-1 研究架構

4.1 實驗資料

首先我們進行資料預處理，我們資料來源台大醫院，資料格式為單極，電位為16個Channel，病患Normal 4個、Seizure 2個，我們總共蒐集了24組已知腦波數據，分別是Normal 8組、Spike 8組、Seizure 8組，如表4-1，每一組都有一串腦波數據分別都是16個Channel，最後為了檢驗實驗的正確性所以我們把數據分成了兩類，每個類別的前4組總共12組為training而後12組為testing。

表4-1 腦波原始數據

Normal	spike	Seizure
16*7200	16*4250	16*2839
16*8000	16*7750	16*6601
16*12000	16*7800	16*6003
16*18000	16*23401	16*5401
16*16000	16*6800	16*13805
16*14000	16*10500	16*12003
16*12000	16*10600	16*12201
16*12000	16*9000	16*7993

4.2 特徵值抽取與選取

首先使用HHT進行腦波轉換，並透過IMF和EMD來抽取特徵值，如圖4-2，總共16個Channel而每一個Channel我們都會抽出8個IMF而每個IMF會抽取出2個特徵所以經過轉換之後我們會得到24組新的數據，每組數據的特徵值都有256個如表4-2

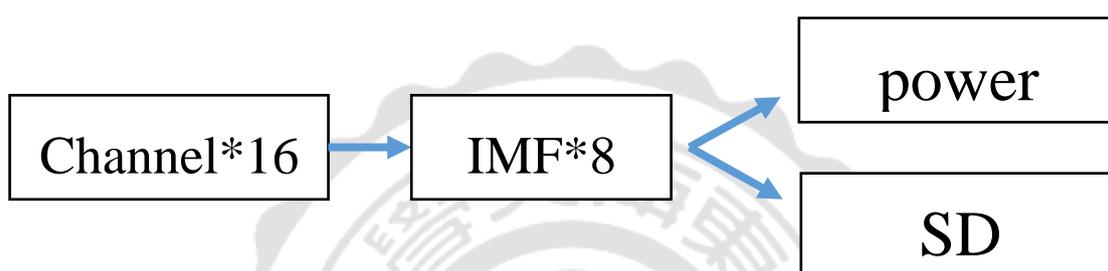


圖4-2 特徵值抽取流程圖

表4-2 處理後的數據

Normal	spike	Seizure
256*680	256*385	256*244
256*760	256*735	256*621
256*1160	256*740	256*561
256*1760	256*2301	256*501
256*1560	256*640	256*1341
256*1360	256*1010	256*1161
256*1160	256*1020	256*1181
256*1160	256*860	256*760

當特徵值後進行特徵值的挑選這裡我們使用的是GA，我們設定群體數為20，挑選個數為隨機，突變率為0.2交配率為0.4，本研究透過GA隨機挑選出了14個特徵值並對照陳(2013)中所選擇的特徵值對比，如表 4-3。

表 4-3 特徵值

基因演算法	7	8	11	16	20	26	34	36	37	39	45	46	50	
													62	
陳(2013)	6	7	8	10	18	20	26	34	36	37	39	40	49	50
														62



4.3 結果

並使用挑選出來的特徵值當作依據進行LDA分類，一開始我們先分類 Seizure也就是發作再來是spike最後剩下的就是Normal，然後首先使用training數據進行分類，再來透過分類完的數據再來對照已知的結果這時我們可以得到一個正確率，用以確認分類的正確性，並透過每一次基因演算法的挑選來重複這個過程，最後我們把基因演算法的代數停在30代，而得到再30代的時候挑選出來的特徵值Seizure 6個、Spike 8個，總共14個特徵值，正確率為Normal 97.12%、spike 85.35%、Seizure 95.1%如圖4-3 4-4，X軸代表的是代數，Y軸代表的是正確率，最後使用training後得到的特徵值進行testing正確率為Normal 97.4%、Spike 84.6%、Seizure 95.6%如表4-3，再對照之前參考的研究室使用貪婪演算法的發現在Seizure和Normal部分都差不多但在spike部分是有成長的如表4-4。

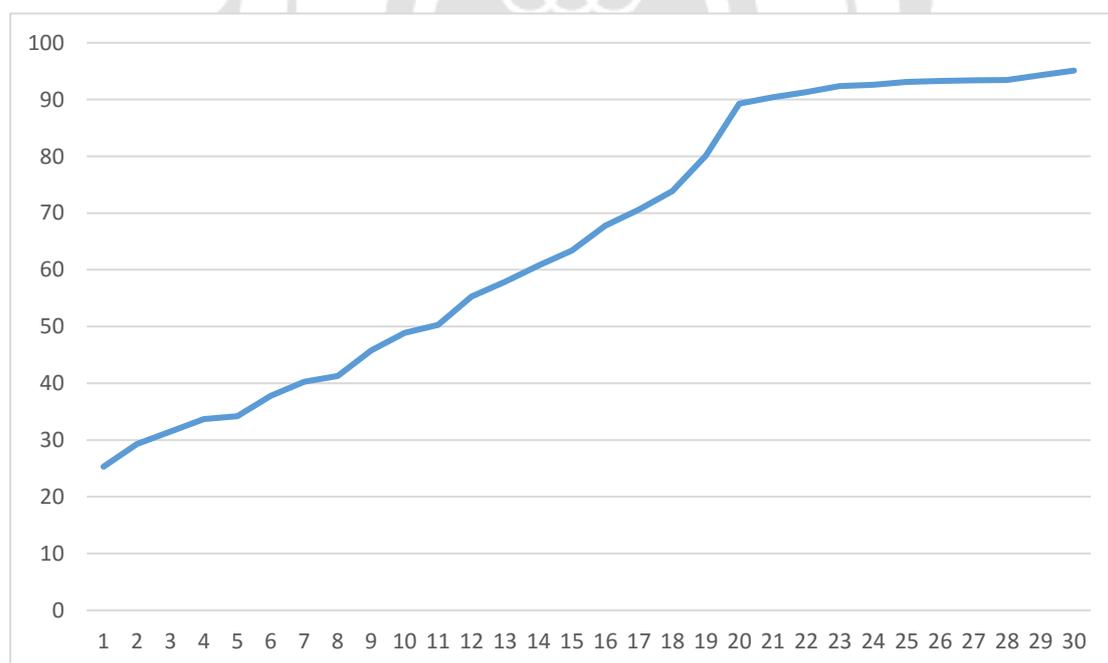


圖4-2 Seizure

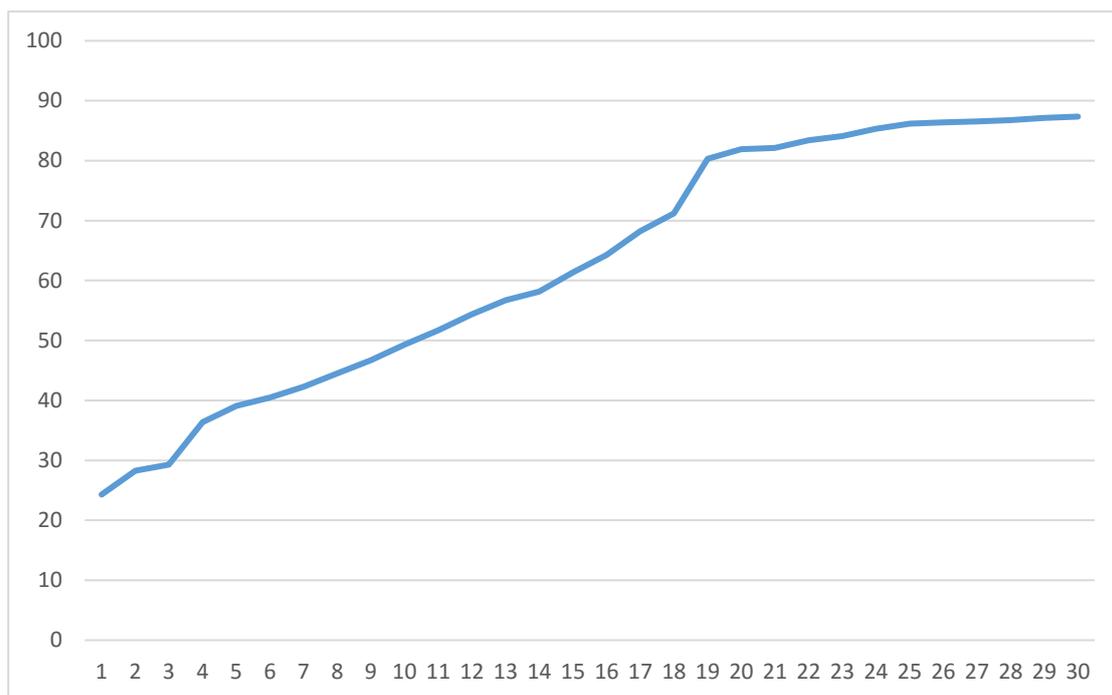


圖4-3 spike

表4-4 實驗結果

		Normal	Spike	Seizure
基因演算法	Training	97.12%	85.35%	95.1%
	Testing	97.4%	84.6%	95.6%
陳(2013)	Testing	97.6%	79.7%	95.3%

第五章 結論

癲癇病患在進行腦電波檢驗時會持續數個小時，產生大量的資料需要許多人力來判讀，癲癇樣腦電波的自動分類技術在臨床上有很大的應用。本研究提出一個系統化流程，透過基因演算法不斷尋找特徵值，並運用線性判斷分析加以分類用以正確的判斷出病人腦波的狀況。

在本文中我們透過基因演算法不斷的演化來產生新的下一代藉此來尋找到最佳的後代，而透過最佳解後，我們能夠清楚的分辨出病人的腦電波是處於何種情況我們能藉此向醫生或是護士提供準確的信息讓他們能透過這些信息來做出正確的判斷。



參考文獻

1. 謝宗廷，應用基因演算法估算塑膠模具射出成型模穴溫度，碩士論文，國立高雄應用科技大學模具工程研究所，高雄，2012。
2. 李穎潔，邱意弘，朱貽盛，”腦電信號分析方法及其應用”，科學出版社，2009.
3. 鄭金林，整合基因演算法與灰預測模型於室內溫度預測之研究，碩士論文 國立台北科技大學 2016
4. 文定宇，使用適應性基因演算法與灰預測模糊PID控制之電力系統穩定器設計，碩士論文，淡江大學，臺北，2003。
5. 林偉仁，基於腦波圖及心電圖辨識情緒的研究，國立中正大學電機工程研究所碩士論文基 2012
6. H. Witte, L. D. Iasemidis, and B. Litt, “Special issue on epileptic seizure prediction,” *IEEE Trans. Biomed. Eng.*, vol. 50, pp. 537–539, 2003.
7. N. E. Huang, Z. Shen and S. R. Long, et al., “Thee DeHilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis”, *Proc. Royal Society*, vol. 454, pp. 903-995, London, 1998.
8. Güler and E. D. Übeyli, “Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients,” *J. Neurosci. Methods*, vol. 148, pp. 113–121, 2005.
9. E. D. Übeyli and I. Güler, “Features extracted by eigenvector methods for detecting variability of EEG signals,” *Pattern Recognit. Lett.*, vol. 28, pp. 592–603, 2007.
10. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, “Epileptic seizure detection in EEGs using time frequency analysis” *IEEE Trans. Inf. Technol. in Biomed.*, vol. 13, pp. 703–710, 2009.

11. S. Ghosh-Dastidar and H. Adeli, "Principle component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection," *IEEE Trans. Biomed. Eng.*, vol. 55, pp. 512–518, 2008.
12. W. Chen, Y. Y. Lam, C. P. Shen, H. Y. Sung, J. W. Lin, M. J. Chiu, and F. Lai, "Ultra-fast Epileptic seizure detection using EMD based on multichannel electroencephalogram," in *IEEE Int. Conf. Bioinformatics and Bioengineering*, pp. 1-4, 2013.
13. B. L. Davey, W. R. Fright, G. J. Carroll, and R. D. Jones, "Expert system approach to detection of epileptiform activity in the EEG," *Med. Biol. Eng. Comput.*, vol. 27, pp. 365–370, 1989.
14. Nonclercq, M. Foulon, D. Verheulpen, C. D. Cock, M. Buzatu, P. Mathys, and P. V. Bogaert, "Spike detection algorithm automatically adapted to individual patients applied to spike-and-wave percentage quantification," *Neurophysiol. Clin.*, vol. 39, pp. 123–131, 2009.
15. W. Ko and H. W. Chung, "Automatic spike detection via an artificial neural network using raw EEG data: Effects of data preparation and implications in the limitations of online recognition," *Clin. Neurophysiol.*, vol. 111, pp. 477–481, 2000.
16. K. P. Indiradevi, E. Elias, P. S. Sathidevi, S. D. Nayak, and K. Radhakrishnan, "A multi-level wavelet approach for automatic detection of epileptic spikes in the electroencephalogram," *Comput. Biol. Med.*, vol. 38, pp. 805–816, 2008.
17. N. Acir and C. Guzelis, "Automatic spike detection in EEG by a twostage procedure based on support vector machines," *Comput. Biol. Med.*, vol. 34, pp. 561–575, 2004.
18. V. P. Oikonomou, A. T. Tzallas, and D. I. Fotiadis, "A Kalman-filter-based methodology for EEG spike enhancement," *Comput. Methods Programs Biomed.*,

vol. 85, pp. 101–108, 2007.

19. M. D. Lucia, J. Fritschy, P. Dayan, and D. Holder, “A novel method for automated classification of epileptiform activity in the human electroencephalogram-based on independent component analysis,” *Med. Biol. Eng. Comput.*, vol. 46, pp. 263–272, 2008.
20. Z. H. Inan and M. Kuntalp, “A study on fuzzy c-means clustering-based systems in automatic spike detection,” *Comput. Biol. Med.*, vol. 37, pp. 1160–1166, 2007.
21. Chunk Anderson, “EEG Pattern Analysis – Sonification”, November 3, 2011, <http://www.cs.colostate.edu/eeg>
22. Giovanni Marco Zaccaria, Sergi Jordà Puig, Sebastià Mealla Cincuegrani, “Sonification of EEG signals – A study on alpha band instantaneous coherence”, 2011
23. Baier, Gerold, and Thomas Hermann, “The Sonification of Rhythms in Human Electroencephalogram,” In ICAD, 2004.
24. Vialatte, François B., Justin Dauwels, Toshimitsu Musha, and Andrzej Cichocki. “Audio representations of multi-channel EEG: a new tool for diagnosis of brain disorders,” *American Journal of Neurodegenerative Disease*, vol. 1, no. 3, pp. 292, 2012.
25. Temko, W. Marnane, G. Boylan, and G. Lightbody, “Clinical implementation of a neonatal seizure detection algorithm,” *Decision Support Systems*, vol. 70, pp. 86–96, 2015.
26. Epileptic EEG Visualization and Sonification Based on Linear Discriminate Analysis Wei Chen, Chia-Ping Shen, Ming-Jang Chiu, Qibin Zhao, Andrzej Cichocki, Jeng-Wei Lin, *Member*,
27. Z. Wu, and Norden E. Huang, “Ensemble empirical mode decomposition: a noise-assisted data analysis method,” *Advances in Adaptive Data Analysis*, vol. 1, no. 1,

pp. 1-41, 2009.

28. <http://rcada.ncu.edu.tw/research1.htm> 中央大學數據中心

29. Schmitt, Lothar M (2004), 遺傳算法理論 (二), Theoretical Computer Science (310), pp. 181-231

